apresentação

Paulo Nordeste
Presidente da Comissão Executiva da
PT Invovação

A revista "Saber e Fazer" é um instrumento de partilha do conhecimento criado pela PT Inovação, nomeadamente nas suas actividades de Investigação e Desenvolvimento.

Fazendo parte da nossa missão, esta partilha constitui um modesto contributo para a formação e actualização das pessoas que trabalham no Grupo PT.

São apresentadas tecnologias, soluções, projectos e experiências levadas a cabo não só pela PT Inovação, mas também por elementos das diferentes empresas com quem trabalhamos e aprendemos diariamente.

A preparação e escrita dos diferentes artigos requer, por parte dos autores, um esforço adicional às actividades quotidianas que importa realçar.

Para todos os que contribuíram para mais esta edição da "Saber e Fazer", vai o nosso reconhecido agradecimento.

01

Da Garantia da Qualidade à Garantia do Negócio

Alcino Lavrador (INB), Edson Takao Kobashi (INB), Fernando Bastos, Filipe Tavares, Gouveia Leal, Hélder Alves, João Pias, Luís Cortesão

palavras-chave:
Garantia do Negócio, Garantia de Receita, Qualidade de Serviço, Indicadores de Desempenho, SLA, Rentabilidade do Negócio e dos Recursos, Controlo de Fraude, Gestão e Registo de Falhas e Problemas, Gestão Integrada

A recolha e processamento de indicadores de desempenho, CDR e IPDR provenientes da rede, bem como de indicadores – quer dos sistemas de informação, quer dos sistemas operacionais – permitem uma gestão *end-to-end* da QoS e SLA, possibilitando uma análise em tempo útil da utilização e respectivo grau de prestação de determinado serviço. Complementarmente, análises de Garantia de Receita e de Gestão de Fraude, permitem uma visão global do Estado do Negócio. É da interligação de sistemas que actuam em ambas as áreas aqui referidas que se conseguem retirar indicadores precisos para a gestão da estratégia do negócio. Neste artigo é apresentada uma abordagem da PT Inovação para a disponibilização duma solução modular e flexível, que permite aos operadores de redes de telecomunicações dispor de uma visão integrada do negócio e dos factores que qualitativa e quantitativamente contribuem em cada instante para a qualidade percebida dos serviços oferecidos, bem como para a defesa do próprio negócio.

1. Introdução

A Garantia do Negócio duma empresa, tem na capacidade contínua de medir a sua evolução contra metas previamente estabelecidas, um dos factores mais fortes de sucesso. Estas metas representam normalmente o aumento de quota de mercado, maximização do lucro e/ou redução de custos. Torna-se possível alargar a carteira de clientes através da aposta em vantagens competitivas, tais como o preço de determinado bem ou serviço, a respectiva diferenciação face à concorrência, o nível de qualidade exibido, a sua universalização (acesso a partir de qualquer lugar), entre outras.

O primeiro passo para aumentar a base de clientes passa por cativar e manter os actuais. Para isso, se é por um lado imprescindível conhecê-los através da utilização dos serviços oferecidos, é por outro necessário conhecer a prioridade e importância que estes dão às características dos mesmos. Dão mais valor ao detalhe e transparência da factura, à disponibilidade, à qualidade, ou ao preço? A partir deste conhecimento, torna-se possível formatar ofertas diferenciadas, que vão ao encontro dos diversos perfis de cliente. Para conciliar este propósito com a sempre necessária racionalização de custos, torna-se fundamental dotar a empresa das ferramentas e indicadores que permitam gerir:

a) Cumprimento contratual:

> Medir em tempo real a qualidade e a prontidão do serviço prestado, disponibilizando alertas para desvios face aos objectivos;

b) Prestação do serviço:

> Medir a eficiência na utilização de recursos básicos, permitindo a sua optimização dinâmica conforme a procura, providenciando deste modo uma melhoria operacional;

> Gerir de forma rápida alterações de configuração resultantes de falhas, nível de utilização, ou mudanças tecnológicas;

> Detectar inconsistências internas nas informações suportadas pelos diversos componentes, na prestação de um serviço.

Nos dias de hoje, não se poderá ignorar os papéis de fornecedores e/ou parceiros: a correcta remuneração destes na partilha do negócio exige, de forma semelhante aos clientes, ferramentas e indicadores que possam medir a quantidade e qualidade do serviço prestado.

Outra vertente na redução de custos passa pela prevenção, detecção e eliminação da utilização fraudulenta de serviços. Quase por certo que não haverá nenhum outro negócio tão sujeito a fraude como seja o dos operadores de telecomunicações. O seu enquadramento de modelo de negócio, em constante mutação, aliado à complexa e evolutiva envolvente tecnológica, bem como à globalização do seu uso, permite a existência de brechas cujo tempo de vida, por muito curto que seja, acaba por resultar em perdas de vulto. Por este motivo a Gestão de Fraude aliada à componente de Garantia de Receita representam um papel fundamental na Garantia do Negócio.

De notar que a Garantia de Receita tem evoluído duma posição relativamente reactiva, em que o objectivo seria a minimização de perdas, para uma posição mais abrangente e proactiva, centrada na gestão de custos (minimização de custos e maximização de margens), entrando na esfera da gestão de produto e de cliente.

A figura 1 ilustra aqueles que constituem os principais vectores de negócio na perspectiva dos dias de hoje:

> Cumprimento contratual com clientes e parceiros;

> Garantia de receita;

> Detecção de fraude;

> Prestação de serviço.

Em conclusão, poder-se-á afirmar que só a monitorização em tempo útil dos diversos aspectos supracitados, possibilitará nos dias de hoje uma maior certeza na Garantia do Negócio, preparando a organização para o seu crescimento racional e sustentado.

2. Agentes da Garantia do Negócio

É de uso comum afirmar que a qualidade tem o seu preço. Qualquer bem transaccionável pode ter na sua qualidade um factor diferenciador, tornando-o mais apetecível, mas eventualmente também mais caro para o cliente final. A existência de mecanismos para verificar se os serviços que estão a ser fornecidos ao cliente se encontram em conformidade com os SLA contratados, providenciando evidências do cumprimento dos mesmos, constituem também uma condição imprescindível.

Nas telecomunicações, esta é também uma regra: quanto melhor o serviço prestado, maior valorização irá ter por parte do cliente, o qual estará mais facilmente disposto a pagar pelo mesmo. No entanto, para um operador de telecomunicações, a qualidade é não só a perceptível pelo cliente final, mas também a que se traduz na eficácia com que se utilizam os respectivos recursos: basta ver o exemplo de uma rede IP, onde a introdução de mecanismos de controlo da QoS só se justifica porque é necessário prioritizar o tráfego, de forma a garantir que os serviços com maior exigência de QoS (classe de serviço *Gold*) têm prioridade no acesso aos recursos da referida rede.

É pois neste equilíbrio, entre a qualidade do serviço a fornecer e os recursos estritamente necessários para o desempenhar, que se encontra a optimização do negócio. Para que um operador se mantenha nesta gestão equilibrada, é necessário ter uma noção exacta, e tanto quanto possível em "tempo útil", da qualidade de serviço prestado e do respectivo grau de eficácia dos recursos usados. Este objectivo só é possível através de uma supervisão de processos, implementada a partir de uma arquitectura de sistemas que abarquem a monitoria e controlo de toda a cadeia de valor. Essa arquitectura deverá ser suportada em tecnologias que permitam agir de forma dinâmica e em tempo real sobre os recursos utilizados, em função dos resultados obtidos.

As secções que se seguem descrevem os vários componentes que fazem parte do "*puzzle*" de agentes necessários à garantia do negócio.

2.1 A Gestão da Qualidade de Serviço

A gestão da QoS surge da necessidade de alinhamento dos objectivos de negócio com os objectivos dos clientes. As telecomunicações não são excepção: há que amortizar os investimentos realizados no mais curto tempo possível, rentabilizando ao máximo os recursos da rede, ao mesmo tempo que têm de se salvaguardar os SLA celebrados com os clientes. O sucesso deste equilíbrio tornou-se possível graças à utilização de Sistemas para Gestão da QoS.

Os operadores de telecomunicações têm actualmente ao seu dispor um conjunto de mecanismos que lhes permitem ter uma visão *customer centric* da QoS, em tempo real, com possibilidade de *drill down* desde o cliente até aos recursos da

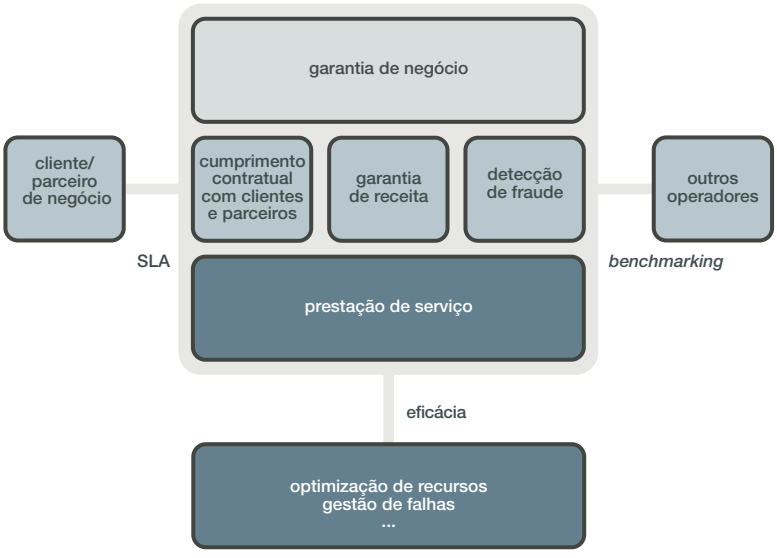

figura 1 - Vectores de negócio

rede que suportam as respectivas comunicações, quer com acesso imediato aos respectivos indicadores de QoS e alarmes, quer recorrendo ao respectivo histórico. A validação de SLA por cliente, despoletando notificações sempre que os indicadores da QoS indiciem uma degradação do serviço, associada à provisão de informação financeira em tempo real, quantificadora da penalização por situações de incumprimento - muito útil na definição das prioridades inerentes aos processos de reposição de falhas - constituem funcionalidades imprescindíveis para a gestão de hoje em dia. Paralelamente, é também relevante a capacidade de validação de SLA celebrados com parceiros de negócio, como será o caso dos fornecedores de alguns dos serviços providenciados pelo operador, onde a ausência de processos de certificação constituiria um elevado risco para a sua própria imagem.

O assegurar de um serviço de qualidade, complementado com um eficiente processo de detecção e resolução de anomalias, permite a disponibilização de relatórios de SLA eficazes, que irão funcionar como promoção periódica da imagem de marca do operador, permitindo não só prevenir o *churn* mas também facilitar o incremento da carteira de clientes.

Deste modo, a implementação de arquitecturas de rede "assessoradas" por sistemas de informação orientados à Gestão da QoS, flexíveis, abrangentes e compatíveis com as permanentes evoluções tecnológicas, constitui um dos aspectos mais relevantes a ter em conta no actual negócio das telecomunicações.

Sendo esta a actual realidade neste sector, fortemente dinâmico e onde a concorrência é cada vez mais agressiva, o não acautelar do factor QoS traduzir-se-á numa inevitável degradação de imagem, que levará ao consequente comprometimento do próprio negócio.

2.2 A Garantia de Receita

A Garantia de Receita traduz-se na identificação e prevenção de problemas que resultam em perdas financeiras nas cadeias de valor de um operador de telecomunicações. Imagine-se uma rede de condutas de água, onde em cada conduta existem gotas que se perdem e que são assim desperdiçadas ao longo da cadeia de valor. A Garantia de Receita corresponde às acções para detecção das referidas fugas, correcção dos mecanismos que as provocam e prevenção de situações semelhantes no futuro.

A Garantia de Receita é um tópico de grande relevância em telecomunicações, devido essencialmente a três factores:

> A velocidade da mudança tecnológica e a intensa concorrência no sector, aumentam a probabilidade de erro. Esta velocidade implica a coexistência de sistemas novos com sistemas legados, suportando o mesmo serviço percebido pelo cliente;

> O número de sistemas e processos envolvidos é muito elevado, tornando muito difícil prever o impacto de alterações, mesmo que pequenas;

> O grande volume de pequenas transacções amplifica as repercussões financeiras de pequenos erros.

Os valores estimados para a Perda de Receita em telecomunicações variam entre 2% e 10%, tendo subjacentes as seguintes causas:

> CDR corrompidos ou perdidos;

> Erros nos processos de tarifação;

> Inconsistências entre as bases de dados de suporte aos serviços e ao negócio.

Nos últimos anos verificou-se um aumento na oferta de ferramentas que permitem detectar e prevenir situações de perda de receita. As ferramentas existentes incluem componentes de *hardware* e *software* para:

> Estimular a rede ou sistemas de informação com tráfego ou dados de entrada, e verificar se os resultados são os esperados;

> Fazer verificações maciças ou estatísticas nas bases de dados de aprovisionamento, e verificar um conjunto de regras de consistência;

> Copiar processos de negócio e verificar se os resultados da cópia são iguais aos do original. Um exemplo é a aplicação de motores de tarifação distintos aos mesmos dados de chamada, pretendendo-se obter os mesmos valores para o custo.

2.3 A Detecção de Fraude

A fraude em telecomunicações – que segundo várias organizações internacionais, pode afectar 3 a 6% da receita líquida dos operadores – é uma questão dinâmica, complexa e transversal à estrutura do operador, assumindo distintas vertentes (subscrição, técnica, interna, agentes, entre outros) em constante mutação. O sucesso do combate à fraude depende, em larga medida, da qualidade e experiência da equipa de analistas, e beneficia claramente da integração de um Sistema de Gestão de Fraude (FMS – *Fraud Mangement System*) adaptado aos processos internos e suficientemente flexível para cobrir as múltiplas vertentes de fraude.

A permanente evolução tecnológica constitui um factor de preocupação, quer pela possibilidade de continuamente poder originar novas técnicas de fraude, quer por exigir a

análise contínua de fontes de informação adicionais. A qualidade de um FMS pode ser aferida pela facilidade de incorporar novas fontes de informação, assim como pelo grau de flexibilidade dos processos de detecção para se adaptarem aos distintos tipos de fraude.

A possibilidade de utilização isolada ou conjugada de diferentes técnicas de detecção - regras, técnicas estatísticas, técnicas baseadas em IA, entre outras - pode constituir um factor diferenciador do FMS, por aumentar a capacidade de detectar fraudes emergentes.

Outras questões chave a considerar na adopção de um FMS:

> O elevado e crescente volume de informação (tráfego, *logs* de SI, entre outros) a processar exigir uma solução facilmente escalável;

> O típico processo de investigação exigir o acesso a informação dispersa em múltiplos tipos de informação;

> A profusão de situações passíveis de análise, justificar a existência de mecanismos capazes de orientar os analistas para as situações de maior risco de impacto financeiro;

> A fidedigna replicação dos distintos serviços e planos comerciais do operador exigir mecanismos fiáveis e flexíveis de valorização de tráfego.

O retorno de um FMS depende da abrangência dos sistemas analisados e das vertentes de fraude a combater.

2.4 A Gestão e registo de falhas e problemas

As ferramentas para gestão e registo de falhas e problemas, são instrumentos fundamentais à Garantia do Negócio, nomeadamente, pela eficácia que imprimem aos processos de notificação dos operacionais responsáveis pela resolução de falhas, com definição de tempos limite para a resolução das mesmas, garantindo ainda mecanismos para confirmação ou rejeição de cada notificação. Como resultado, é garantido o controlo e a monitoria em todas as fases e tarefas associadas, sendo possível conhecer a cada instante o estado de cada notificação ou eventual reparação. Quando o tempo de resolução está a atingir os limiares predefinidos para cada fase, torna-se possível notificar os responsáveis com um alarme, ao mesmo tempo que se monitorizam os SLA acordados com os clientes.

Este tipo de ferramentas funciona também como complemento à avaliação do desempenho das equipas de operação, permitindo avaliar em cada instante se a empresa está a reagir em tempo útil e de acordo com as expectativas do negócio e dos clientes. Permite ter ainda um registo único de toda a informação que circula na empresa sobre falhas e problemas, mantendo o seu histórico, e facultando indicadores com especial interesse para a gestão, permitindo a melhoria contínua dos processos associados.

3. A Inovação na Perspectiva do Negócio

A PT Inovação, em regra, transpõe as actuais e futuras necessidades dos seus clientes – Empresas do Grupo PT – para os requisitos de concepção de cada um dos seus produtos. Esta é uma realidade que torna único o portfolio PTIN, ainda mais quando estes produtos se articulam também entre si, numa perspectiva transversal a todo o negócio, formando um "*puzzle*" tecnológico (figura 2).

Os sistemas acima identificados, que têm já implementações consolidadas em vários clientes, constituem-se também como parte integrante duma solução global que monitoriza toda a cadeia do negócio

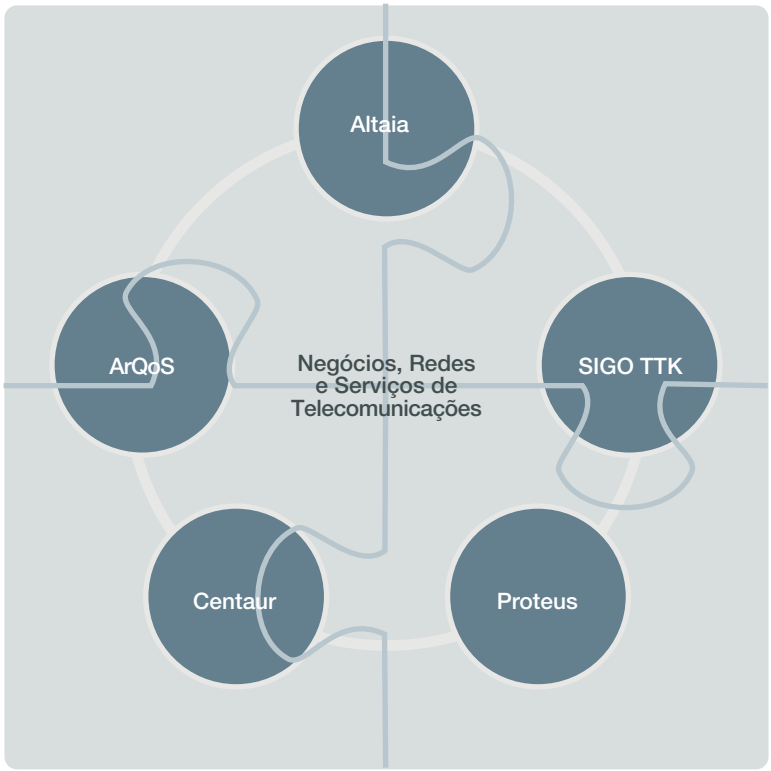

figura 2 - *Puzzle* Tecnológico para a Garantia do Negócio

de um operador de telecomunicações. Esta é uma realidade que permite aos operadores de redes de telecomunicações dispor de uma visão integrada do respectivo negócio, assim como dos factores que qualitativa e quantitativamente contribuem em cada instante para a qualidade percebida dos serviços oferecidos, tendo como missão a permanente defesa do próprio negócio.

Em termos de enquadramento funcional, os sistemas integrantes deste "puzzle" tecnológico podem ser contextualizados ao nível do *Telecom Application Map* (TMFC3875) – matriz que correlaciona os vários *layers* de negócio com as diversas perspectivas funcionais.

Foi até agora referido que para garantir o negócio se torna indispensável apostar na qualidade do serviço como factor diferenciador, ao mesmo tempo que se minimizam os custos, garantindo a plena utilização dos activos, e evitam as perdas de receita, erros de facturação, fraude, *churn*, etc. Foram também mencionadas as principais características e funcionalidades de cada um dos sistemas indicados na figura 3, já sumarizadas no capítulo 2. O presente capítulo pretende descrever as principais vantagens que advêm das sinergias resultantes da articulação entre os diferentes sistemas integrantes do *portfolio* PTIN.

3.1 ArQoS (*Service Quality Monitoring*)

O ArQoS é o sistema de *probing* da PT Inovação. Permite o acesso, captura e/ou simulação do tráfego que transita nas redes de telecomunicações, com o principal objectivo de disponibilizar KPI e KQI relativos ao negócio, serviço e/ ou rede (próprios ou de operadores concorrentes – *benchmarking*), enviando notificações automáticas nos casos de violação de limiares pré-definidos. O ArQoS constitui-se também como uma fonte privilegiada de informação para o Altaia, Proteus, Centaur e SIGO TTK.

Com a utilização de dois tipos de *probes* sob a mesma aplicação de gestão, o sistema combina as vantagens do *probing* intrusivo com o *probing* não intrusivo. As *probes* intrusivas suportam equipamentos terminais compatíveis com as principais tecnologias de comunicação fixa e móvel, bem como clientes de voz sobre IP e acessos à Internet de banda larga. A integração de várias destas tecnologias no ArQoS permite testar a interoperabilidade entre as mesmas. Estas *probes* simulam a actividade típica de um cliente, e registam a qualidade de serviço do ponto de vista do mesmo. Assim, o operador de telecomunicações dispõe de uma rede de clientes virtuais instalados estrategicamente pela área de cobertura da rede, podendo utilizá-los para uma constante utilização e monitorização da qualidade de serviço. O sistema notifica imediatamente o operador na ocorrência de eventuais problemas ou degradação de qualidade, face a limiares pré-definidos, antecipando-se às reclamações e insatisfação dos clientes reais.

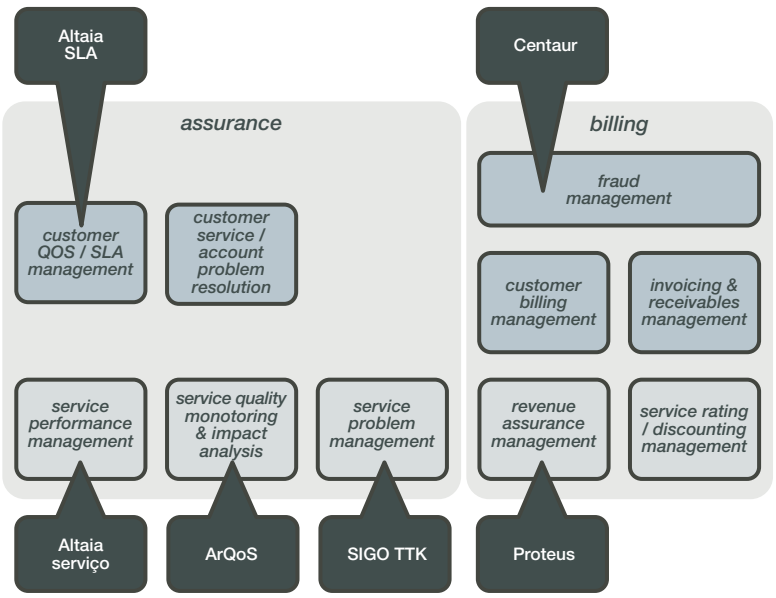

figura 3 - Enquadramento no *Telecom Application Map*

As *probes* não intrusivas ligam-se directamente à rede *core*. São equipadas com um *Network Processor*, o que lhes permite processar os elevados débitos das interfaces *Gigabit Ethernet*. Distribuídas por vários pontos-chave da rede, as *probes* capturam o tráfego das chamadas nas suas diferentes fases, sinalização e dados, e em vários pontos de passagem. Os dados são posteriormente recolhidos e processados numa unidade central onde é feita uma análise global do processo de cada chamada. Em caso de falha é possível identificar em que fase e em que troço da rede ocorreu. Com vários níveis de detalhe e filtragens disponíveis, o sistema pode registar desde estatísticas gerais de todas as chamadas, até capturar e gravar pacotes previamente seleccionados para uma análise mais detalhada.

Os dois tipos de *probes* podem trabalhar em conjunto. As *probes* intrusivas geram actividade que é analisada do lado da rede pelas *probes* não intrusivas. Os indicadores de qualidade recolhidos dos diferentes pontos de vista podem depois ser correlacionados. As duas informações complementam-se.

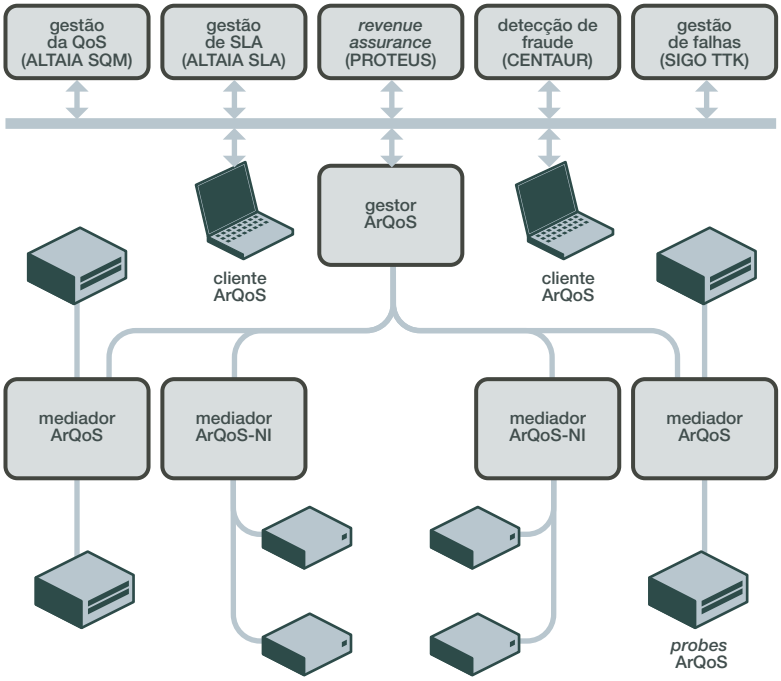

figura 4 - Arquitectura ArQoS

Enquanto que o *probing* não intrusivo permite identificar e localizar a origem das falhas, o *probing* intrusivo permite analisar o impacto real das mesmas no serviço prestado ao cliente.

O Gestor ArQoS é constituído por uma aplicação modular e escalonável, compatível com uma arquitectura de processamento distribuído por uma ou várias máquinas em função da carga do sistema, i.e. análises simultâneas (figura 4).

3.2 ALTAIA (SQM, SLA)

O Altaia é o sistema de gestão de Qualidade de Serviço da PT Inovação. Tendo por base uma plataforma modular e flexível baseada em J2EE e usando API normalizadas de gestão do TMForum e adaptadas para Java pelo consórcio OSS/J, tem a capacidade de se integrar com facilidade com os sistemas de rede e de serviços, obtendo todos os indicadores necessários para uma análise *end-to-end* dos serviços.

O Altaia SQM permite a visão integrada da QoS *end-to-end*, possibilitando uma análise correlacionada dos indicadores, quer provenientes directamente da rede, quer do sistema ArQoS. Do ArQoS obtém novos indicadores de serviço, complementares aos indicadores da rede.

O Altaia SLA permite a gestão de SLA de cliente, monitorando o seu cumprimento face ao contratado (figura 5) . Esta monitoria assenta em dados provenientes do Altaia SQM, ArQoS, Proteus. Do Altaia SQM recebe a QoS de cada serviço contratado pelos clientes; do ArQoS obtém informação relativa à utilização do serviço por cliente e a respectiva qualidade perceptível pelo cliente; do Proteus tem toda a indicação relativamente a dados de receita por cliente e eventuais problemas de facturação, com impacto no cumprimento dos SLA.

figura 5 – Gestão de QoS no Altaia

O seu papel central numa arquitectura de gestão com o objectivo de assegurar a garantia do negócio passa, numa primeira fase, pela captura e análise em tempo-real de indicadores de desempenho de rede e de indicadores complementares provenientes do ArQoS. Esta visão *on-line*, integrada e consolidada do estado do serviço, designada na arquitectura eTOM de SQM (Gestão da Qualidade do Serviço), constitui um dos pilares de construção do "*puzzle*" de gestão das operações.

A partir destes indicadores de serviço e usando dados de cadastro provenientes da integração com sistemas de provisão de SLA, o passo seguinte do Altaia é efectuar a monitoria e vigilância dos níveis de contratação dos clientes, detectando e alarmando situações de degradação que poderão degenerar na violação de limiares contratuais. Este processo é auxiliado na sua actividade a partir da obtenção de indicadores de utilização provenientes do ArQoS. Complementarmente, obtém do Proteus informação sobre possíveis problemas de facturação que poderão ser relevantes para a gestão dos SLA.

O controlo apertado dos SLA, com dados em tempo-real proveniente da rede, permite ao operador monitorar os contratos e a qualidade fornecida de forma a garantir a melhor relação entre o serviço e a qualidade oferecida. **3.3 CENTAUR**

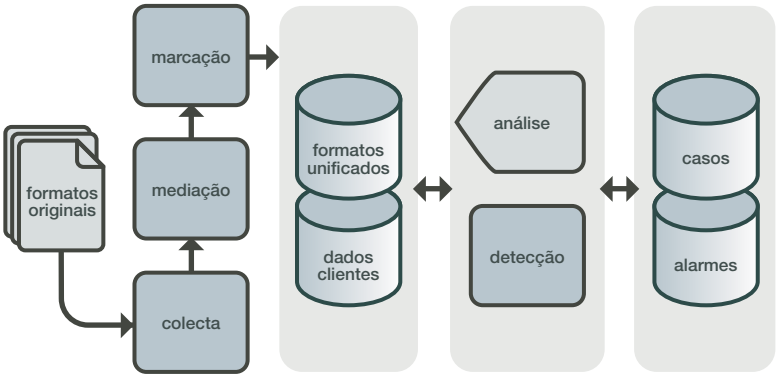

figura 6 – Fluxo de informação Centaur

(Gestão de Fraude)

O Centaur é o sistema de detecção de fraude da PT Inovação. Para o efeito, procede ao cruzamento de informação oriunda de diferentes pontos da rede e/ou sistemas de informação, posteriormente correlacionada com um conjunto de regras definidas pelo operador. O ArQoS poderá constituir uma das diversas fontes de informação que alimentam o Centaur, permitindo a configuração das respectivas *probes* não intrusivas para efectuar o *forward* selectivo de informação capturada ao longo dos diversos nós da rede.

O sistema de gestão de fraude Centaur foi concebido segundo uma arquitectura cliente-servidor, baseada em componentes para maior flexibilidade de configuração e reutilização de *software* figuraa 6).

Após a recolha e carregamento dos dados do operador considerados relevantes para as vertentes de fraude a combater (colecta), procede-se à conversão dessa informação para os formatos internos Centaur (mediação). Na fase de marcação, para além de se enriquecer a informação, cruzando com dados de diversas proveniências, efectua-se também a tarifação. O Centaur utiliza um motor de tarifação independente de elevada precisão, que permite reflectir na solução qualquer alteração da estratégia comercial do operador e desta forma garantir maior fiabilidade nos alarmes gerados.

Os processos de detecção Centaur, responsáveis pela geração de alarmes, têm a possibilidade de utilizar diversas fontes de informação em simultâneo. As opções avançadas de cruzamento de informação contribuem para uma redução significativa do número de falsos alarmes. Adicionalmente, a extrema flexibilidade na implementação de detectores baseados em diferentes técnicas permite a fácil adaptação a diferentes segmentos de clientes, distintos SI do operador, novas técnicas de fraude, assim como às novas tecnologias emergentes.

No Centaur, as tarefas de investigação dos alarmes gerados na fase de detecção (análise) são extremamente eficientes dada a disponibilização centralizada da informação normalmente dispersa pelos distintos SI do operador. A isso acresce a facilidade de inter-relação de acordo com os processos mentais dos analistas e a existência de mecanismos de prioritização multi-variável dos alarmes, que permitem orientar os analistas para os casos com maior risco e probabilidade de impacto financeiro.

3.4 PROTEUS (*Revenue Assurance*)

O Proteus é o sistema de *Revenue Assurance* da PT Inovação. Foi projectado para comparar informações distintas provenientes de diversas fontes de dados, denominadas "pontos de colecta", entre as quais se encontra o ArQoS. Por análise e confrontação de CDR gerados nos vários sistemas envolvidos, efectua a auditoria e validação a cada sistema assim como a toda a cadeia de geração de receita. Por outro lado a verificação dos dados de aprovisionamento entre bases de dados operacionais permite detectar inconsistências que resultam em perdas de receita.

O Proteus permite ao utilizador acompanhar a trajectória e detectar inconsistências nos serviços utilizados na rede do operador em questão, nas tecnologias UMTS, GSM/GPRS, TDMA, CDMA/1x, nas plataformas de pré-pago e pós-pago (figura 7).

A análise de CDR e dados de aprovisionamento de forma controlada divide-se em 3 macro processos:

1 **Extracção:** A ferramenta permite ao utilizador configurar regras para extracção de dados, definindo, por exemplo, o horário da chamada, o número chamador, o número chamado, a duração, entre outros;

figura 7 – Menu principal do Proteus

2 **Correlação:** Este processo permite identificar os CDR e dados de aprovisionamento entre os diferentes pontos de colecta e definir critérios para batimento e uniformização dos mesmos;

3 **Confrontação:** Através das regras de confrontação, o utilizador pode criar regras que quantifiquem CDR com falha, falta de CDR, inconsistências nas bases de dados operacionais ou outros cenários que atendam a necessidade da análise a ser feita.

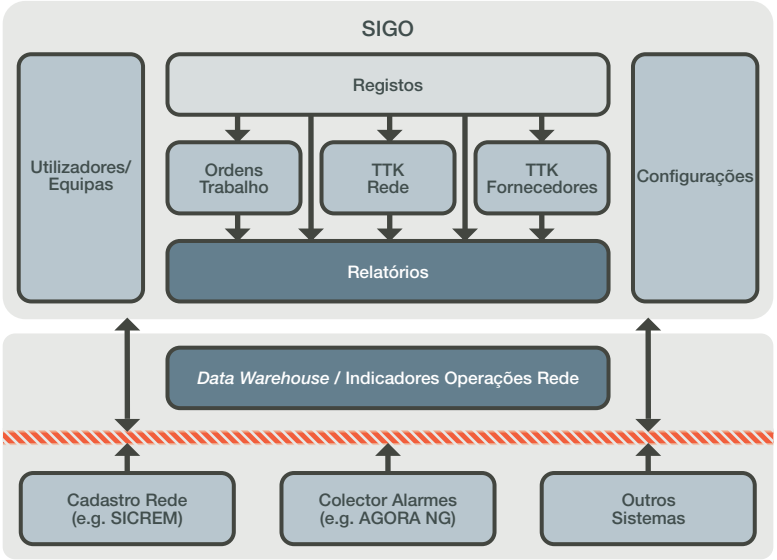

figura 8 – Plataforma SIGO

3.5 SIGO TTK (Gestão de falhas)

O SIGO TTK é o sistema de gestão de falhas da PT Inovação, sendo responsável pelo registo e controlo da execução de todas as actividades envolvidas no processo de gestão de avarias da rede. Implementa todo o conjunto de funcionalidades mencionado no capítulo 2.4, contemplando várias fases e regras associadas ao tratamento de avarias figuraa 8).

Este módulo permite, entre outras opções:

> Tipificação de avarias;

> Criação e registo de avarias, assim como associação de ficheiros de anexo;

> Atribuição directa da avaria à entidade responsável pela sua resolução;

> Relacionamento de avarias;

> Definição do tipo de intervenção, prioridade e tipo de corte na rede;

> Gestão global e consulta do estado da avaria;

> Consulta flexibilizada do conteúdo das avarias, sua exportação e impressão;

> Controlo dos prazos da resolução da avaria, nomeadamente, gestão de SLA;

> Escalonamento e notificação;

> Controlo dos prazos para execução de tarefas;

> Associação de avarias a registos de participação, indisponibilidade e intervenção.

Ao registar e ao controlar o estado de resolução de todas as avarias (TTK), este sistema oferece aos utilizadores envolvidos o controlo integral das actividades constituintes. Permite ainda a gestão de serviços de suporte e de reparação/ substituição de unidades pelos fornecedores, possuindo mecanismos de registo e encaminhamento de *tickets* para os mesmos, também considerados FR (*Fault-Report*).

4. Conclusão

Face à constante evolução do negócio das telecomunicações e à forte concorrência no sector, obter o máximo de rentabilidade dos recursos disponíveis, garantindo a qualidade aos clientes, constitui um factor crítico de sucesso para qualquer operador. Ter uma gestão integrada das plataformas e dos serviços, recolher, processar a analisar indicadores de funcionamento dos sistemas e em paralelo comparar dados dos sistemas operacionais com os de negócio para avaliar da consistência da informação, são factores determinantes para ter sucesso.

A PT Inovação posiciona-se, neste contexto, como um fornecedor privilegiado, propondo um "*puzzle*" tecnológico, composto de um conjunto integrado de sistemas, que facilitam a implementação dos agentes necessários para a garantia do negócio dos operadores.

Referências
TMFC3875: Telecom Application Map

Alcino Lavrador, licenciado em Engenharia Electrotécnica pela Faculdade de Ciências e Tecnologia da Universidade de Coimbra em 1985. Director Executivo da PT Inovação Brasil desde Março de 2003, ingressou no Centro de Estudos de Telecomunicações/PT Inovação em 1985 tendo exercido funções no desenvolvimento e coordenação de projectos nas áreas de protocolos de sinalização Nº 7 e Redes Inteligentes. Coordenou grupos de desenvolvimento de serviços e de implementação de diversos projectos de soluções de plataformas de Redes Inteligentes multi-serviços tendo sido director de Integração de Sistemas e Serviços na PT Inovação de Janeiro de 2002 a Março de 2003.

Edson Takao Kobashi, bacharel em Matemática Aplicada pela Universidade de São Paulo. Atua na área de *Business Intelligence* desde 1999 onde já participou de vários projetos em diferentes segmentos de mercado tais como varejo, banco e telecomunicações. Atualmente é o responsável pela área de Suporte à Decisão e Inteligência no Negócio da PT Inovação Brasil.

Fernando Bastos, licenciado em Eng. Electrotécnica, pela Universidade de Coimbra em 1984, e MSC em Telecomunicações pela Universidade de Aveiro em 2001. Trabalha na PT Inovação desde 1985, e é actualmente responsável pela Unidade Funcional "Plataformas OSS". Coordena as equipas de desenvolvimento dos sistemas Altaia (Sistema de Gestão de QoS e SLA) e Centaur (Sistema de Gestão de Fraude). Tem uma vasta experiência na área de sistemas de gestão, tendo participado em grupos de normalização europeus e em projectos de desenvolvimento nacionais e internacionais.

Filipe Tavares, licenciado em Engenharia Electrotécnica e de Computadores pela Faculdade de Engenharia da Universidade do Porto. Iniciou a actividade profissional na PT Inovação em 2001 no grupo de Sistemas e Infra-estruturas de Rede, tendo participado no desenvolvimento de terminais de operação e manutenção de equipamentos de rede, bem como no desenvolvimento de soluções de qualidade de serviço.

Gouveia Leal, Engenheiro de Telecomunicações pela UC em 1983 e desde Janeiro de 1985 no grupo PT. Inicio de actividade em 1984 na Central Térmica da EDP em Setúbal, ingresso como projectista nos Telefones de Lisboa e Porto no grupo de Engenharia e Projecto de Redes da Direcção de Engenharia de Lisboa (DETU), passando pela Direcção Regional de Telecomunicações Centro dos CTT Telecomunicações em Coimbra, como responsável pela análise técnica e económica de projectos de telecomunicações e apoio à implementação de Sistemas de Suporte às Operações (OSS).

Desde 1993 no CET e posteriormente na PT Inovação, sempre na área de planeamento e projecto de Redes de Telecomunicações e Sistemas de Suporte e Gestão de Redes, acumulando entre 1988 e 2001 a responsabilidade de director do projecto SIGRA na Direcção Geral de Infraestruturas da PT Comunicações.

Desde 1999 responsável pela Área de Negócios de Sistemas de Suporte e Gestão de Redes e consultor da Agência de Inovação como avaliador externo de projectos de telecomunicações."

Hélder Alves, licenciatura em Eng. Electrónica e de Telecomunicações pela Universidade de Aveiro, trabalha na PTIN desde 1990, é actualmente responsável pela Unidade Funcional de Teste de Integração de Soluções de Rede. Coordena os processos de teste ao nível da gestão das soluções NetB@nd, implementando um dos importantes elos de articulação entre o Desenvolvimento e os Serviços de Engenharia. De 2002 a 2006, foi responsável pela Unidade Funcional de Gestão Local e Qualidade de Serviço, tendo coordenado equipas de desenvolvimento de produtos de QoS no âmbito do portfolio NetB@nd – projectos ArQoS, SCP-RDIS, entre outros – bem como de terminais Web e agentes SNMP *embedded*, no contexto da gestão do Elemento de Rede aplicado a todas as soluções NetB@nd. Possui uma vasta experiência de trabalho e liderança em projectos de desenvolvimento de sistemas, prospecção e consultoria.

João Pias, actualmente, e desde Fevereiro de 2002, responsável pela Unidade Sistemas de Suporte à Decisão, na área de negócios Integração de Sistemas e Serviços. Durante este período esteve envolvido em vários projectos dos quais se destacam:

> Plataforma NGIN (*Next Generation Intelligent Networks*): Responsável pelo desenvolvimento de *Business Intelligence Tools*, abarcando temas como Monitoria de Desempenho, Garantia de Receita e Gestão de Fraude.

> *Datawarehouse* de grandes dimensões do operador VIVO: Responsável pelo projecto, abarcando tarefas como definição e avaliação dos recursos humanos necessários, elaboração e faseamento do cronograma geral, relacionamento com o cliente e definição da arquitectura.

Luis Cortesão, licenciatura em Engenharia Informática pela Universidade de Coimbra. Como colaborador do CET e posteriormente da PT Inovação, desenvolveu trabalho nas áreas de sistemas de informação geográfica, usabilidade, gestão de competências e gestão do conhecimento. Actualmente, responsável pelo desenvolvimento de sistemas de gestão de fraude e envolvimento no desenvolvimento de soluções de *revenue assurance.*

02

Plataforma de Rede Inteligente para Operações Móveis Virtuais - o Caso UZO®

José Domingues
colaboraram igualmente para a edição deste artigo:
Jorge Pinto, José Bonnet, Nuno Honório, João Ribeiro da Silva, Paulo Sousa, José Maio, João Mendes (TMN), Jorge Frazão (TMN), Jorge Teixeira da Silva (TMN).

palavras-chave:
MVNO, UZO, NGIN (*Next Generation Intelligent Networks*), Rede Inteligente , Serviço Pré-Pago, UZO-NGIN CORE, UZO-NGIN CARE, UZO-NGIN PORTAL, UZO-NGIN BIT, UZO-GW SIBS, UZO-SIGO, UZO_IVRs.

Este artigo descreve, de forma sumária, o conjunto de plataformas, sistemas de informação e aplicações desenvolvidas pela PT Inovação para a UZO, bem como a sua integração com os processos e sistemas de informação envolventes do seu operador de rede hospedeiro, a TMN.

1. Introdução

Atenta ao fenómeno crescente do aparecimento de Operadores Móveis Virtuais (MVNO – *Mobile Virtual Network Operator*) na Europa, a TMN, empresa operadora de serviços móveis do Grupo Portugal Telecom, optou pela PT Inovação como seu parceiro tecnológico para o desenvolvimento, instalação e operação de toda uma plataforma de serviços, aplicações e sistemas de informação para viabilizar a entrada em produção, num prazo recorde, de uma nova Unidade de Negócios (autónoma), operando na lógica de um Operador Móvel Virtual – a UZO.

Ancorada na Plataforma de Rede Inteligente da PT Inovação (NGIN – *Next Generation Intelligent Networks*), foram desenvolvidas e(ou) adaptadas novas lógicas do serviço IN Pré-pago, bem como todo um conjunto de módulos adicionais da família NGIN que permitiram o desenvolvimento e a entrada em exploração, em menos de quatro meses, do primeiro Operador Móvel Virtual em Portugal – a UZO.

2. Conceitos

Antes de avançar com uma descrição mais exaustiva da plataforma de suporte à UZO – NGIN para suporte a MVNO e MVNE (*Mobile Virtual Network Enabler*) – apresentam-se alguns conceitos fundamentais para a sua completa compreensão.

2.1 MVNO

Um MVNO é, tipicamente, uma empresa sem infra-estruturas de comunicação (comutadores, rede, plataformas) que fornece serviços móveis sob uma "marca" própria, enfocando o seu negócio na relação directa com o cliente final e na oferta de diferentes propostas de valor, sejam elas baseadas em "preço", segmentação de mercado, diferenciação de serviços ou *bundles* de serviços.

Qualquer empresa pode tornar-se num MVNO, não necessitando de ser perita em comunicações móveis ou em operações de *back-office* para operar o negócio.

Actualmente encontramos empresas com operações MVNO tais como operadores de telecomunicações, indústria do entretenimento, cadeias de retalho, clubes de futebol, etc..

Os factores âncora para o sucesso de uma operação MVNO assentam em:

> Aluguer de infra-estrutura (rede) a um operador de rede "eficiente";

> *Outsourcing* de serviços a empresas facilitadoras (MVNE);

> Uma proposta de valor ancorada numa "marca" forte;

> Estratégia clara e bem definida (e.g. preço, segmentação de mercado, *portfolio* de serviços, *bundles*,…).

2.2 MVNE

Um MVNE é, essencialmente, uma empresa que fornece os serviços de *back-office* que os MVNO necessitam para a sua Operação. É um conceito abrangente para definir uma empresa que, dispondo de

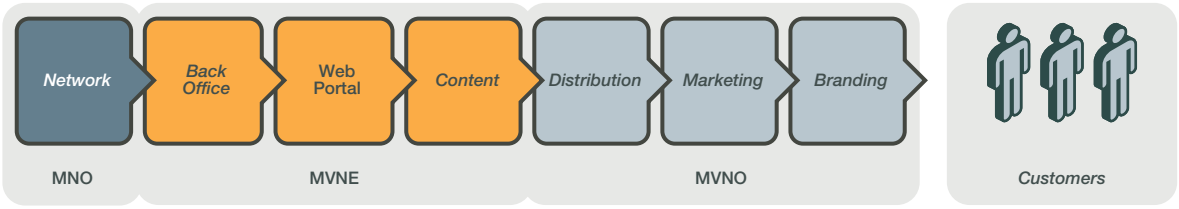

figura 1 – Cadeia de valor

know-how, plataformas de serviços e de sistemas de informação apropriados, pode fornecer qualquer tipo de serviço de *back-office* de suporte a vários ou à totalidade de processos, desde a gestão da interacção com o MNO (Operador Hospedeiro) até ao apoio ao MVNO na definição de terminais ou na aquisição de clientes.

O MVNE garante todo o Apoio Operacional ao MVNO, resolvendo e tratando de todos os processos complexos (e muitas vezes morosos), libertando o MVNO para desempenhar aquilo que melhor sabe fazer, e.g. – *branding*, *marketing* e aquisição/gestão de clientes.

Uma operação MVNO recorrendo a uma empresa MVNE apresenta vantagens significativas tais como:

> Oferece maior flexibilidade e controlo que um MVNO conseguiria *per si*;

> Aumenta a rapidez de lançamento de serviços de um MVNO;

> É mais tecnico/economicamente eficiente, reduzindo o investimento aos MVNO;

> Permite que os MVNO se concentrem no seu *core business.*

A figura 1 reflecte, em termos genéricos, a Cadeia de Valor do negócio móvel associada a MVNO, MVNE e MNO – fonte MVNGO.

3. Descrição da Plataforma NGIN/UZO

A criação da UZO obrigou ao desenvolvimento de um conjunto de serviços de rede inteligente, do tipo pré-pago móvel, a serem suportados numa plataforma NGIN da TMN, (NGIN/UZO), bem como por um conjunto de plataformas computacionais específicas para instalação de serviços adicionais que permitiram a criação de um novo *brand*, suportado pela rede da TMN. Pretendeu-se fornecer aos clientes um serviço pré-pago simples, controlado em tempo-real, com funcionalidades de auto-aprovisionamento, com esquemas de tarifação simples e facilmente percebidos pelo cliente, endereçando um segmento de clientes que paga essencialmente Voz e SMS e que privilegia os canais de contacto com o operador via IVR (*Interactive Voice Response*). O portal *Web* deveria permitir:

> Auto-aprovisionamento;

> Activação e desactivação automática de serviços – e.g.

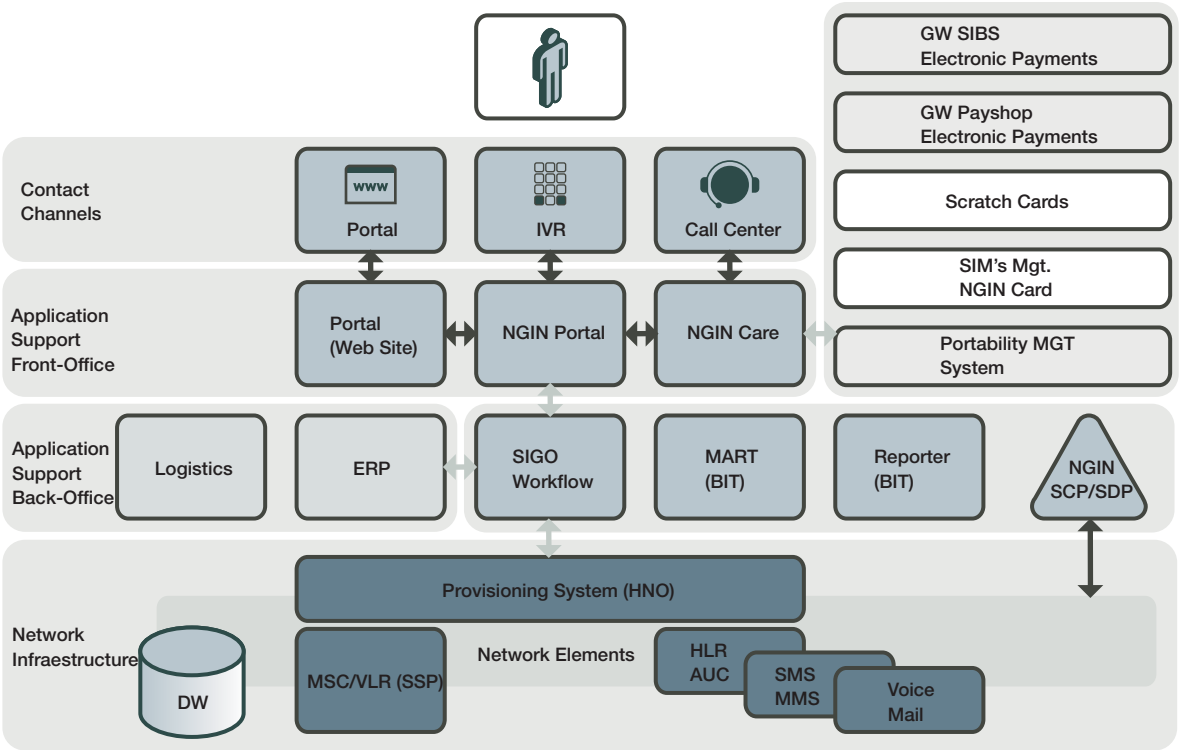

figura 2 – Solução UZO: Arquitectura funcional

Roaming, Saldo no Final de Chamada, *Voice Mail*, MMS...;

> Recargas;

> Consultas sobre utilização de serviços (consumos, débitos, créditos,…);

> Compras.

A PT Inovação actuou como integrador no projecto fornecendo *hardware*, *software* e serviços e participando também, nas equipas que integraram a solução com as redes e sistemas de informação da TMN.

Para o projecto UZO a plataforma baseia-se numa arquitectura funcional, composta por vários elementos figura 2, cada um deles com uma função específica, com alto nível de integração, optimizada para reduzir os custos operacionais do sistema.

O esquema ilustrado pretende demonstrar com um detalhe de alto nível, a actual arquitectura funcional da UZO.

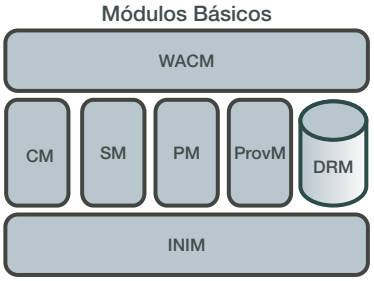

figura 3 – NGIN Care: Arquitectura Funcional

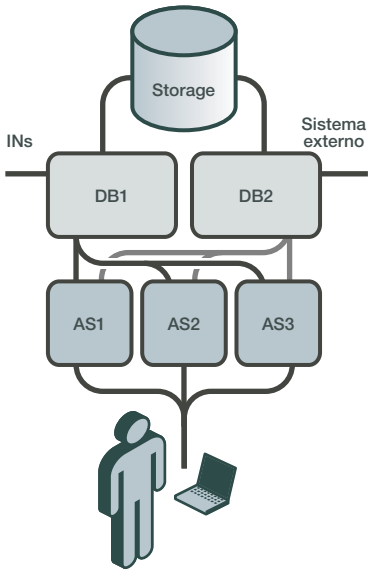

figura 4 – NGIN Care: Arquitectura física

3.1 NGIN *Core*

A plataforma base NGIN (NGIN *Core*) é o ambiente de execução da lógica de serviço(s). Através da interacção com a rede de telecomunicações (móvel), de um modo normalizado, é possível implementar variados serviços. No caso do UZO foi implementado o Serviço Pré-Pago UZO, controlando tráfego de voz, dados e eventos.

A plataforma pode coexistir com redes móveis GSM/GPRS e UMTS, bem como com redes *Wi-Fi.*

3.2 NGIN *Care*

O NGIN *Care* tem como principais funções o suporte ao *Customer Care*, do Aprovisionamento de Cartões e Serviços (incluindo plataformas externas), implementação das Regras de Negócio, Gestão da Portabilidade, Suporte ao *Front & Back-office*, Interfaces com plataformas de serviços (*Data Wharehouse*, MMS, SMS, Jogos,….) e com sistemas de informação envolventes e processos de suporte.

A arquitectura lógica deste componente é a representada na figura 3 e é constituída por:

> **INIM:** Interface com a(s) plataforma(s) IN;

> **CM:** Gestão de Clientes;

> **SM:** Gestão de Serviços;

> **PM:** Gestão de Pagamentos e Bónus;

> **ProvM:** Aprovisionamento;

> **WACM:** Controlo de Acessos *Web*;

> **DRM:** Registos Detalhados.

Do ponto de vista da arquitectura física, o NGIN *Care* é constituído por *storage* independente, bases de dados (DB) e *Application Servers* (AS), funcionando em partilha de carga e de forma totalmente redundante (figura 4).

Suporte ao Customer Care

> Informação do Cliente;

> Informação do Serviço/Cliente;

> Eventos *Inbound/outbound*;

> Operações: (mudança número, Mudança de Planos, Serviços, Créditos, Débitos,…);

> Subscrição de Serviços;

> Contas;

> Coexistência de contas de Grupo e Privadas);

> Movimentos detalhados, Conta-corrente, Saldos em múltiplos formatos HTML (*e-mail*, web, *finishing*,…)

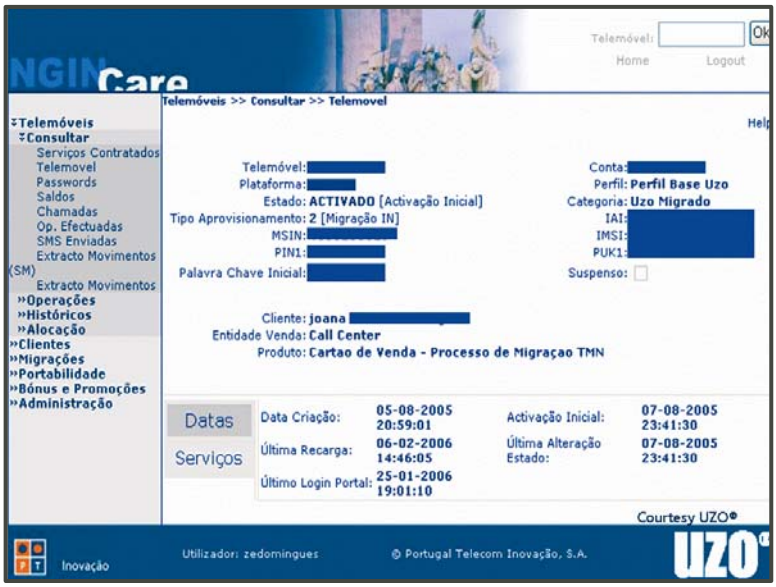

figura 5 – Aprovisionamento

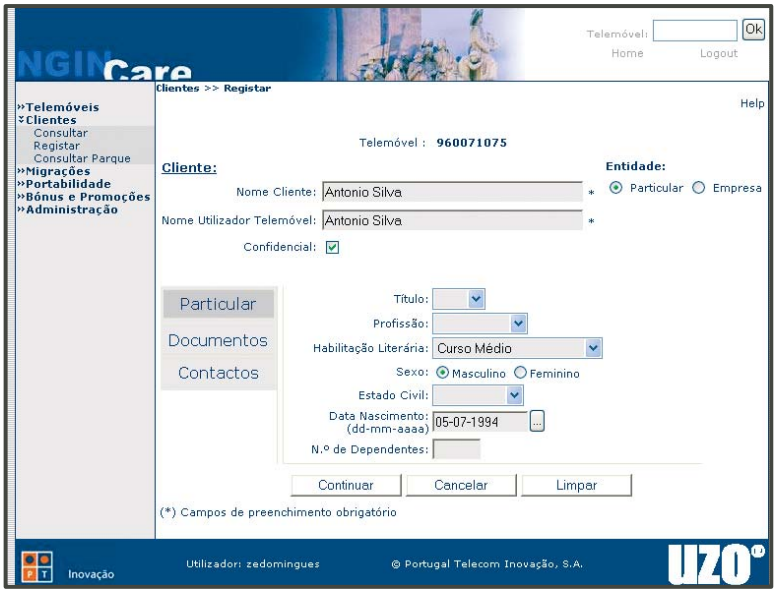

figura 6 – Registo de cliente

Aprovisionamento (figura 5)
O aprovisionamento de serviços implica uma interface (directa ou através de mediadores) com plataformas da rede tais como o HLR (*Home Location Register*), o WAP *Gateway*, o SMS-*Center*, o *Voice-Mail*, etc. Um número aprovisionado é activado com a primeira chamada (efectuada ou recebida) que passe na plataforma.

A contabilização das activações efectuadas e dos movimentos de recargas e chamadas são disponibilizadas em ficheiros diários para processamento pelos sistemas da empresa.

O aprovisionamento no NGIN *Care* possui as seguintes funcionalidades principais:

> Aprovisionamento de número avulso: com base numa lista de números disponíveis, é possível aprovisionar um número específico;

> Aprovisionamento de números em lote: lista de números para aprovisionar;

> Aprovisionamento de serviços: solicitações do cliente para a activação ou desactivação de serviços são encaminhados por este módulo;

> Validação de números roubados;

> Gestão do ciclo de vida do número.

Recargas
As diferentes formas de pagamento disponíveis são agregadas num único ponto de entrada, facilitando assim a respectiva manutenção e evolução. Após validação/confirmação de cada pagamento, é determinado o eventual bónus disponível e atribuído o crédito na plataforma IN em que o número reside, permitindo assim intervalos muito curtos entre a disponibilização do pagamento e a sua entrada na

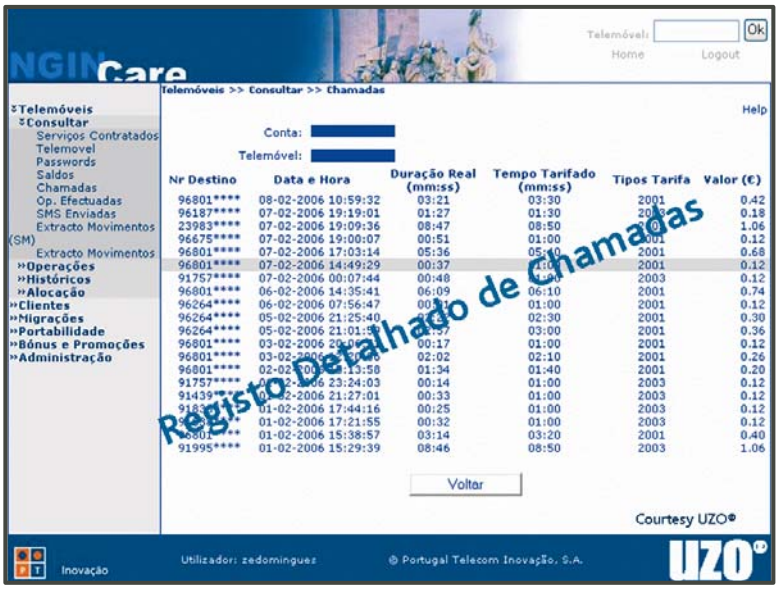

figura 7 – Extracto detalhado

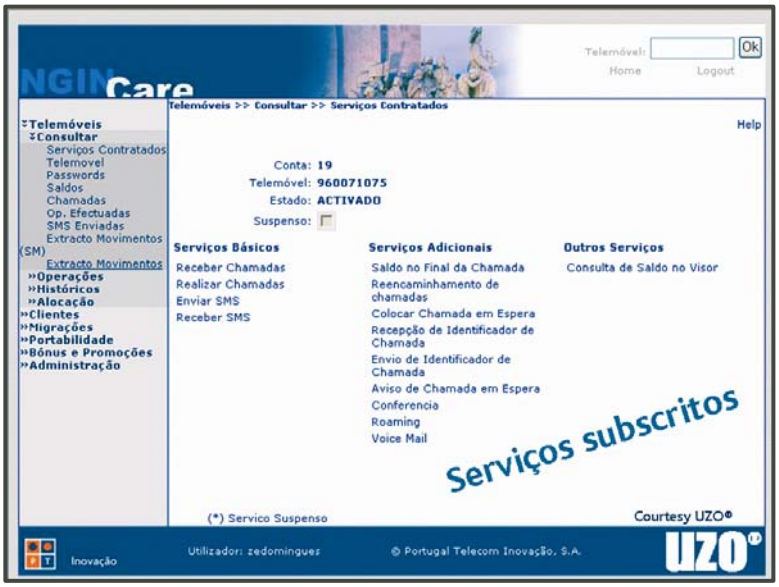

figura 8 – Serviços subscritos

conta do cliente, poucos segundos após a sua realização.

As recargas podem ainda ser transferidas entre contas, quer de forma informal (do telemóvel dum pai para o de um filho, por exemplo) quer de forma mais comercial, por exemplo, com uma rede de parceiros que adquire crédito grossista, que depois revende cobrando uma taxa.

Para utilizadores com o perfil adequado (tipicamente, supervisor de *back office*), é ainda possível efectuar ajustes em saldos de cliente, quer individualmente quer em lote.

A contabilização das recargas efectuadas é disponibilizada em ficheiros diários para processamento pelos sistemas de informação da empresa (e.g. NGIN *Mart*, *Datawarehouse*,...).

Clientes

Algumas das principais funcionalidades deste módulo:

> **Registo e dados de cliente (figura 6):** O registo dos dados de clientes, complementando informação que eventualmente já exista no sistema;

> **Extracto detalhado (figura 7):** O extracto detalhado (fonte XML) com possibilidade de emissão para os 30, 60 ou 90 dias anteriores à data do pedido) e envios para correio electrónico, impressoras de grande porte e para disponibilização no portal;

> **Gestão de contas de Grupos:** grupos de clientes podem associar-se em grupos, beneficiando de tarifas mais favoráveis nas chamadas entre os membros do grupo, com uma conta de grupo que pode ter o seu próprio perfil de tarifação, sem que cada um dos membros do grupo perca as suas características individuais de tarifação e serviços.

Ficam ainda disponíveis as seguintes funcionalidades:

> **Consulta de chamadas efectuadas:** todas as chamadas a cobrar e que passem pela plataforma são disponibilizadas aos utilizadores da aplicação com todo o pormenor necessário ao esclarecimento do cliente ou ao despiste de reclamações ou tentativas de fraude;

> **Consulta de serviços subscritos (figura 8):** todos os serviços subscritos, cancelados ou suspensos, bem como o respectivo histórico de alterações;

> **Consulta de transacções:** todas as transacções realizadas pelo cliente (uma troca de plano, um bónus que foi atribuído, etc.) são registadas e podem ser mostradas;

> **Troca de número;**

> **Troca de Plano.**

Serviços
A gestão de serviços, como por exemplo a sua disponibilização ou retirada a todos os clientes, a sua tarifação, os bónus associados à respectiva utilização, etc., são algumas das funcionalidades disponíveis neste módulo. As interfaces com os diversos sistemas de aprovisionamento de serviços, como são as plataformas WAP, de SMS, etc, estão também concentradas neste módulo.

Os serviços são a parte de maior dinamismo que um operador possui para se diferenciar positivamente da concorrência. O NGIN *Care* possui disponíveis os mecanismos de suporte a diversos serviços, quer ao nível de interfaces com os sistemas de suporte externos, quer ao nível da respectiva tarifação.

Exemplos desses serviços são:

> **Listas de números favoritos:** mais do que uma lista pode ser suportada simultaneamente;

> **Serviços de rede clássicos**, como sejam *Call Forwarding*, Identificador do Chamador, *Roaming*, …;

> **Pacotes de serviços** compostos arbitrariamente a partir de outros, como por exemplo dando 50 SMS e acesso WAP a uma tarifa económica;

> **…**

Os serviços podem ser configurados por plano tarifário e por estado do cartão, para serem activados/desactivados quando o cartão muda de estado.

Um serviço pode ainda ser atribuído através duma promoção, com ou sem data limite de utilização.

Bónus e Promoções
Uma promoção tem sempre um período de validade, durante o qual, ao acontecer o evento pré-configurado (uma activação do telemóvel ou dum serviço, uma recarga, troca de número ou de aparelho, etc.), é atribuído ao cliente um bónus ou um serviço. Existem vários tipos de possibilidades para os bónus tais como créditos atribuídos para determinado tipo de chamadas, créditos em saldos específicos, parcelados, *steps* incrementais, atribuição de prémios,…

A plataforma NGIN providencia saldos de vários tipos que, conjugados com os bónus e promoções, possibilita uma grande diversidade de configurações.

Cada promoção pode ter associadas SMS, que são enviadas ao cliente como estímulo para a participação em determinada promoção (por exemplo, incentivando à recarga, no prazo limite para o fazer).

A conjugação de todos estes parâmetros possibilita a configuração de um conjunto bastante vasto de bónus e promoções, que fazem a diferença dos operadores com o NGIN.

Configuração do NGIN Care
O NGIN Care é configurado através de um módulo próprio que disponibiliza, entre outras, as seguinte funcionalidades:

> **Autenticação:** cada utilizador tem uma palavra de acesso atribuída que tem obrigatoriamente que trocar no primeiro acesso;

> **Configuração do menu da aplicação:** cada nova funcionalidade deve ser colocada num de quatro níveis do menu. O nome do próprio menu é também alterável com um simples duplo clique;

> **Controlo de Acessos baseado em Perfis:** cada utilizador da aplicação possui um e um só perfil atribuído (um perfil de utilizador é, por exemplo, Operador de *Back-Office*, etc.). Todas as funcionalidades são depois incluídas nos menus com permissões atribuídas a cada um dos perfis definidos;

A opção de controlo de acessos por perfil de utilizador permite que a mesma aplicação tenha um vasto leque de tipos de utilizadores, podendo inclusive alguns utilizadores acederem via *extranet* (em redes de parceiros, por exemplo).

3.3 NGIN Bit

As ferramentas NGIN BIT (*Reporter* e *Mart*), têm como principal função, entregar ao cliente, e utilizador, informação, em tempo-real, numa forma adequada para responder a questões do tipo:

> Qual o número total de chamadas em Janeiro?

> Qual o número de chamadas ao longo dos dias de Janeiro?

> Qual o número de tentativas de acesso ao serviço em Janeiro dos clientes registados em cada SCP (*Service Control Point*)?

> Quanto e quando recarregam mais os clientes ao longo da semana?

> Qual a eficácia em termos de chamadas completas para o destino X? E para o destino Y?

> Como tem evoluído o n.º de clientes do parque ao longo do tempo?

> ...

Estas ferramentas processam e armazenam dados operacionais e de negócio NGIN UZO num repositório central, possibilitando aos gestores NGIN UZO o acesso às informações que procuram, por recurso a um *web browser*.

A informação operacional é extraída da plataforma e processada até ao seu armazenamento sob a forma de um Modelo Dimensional (entenda-se por Modelo Dimensional, um conjunto de indicadores e dimensões de análise logicamente relacionados entre si, num determinado contexto de negócio – consiste em tabelas de factos centrais, que cruzam com um conjunto de tabelas de dimensões, constituindo estrelas ou *Data Marts*).

Assim, para uma determinada área do negócio será possível criar relatórios, recorrendo à visualização dos indicadores numéricos (métricas) constantes da tabela de factos e às dimensões que são relevantes e estão relacionados com o conceito de negócio em causa.

Os NGIN *Bits* (*Mart* e *Reporter*) permitem assim fornecer, em tempo real, Informação de Negócio (e.g. recargas, clientes (ARPU), tarifação) e de desempenho (e.g. tráfego), por tempo, tipo de recarga, canal de recarga, SCP, tipo de serviço, estado dos clientes, destino, tipo de chamada ou tipo de tarifa.

3.4 NGIN SIGO

A Aplicação SIGO-UZO constitui uma solução preparada para dar resposta à complexidade dos ambientes de Gestão Operacional. Baseando-se em conceitos de *Workflow*, a solução visa gerir e cadastrar todas as Operações (Pedidos, Anomalias, Pontos de Situação, Encomendas, ...), sendo possível criá-las e acompanhar a sua evolução relativamente aos seus estados e respectivos intervenientes.

O SIGO-UZO encontra-se construído sob uma arquitectura em 3 camadas seguindo uma abordagem orientada por objectos. A adopção de tecnologias Internet permite ao SIGO-UZO disponibilizar uma abordagem cliente-servidor adequada a operar em Intranets. São disponibilizados interfaces tipo *web browser*, com soluções de interacção gráficas ajustadas e especialmente vocacionadas para acessos a partir de locais com recursos de transmissão escassos.

3.5 NGIN *Gateway* de Pagamentos

O *Gateway* de Pagamentos UZO é o componente UZO que gere e garante todas as transacções associadas a recargas UZO. Implementa diferentes protocolos que permitem o carregamento de cartões recorrendo à SIBS (quer recorrendo ao menu "Pagamentos de Serviços", quer através do carregamento personalizado debaixo do actual menu das ATM para carregamentos TMN), ao Sistema de Pagamentos PayShop e ainda a um novo quarto canal de recarga, a lançar durante o quarto trimestre de 2006.

Gerida através de uma interface de administração (*web based*) é possível monitorar, controlar e garantir a segurança dos pagamentos/carregamentos através dos diferentes canais de recarga.

Dada a criticidade deste componente, a *Gateway* de Pagamentos UZO é baseada numa arquitectura dotada de elevados graus de segurança, eficiência, robustez e redundância.

3.6 NGIN IVR (InoVox)

Um dos canais de interacção entre a plataforma e os utilizadores é executada através dos chamados Periféricos Inteligentes, módulos sofisticados de *hardware* e *software* capazes de interagir com os clientes através de anúncios de voz e de recolher as suas informações através do premir dos dígitos do telefone ou mesmo através de reconhecimento automático de palavras (IVR – do inglês I*nteractive Voice Response*).

Na base dos serviços interactivos de voz está o componente InoVox, que disponibiliza as facilidades básicas de IVR (interface com as redes telefónicas, capacidade para tocar e gravar anúncios vocais, recolha de marcação DTMF, etc.) bem como funcionalidades mais avançadas (ex: reconhecimento e síntese fala, interface para redes de terceira geração 3G, VoIP, vídeo).

Para além dos IVR InoVox há ainda uma série de componentes adicionais, que acompanham os serviços, e que são necessários para o armazenamento de dados, disponibilização de estatísticas, alarmes, monitorização, configuração, etc., que fazem parte da arquitectura de referência da plataforma de serviços interactivos de voz da PT Inovação.

3.7 NGIN Portal

O *website* da UZO, seguindo igualmente o paradigma da "descomplicação", foi desenvolvido em parceria com uma outra empresa

figura 9 – Meu UZO

nacional, cabendo à PT Inovação a responsabilidade do desenvolvimento da sua componente Portal (Meu UZO).

Para além das funcionalidades e opções disponíveis no *website* (*e.g.* informações; adesão, escolha de número, loja, FAQs, suporte, *on-line chat*, contactos,…) é na componente Portal (Meu UZO)figurara 9) que residem as funcionalidades mais directamente associadas ao cliente tais como:

> Autenticação (*Login/password*);

>Subscrição de Serviços/Planos/ Planos/Promoções);

> As minhas Comunicações (consultas);

> Recargas;

> Envio de SMS/MMS (a disponibilizar em breve);

> Roubo de Telemóvel;

> ....

A componente Portal do *website* interage directamente com o módulo NGIN *Care* permitindo ao cliente, em tempo real, o acesso a toda a informação detalhada que lhe diz respeito, nomeadamente as relativas à sua informação pessoal e às suas comunicações (voz, SMS, MMS, recargas, operações, serviços, planos e promoções subscritos, ...).

3.8 NGIN *Manager*

O NGIN *Manager* é a componente de Gestão global da solução NGIN, responsável por todo o cadastro e toda a análise, monitoria, controle, alarmística e *reporting* das restantes componentes da solução, quer ao nível do Serviço quer ao nível das Plataformas (*Hardware* e *Software*). É através deste módulo, utilizado essencialmente pelos departamentos que garantem a Operação, Manutenção e Suporte das plataformas, que são garantidos os níveis adequados de desempenho e de qualidade de serviço das Plataformas NGIN.

4. Conclusões

Foram apresentados alguns conceitos base sobre Operadores Móveis Virtuais bem como descrita, sumariamente, a solução de Rede Inteligente da PT Inovação NGIN MVNO que actualmente suporta o negócio da UZO.

A Solução NGIN disponibilizada para a UZO, pela sua flexibilidade, permite a esta Unidade de Negócio da TMN deter uma vantagem competitiva clara sobre outras eventuais operações do tipo MVNO levadas a cabo pela concorrência, não só ao nível da operação do negócio MVNO UZO *per si*, como também potenciar à TMN, de uma forma fácil e rápida o lançamento de novas áreas de negócio do tipo MVNE, fazendo o *hosting* de outros potenciais MVNO.

José Domingues
Ingressou no Centro de Estudos de Telecomunicações (CTT) em Aveiro 1980 para a Divisão de Engenharia de Tráfego onde esteve ligado à concepção e desenvolvimento de sistemas de informação de apoio ao planeamento, projecto e gestão técnica de redes de telecomunicações. Foi responsável por acções de formação sobre planeamento e projecto de redes locais de telecomunicações realizadas para as empresas do Grupo Portugal Telecom. Ascendeu ao cargo de Chefe de Divisão de Engenharia de Tráfego em 1986, assumindo posteriormente os cargos de Chefe de Divisão de Recursos Estratégicos, Chefe de Divisão de Integração de Tecnologias e Serviços, Director de Comunicação e Imagem, Director de Marketing & Comercial e actualmente é Director no Gabinete da Comissão Executiva da PT Inovação SA com responsabilidades específicas pelo desenvolvimento estratégico, parcerias e pela gestão dos clientes Grupo Telefónica e TMN/UZO.

03

NGIN Pack ISP

Carlos Rodrigues,
Ricardo Gonçalves

palavras-chave:
Next Generation Network (NGIN),
Controlo de Acessos IP,
Tarifação *Near Real Time*

Este artigo apresenta resumidamente o projecto NGIN Pack ISP, a nova solução de controlo de acessos e tarifação para os Pré-Pagos da Telepac.

Este projecto tinha como objectivo a substituição da antiga plataforma por uma que satisfizesse as necessidades da Telepac em termos de fiabilidade, flexibilidade e de capacidade de evolução para novas funcionalidades, mantendo inalteradas as interfaces de comunicação com todos os componentes externos.

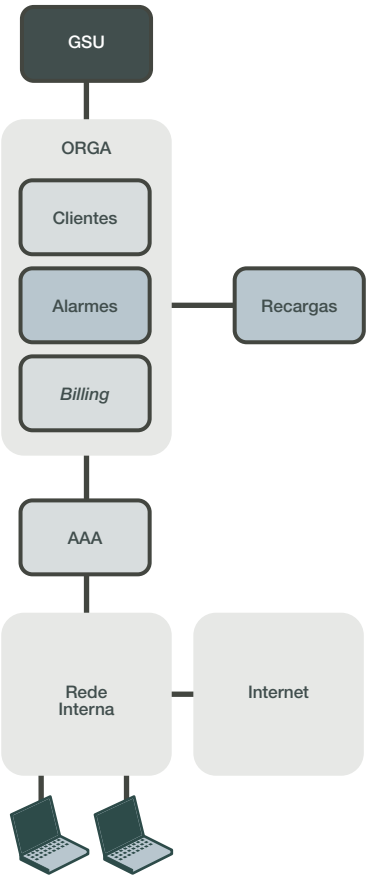

figura 1 – Estrutura antiga da plataforma de controlo de acessos da Telepac

1. Introdução

Em 2005 a Telepac solicitou à PT Inovação o desenvolvimento de uma solução suportada na plataforma NGIN Pack ISP, a operar em modo *Near Real Time*, que satisfizesse a necessidade de evolução para novas funcionalidades do sistema utilizado actualmente no controlo de acessos *Dial-Up* e ADSL Pré-Pago. Este desenvolvimento apresentava diversos desafios, entre os quais a completa substituição da antiga plataforma sem qualquer impacto nos módulos com os quais interagia, requisitos de performance no controlo dos acessos e a implementação das funcionalidades existentes e de outras novas.

O NGIN Pack ISP é, desde 1 de Abril de 2006, a nova peça chave no controlo e tarifação dos produtos Pré-Pagos da Telepac, com mais de 300.000 clientes aprovisionados.

2. Estrutura Inicial e Requisitos

A antiga plataforma da Telepac, designada por ORGA, não fornecia as condições necessárias para a implementação de novos produtos e novas funcionalidades devido à sua antiguidade, rigidez ao nível de configuração e impossibilidade de evolução, qualidades que são apanágio da solução NGIN *Pack*.

A primeira fase do projecto consistiu no levantamento da estrutura existente em produção e dos requisitos necessários para o desenvolvimento do NGIN *Pack* ISP. Em termos gerais, a estrutura poderia ser representada como ilustra a figura 1.

No centro, a plataforma ORGA, responsável por executar quase todas as operações necessárias ao funcionamento do sistema de Pré-Pagos. Essas funções iam desde o controlo de acessos através da interface com os servidores Radius, tarifação das sessões dos clientes, operações de aprovisionamento via o sistema de gestão de utilizadores GSU, geração de alarmes *e-mail* relacionados com os carregamentos e prazos e processamento da informação das recargas dos clientes.

O planeamento e concepção do NGIN Pack ISP foram delineados com base em três grandes requisitos, descritos em seguida, cujo cumprimento dependia o sucesso da implementação da solução.

1.º – O modo de funcionamento do ORGA teria de ser replicado integralmente e as interfaces com os vários componentes externos teriam de ser mantidas para que não houvesse o mínimo impacto na transição entre plataformas;

2.º – Performance da plataforma. Devido à existência do *Service Level Agreement* (SLA) relativo às respostas das invocações dos servidores Radius em que estas teriam de ter uma duração máxima de 40 milissegundos, caso contrário, seriam descartadas afectando desta forma os acessos dos clientes em causa, o planeamento teve de contemplar a optimização dos recursos e da arquitectura;

3.º – O sistema teria de integrar inúmeras novas funcionalidades mas ao mesmo tempo manter uma elevada capacidade de

configuração do serviço que permitisse à Telepac a criação dos produtos pretendidos. Para além disso, o sistema deverá ser bastante modular e estruturado de forma a permitir, no futuro, a fácil implementação de novas funcionalidades.

3. Desenvolvimento do NGIN *Pack* ISP

Após a análise dos requisitos do projecto, a estrutura definida para o NGIN *Pack* ISP tomou o aspecto representado na figura 2.

No centro desta nova plataforma encontra-se o *core* do NGIN *Pack* ISP, residindo numa base de dados Oracle, que alberga toda a informação dos clientes, a configuração do sistema, a informação de *rating* e é onde são executadas as lógicas de Serviço, Aprovisionamento e *Billing*. Estas lógicas reproduzem o funcionamento interno do ORGA, implementando as mesmas regras de negócio e processando a infor-maçã̃o disponibilizada pelos vários sistemas externos.

À volta do *core* encontram-se os vários subsistemas que disponibilizam as interfaces de configuração e comunicação com os componentes externos. Um dos principais módulos encontra-se nas duas máquinas de *Front End*, o Relay2db. Desenvolvido com recurso à linguagem Perl, é o ponto de contacto do NGIN *Pack* ISP com os vários componentes externos que interagiam com o ORGA, disponibilizando para tal interfaces semelhantes. Tanto o GSU (Sistema de Gestão de Utilizadores) como os servidores AAA comunicam directamente com este módulo, sendo as suas mensagens interpretadas e enviadas via SDP para o *core* para processamento.

Tal como nas restantes plataformas NGIN, o sistema de *rating* é baseado no NRF (*NGIN Rating Function*). Por razões de *load balancing* e *failover*, este módulo encontra-se instalado nos dois *Front Ends*.

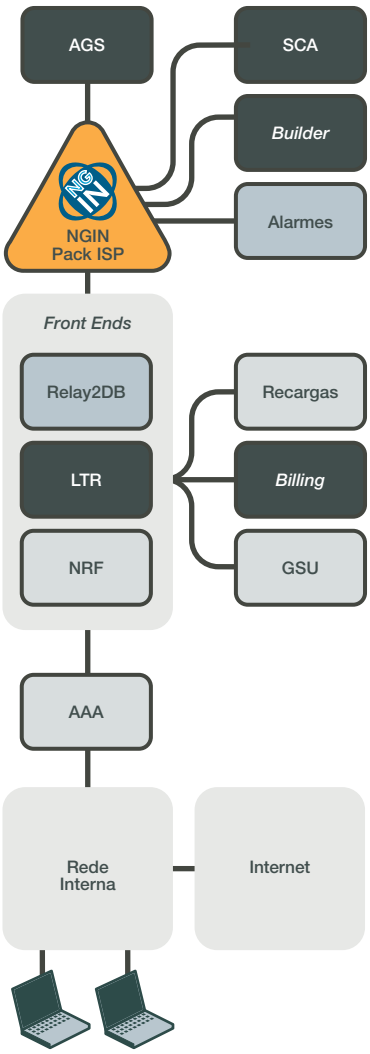

figura 2 – Estrutura da nova plataforma de controlo de acessos da Telepac.

Como se trata de um sistema *Near Real Time*, apenas existe um pedido de autenticação no início de cada sessão. É sobre a duração da resposta desse pedido que foi definido o SLA referido anteriormente. Durante a duração da sessão não existe mais nenhum envio de informação para o NGIN indicando consumos ou informação relevante da sessão. Apenas após o fim da sessão é que, através de uma operação disponibilizada no *Relay2db* e invocada pelo GSU, o NGIN recebe e processa todos os dados de *Billing*.

As recargas do saldo dos clientes são colocadas nos *Front Ends* em formato de ficheiros que serão posteriormente processados pelo módulo de *Lost Ticket Return* (LTR), existente noutras plataformas NGIN e que foi alterado especificamente para o processamento de informação contida nesses ficheiros.

Devido ao facto dos clientes aprovisionados nesta plataforma serem Pré-Pagos, as recargas dos utilizadores, para além da actualização do respectivo saldo, irão despoletar inúmeras operações que poderão alterar o estado da conta do cliente e dar origem ao envio de alarmes aos clientes se as condições assim o exigirem. Actualmente, o NGIN *Pack* ISP apenas dispõe na sua interface de alarmes do envio de mensagens de *mail*. Futuramente, está previsto o envio igualmente de SMS.

Um sistema não fica completo sem uma interface de configuração. Os próximos três módulos, o *Builder*, o SCA e AGS são dedicados a esse propósito.

O *NGIN Builder* (figura 3) é a aplicação utilizada para a configuração do *Rating* permitindo a criação de regras complexas de tarifação com base nas características de cada produto, no horário de utilização, no tipo de tarifação (tempo ou volume) e nos vários saldos atribuídos a cada cliente.

A configuração dos vários produtos é feita através da aplicação *Web* designada por Aplicação de Gestão do Serviço, AGS (figura 4). Nela é possível a criação de várias entidades que contêm e definem as características das contas, nomeadamente os Perfis de Conta, Perfis de *Plafond*, Perfis de Serviço e os Serviços Básicos. É nestas entidades que são definidos os limiares de saldo e de recarga, os *plafonds* de cada conta, o conteúdo e as condições de cada alarme e notificações aos clientes, etc.

Configurações de sistema, *settings*, repositórios de ficheiros e outras características inerentes ao funcionamento do sistema são definidos

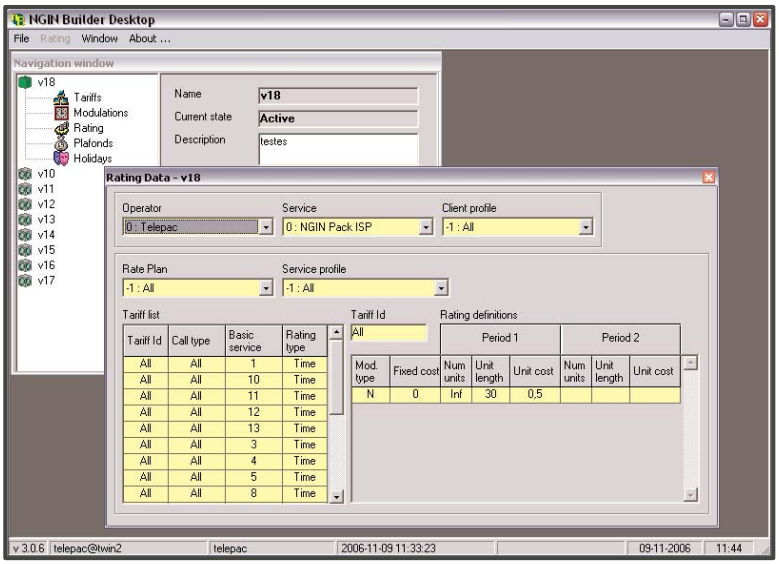

figura 3 – Aplicação NGIN Builder utilizada para a configuração do Rating

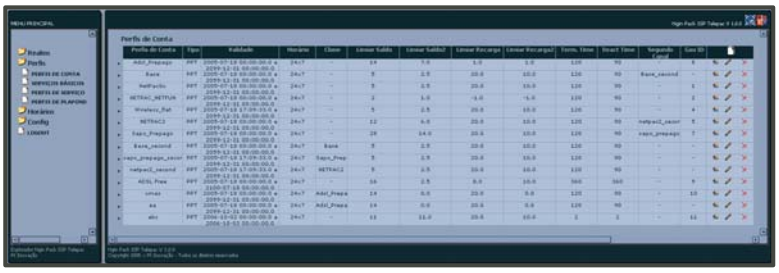

figura 4 – Página da AGS dos Perfis de Conta

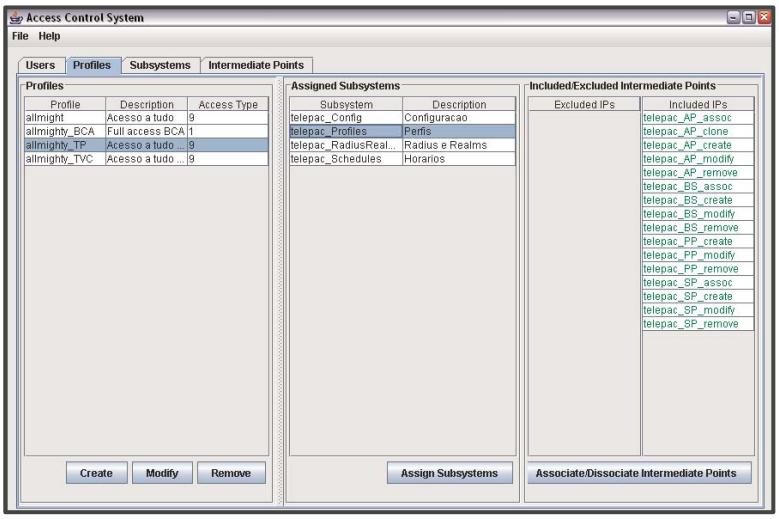

figura 5 – Sistema de Controlo de Acessos

igualmente através desta interface, permitindo desta forma um elevado grau de liberdade e flexibilidade.

O Sistema de Controlo de Acessos, o SCA (figura 5) permite a criação de *logins* para a AGS e de perfis para cada *login* que definem o conjunto de páginas e operações que lhe são apresentadas, permitindo desta forma adaptar a AGS conforme o tipo de funções que o utilizador desempenha.

4. Conclusão

O sucesso do NGIN *Pack* ISP dependia em grande parte da replicação de todas as funcionalidades da antiga plataforma para que a sua substituição não implicasse alterações significativas nos sistemas com os quais interagia. Para além do desenvolvimento do *core* da plataforma, foi necessário desenvolver todo um conjunto de módulos e interfaces que permitissem a configuração de todas as variantes do sistema mas, ao mesmo tempo, com grande flexibilidade e autonomia.

Todo o trabalho feito neste projecto deu os seus frutos quando o NGIN *Pack* ISP substituiu a plataforma ORGA sem que fossem necessárias mudanças ou ajustes nos módulos externos e pela reacção extremamente positiva por parte da Telepac relativamente às capacidades do sistema.

Referências

http://www.ptinovacao.pt/produtos/ngin.htm
http://www.telepac.pt
http://adsl.sapo.pt/

Carlos Rodrigues, licenciado em Engenharia Electrónica e Telecomunicações pela Universidade de Aveiro, em 2001. Iniciou a sua actividade profissional na Portugal Telecom Inovação em 2001, desempenhando actividades na área das redes IP, mais concretamente IPv6, Mobilidade IPv6, redes de acesso, VoIP e qualidade de serviço no âmbito de diversos projectos de investigação (LONG, Euro6IX, Torrent, etc.). Desde Outubro de 2004 participa no desenvolvimento de aplicações e arquitecturas para controlo de sessões de dados em tempo-real.

Ricardo Gonçalves, licenciado em Engenharia Electrónica e Telecomunicações pela Universidade de Aveiro em 2002. Pós Graduação em Sistemas de Informação pela Universidade de Aveiro decorrida durante 2003. Estágio profissional na empresa Portugal Telecom Inovação sob o tema Pagamentos Electrónicos em 2004 Colaborador da Portugal Telecom Inovação desde 2005 tendo desempenhado funções de programador de PL/SQL no âmbito dos projectos NGIN *Pack WiFi e* NGIN *Pack* ISP sendo actualmente Responsável de Acção dos mesmos.

04

PT Inovação, bom dia! ORADIGA

Serviço de Atendimento Automático com Reconhecimento da Fala em Português

Luís Almeida, Alexandre Sapage, Joaquim Azevedo, Glória Branco, Nuno Beires

palavras-chave:
Serviços Interactivos de Fala, Atendimento Automático, Reconhecimento Automático de Fala, Síntese de Texto para Fala, VoiceXML.

O ORADIGA é um serviço de atendimento automático para a PT Inovação que permite aos utilizadores estabelecer uma ligação telefónica com qualquer colaborador da empresa dizendo apenas o seu nome. O serviço é suportado na plataforma InoVox-MSP – InoVox *Multimedia Service Platform* – desenvolvida pela PT Inovação e é baseado em tecnologia de reconhecimento de fala independente do orador. O ORADIGA é intuitivo, de fácil utilização e evita a memorização de números de telefone e extensões – apenas um número para toda a empresa, qualquer que seja o acesso (interno e externo, fixo ou móvel).

O ORADIGA é um serviço interactivo de voz que ouve e fala em Português.

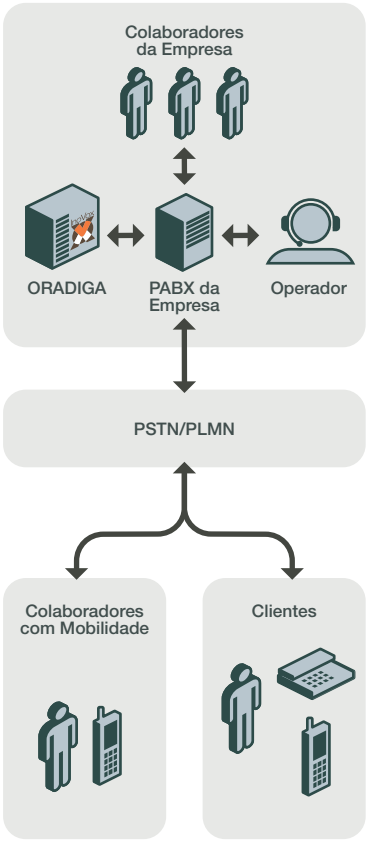

figura 1 – Serviço ORADIGA

1. Introdução

Serviços comerciais baseados em Sistemas Interactivos de Voz, *Interactive Voice Response* (IVR) na terminologia inglesa, são hoje uma realidade e estão presentes em muitas das operações do nosso dia-a-dia. Serviços financeiros, de informações, de atendimento ao cliente utilizam cada vez mais este tipo de tecnologias.

O grau de maturidade atingido pelas tecnologias da fala aplicadas em redes de telecomunicações está a evoluir gradualmente da tradicional interacção baseada na marcação de tons multifrequência (DTMF) para uma interacção mais natural e amigável com recurso ao reconhecimento da fala. O reconhecimento automático de fala (ASR) permite aos computadores escutar e interpretar o que foi dito e, por sua vez, a síntese de texto para fala (TTS) permite que estes falem.

A adopção destas soluções revela-se vantajosa quer para as empresas, quer para os clientes. Do ponto de vista empresarial permitem ganhos de eficiência, pois, de uma forma mais amigável para o Utilizador/Cliente, facilitam a automatização de um conjunto de processos, permitem uma resposta mais rápida às solicitações e libertam os operadores humanos para actividades mais complexas. Na perspectiva do Utilizador os serviços de fala são mais amigáveis e fáceis de utilizar. São naturais, intuitivos, mais rápidos e flexíveis que os sistemas DTMF e são facilitadores da mobilidade, pois podem ser utilizados em qualquer lugar, em qualquer circunstância e por todos os utilizadores. Para além disso, são mais acessíveis para pessoas com necessidades especiais, visuais ou motoras.

A PT Inovação, por intermédio do Pólo do Porto, tem acompanhado a evolução das tecnologias da fala e contribuído activamente para o desenvolvimento e utilização dos recursos tecnológicos e linguísticos necessários ao desenvolvimento de serviços de fala em português. A plataforma InoVox-MSP – solução IVR desenvolvida pela PT Inovação – e as tecnologias da fala que a plataforma integra e suporta, são a base dos serviços interactivos de voz já desenvolvidos para o Grupo PT, e concretizam, literalmente, "a tecnologia que fala Português".

É neste enquadramento que surge o ORADIGA (OpeRadorA Digital Interactiva portuGuesA), um serviço de atendimento automático para a PT Inovação. Com base na tecnologia de reconhecimento de fala e na integração computador-telefonia (CTI), disponibilizadas pela plataforma InoVox-MSP, foi desenvolvido o serviço ORADIGA (figura 1), que fornece uma interface de fala para acolher e encaminhar os clientes e que gere as chamadas internas dos colaboradores da empresa.

Os serviços de atendimento automático têm vindo progressivamente, e com maior incidência nestes últimos anos, a conquistar a sua quota de mercado quer como solução *standalone*, empresarial ou de operador (com alojamento do serviço), quer como funcionalidade avançada dos (V)PABX, IP PABX.

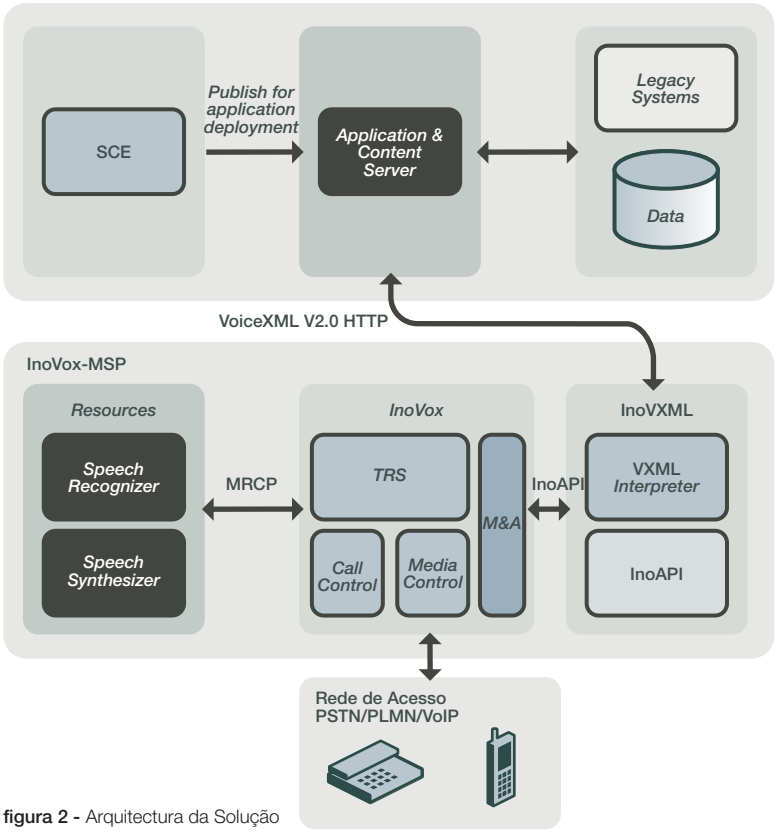

figura 2 - Arquitectura da Solução

Podem até, muito facilmente, ser portados ou integrados, como um *service enabler*, numa arquitectura *IP Multimedia Subsystem* (IMS).

2. Serviço ORADIGA

2.1. Conceitos e motivação

Os serviços de atendimento automático em geral, de qual o ORADIGA é um exemplo, são serviços *call-in* que permitem aos utilizadores, quer da rede fixa, quer da móvel, a marcação por fala usando linguagem natural. Assentes em tecnologia de reconhecimento de fala independente do orador, intuitivos e de fácil utilização, permitem aos clientes a realização de chamadas para os contactos pretendidos através de simples comandos vocais.

Os factores de motivação que sustentam e fomentam o crescimento deste tipo de serviços estão relacionados com:

> **redução de custos –** diminuindo a componente do operador humano nos serviços de atendimento;

> **aumento de produtividade –** ajudando os colaboradores a encontrarem-se mais rapidamente e garantindo serviços de encaminhamento na ausência das recepcionistas de atendimento;

> **mobilidade –** complementando os serviços de contactos *online* e acomodando o aumento da utilização dos telemóveis;

> **segurança na condução –** simplificando o estabelecimento do contacto telefónico, permite manter as mãos no volante e focar a atenção na condução.

2.2 Abordagem Técnica

Para a implementação deste serviço foi acordada a utilização de protocolos e metodologias de desenvolvimento normalizadas. Uma das decisões iniciais prendeu-se com a escolha da interface de programação de serviços (API) a utilizar.

A solução proprietária é tipicamente mais granular e possibilita um melhor e mais refinado controlo das várias funcionalidades disponibilizadas pelo IVR no desenvolvimento dos serviços. Requer, contudo, conhecimentos profundos por parte do programador em termos das funções para controlo de chamada, processamento e síntese de fala, conferência, etc. Por sua vez, a solução *standard* simplifica o processo de codificação, pois abstrai a complexidade inerente às API proprietárias, reduzindo significativamente a curva de aprendizagem e o tempo de desenvolvimento.

O serviço a desenvolver teria de utilizar tecnologia de fala independente do orador e ser flexível em termos da interacção com o utilizador. Pretendia-se que o utilizador pudesse usar uma linguagem próxima da natural e que o serviço fosse capaz de compreender quer comandos simples, quer frases mais complexas. A iniciativa da interacção teria de ser mista de forma a torná-la mais natural, podendo o utilizador interromper o serviço, em qualquer altura, para efectuar a marcação se assim o desejasse.

Analisando os requisitos do serviço em termos de transparência do protocolo para transporte da voz sobre as redes PSTN e VoIP, e atendendo às funcionalidades específicas que requeriam reconhecimento e síntese de fala em Português, a decisão recaiu sobre a utilização, sempre que possível, de protocolos *standard* e em particular do *Voice*XML[1] suportado na plataforma InoVox-MSP.

2.3 Arquitectura da Solução

A arquitectura da solução para o serviço ORADIGA assenta na plataforma InoVox-MSP para o *front end* telefónico (IVR) e num servidor *Web* como é ilustrado na figura 2.

A solução é totalmente modular e escalável. Os módulos são independentes e podem ser instalados num ou em vários servidores consoante as necessidades, suportando múltiplos serviços e optimizando a partilha de recursos. O reforço de capacidade e a redundância da solução são possíveis aumentando o número de canais telefónicos ao nível do IVR ou adicionando mais

IVR e instalando novos servidores de aplicações e *Web*. Segue-se uma breve descrição de cada um dos principais módulos da solução ORADIGA.

O ***InoVox Multimedia Service Platform*** (InoVox-MSP) é uma plataforma multi-serviço, de arquitectura aberta e flexível, escalável e de elevado desempenho, visando o rápido desenvolvimento de serviços interactivos multimédia avançados sobre redes PSTN, convergentes ou totalmente IP. Socorrendo-se duma combinação de tecnologias de fala – reconhecimento da fala (ASR), síntese de fala (TTS) e verificação do orador (SV), conjugadas com funcionalidades PSTN e capacidades de VoIP avançadas, a plataforma InoVox-MSP está optimizada para as redes de hoje bem como para as redes *All IP* da próxima geração, traduzida pela evolução para IMS.

O InoVox-MSP integra e implementa *standards* abertos do W3C e do IETF, que revolucionam a metodologia de desenvolvimento e de disponibilização de serviços interactivos de voz, fazendo evoluir o IVR do mundo das plataformas fechadas para o paradigma da Internet. Estes *standards* incluem o *Voice*XML 2.0, o SSML[2] e o SRGS[3]. Para integração das tecnologias de fala, nomeadamente o reconhecimento e a síntese, o InoVox-MSP utiliza o *Multimedia Resource Control Protocol* (MRCP).

O módulo **InoVXML** é baseado no interpretador de *Voice*XML *open-source Open*VXI[4] e permite a especificação de vários serviços telefónicos baseados em *Voice*XML 2.0 na plataforma InoVox-MSP. Este módulo estabelece associações entre uma chamada ou sessão e um determinado serviço, tomando a iniciativa de solicitar ao *Application & Content Server* (via HTTP) as páginas *Voice*XML respectivas, sendo responsável pelo seu carregamento e posterior processamento. De acordo com a lógica do serviço, o InoVXML invoca os recursos do InoVox-MSP e acede aos vários tipos de *media* através da InoAPI. Esta interface disponibiliza mecanismos de controlo e monitorização dos vários recursos da plataforma InoVox-MSP.

O ***Application & Content Server*** é o módulo responsável pela implementação da lógica do serviço e pela geração do respectivo código *Voice*XML. É igualmente responsável pelo armazenamento dos vários recursos que incluem: ficheiros de áudio pré-gravado, gramáticas de voz ou DTMF, ficheiros ECMAS*cript*, bibliotecas para acesso a aplicações ou serviços, entre outros. Este módulo recorre a um J2EE *Container* para a geração do *Voice*XML em resposta aos pedidos efectuados pelo interpretador e acede aos repositórios de dados necessários à implementação do fluxo do serviço.

O ***Service Creation Environment*** (SCE) é o ambiente gráfico para o desenvolvimento e depuração do serviço. O SCE disponibiliza um conjunto de funcionalidades que facilitam o desenvolvimento de serviços em *Voice*XML. Estas facilidades incluem, entre outras, a geração automática de código fonte a partir de *templates*, a verificação automática da sintaxe, a análise de eventuais problemas e usabilidade do serviço, dispondo ainda de ajuda contextualizada. Um mecanismo para transferência e instalação dos serviços no *Application & Content Server* está também disponível. Este ambiente de desenvolvimento integrado é constituído pelo Eclipse IDE [5], com um *plug-in* específico para desenvolvimento de aplicações em *Voice*XML, e está integrado com um J2EE *container* para testar a geração dinâmica de *Voice*XML.

2.4 Tecnologias

A utilização das tecnologias de reconhecimento e síntese de fala é inerente ao desenvolvimento de serviços interactivos de atendimento automático. O reconhecimento de fala para este tipo de serviços caracteriza-se pela independência em relação ao orador, por ser contínuo e por um vocabulário de pequena ou média dimensão. Por vocabulário entende-se o conjunto de palavras e/ou expressões que podem ser reconhecidos pelo sistema em cada instante. Estas palavras e/ou expressões são definidas e especificadas num ficheiro de texto, designado por gramática. As gramáticas têm como função delimitar as frases que são geradas pelo reconhecedor, melhorando o seu desempenho para o domínio semântico desejado.

A gramática obedece a um formato XML standard de acordo com a especificação SRGS. A principal gramática do serviço ORADIGA contém os nomes dos colaboradores da PTIN e inclui entradas para os seus nomes abreviados. Esta gramática foi afinada de forma a flexibilizar a marcação por voz para as situações em que se podia reduzir o nome abreviado do colaborador de três para dois nomes.

Um outro aspecto que teve que ser acautelado foi a pronúncia de nomes estrangeiros. Para cada um destes nomes foi necessário gerar a respectiva transcrição fonética em SAMPA, tendo sido criado um dicionário com as várias transcrições.

Para além do reconhecimento de fala (ASR) para processar o que foi dito pelo utilizador, o serviço recorre à síntese de fala (TTS) para gerar anúncios, relativos a informação dinâmica, que complementam os anúncios pré-gravados com informação estática. Para ambas as tecnologias, o produto comercial utilizado foi o da *Nuance*: OSR 3.0[6] para o ASR e o *RealSpeak* 4.0[7] para TTS.

O amadurecimento destas tecnologias, conjugado com a normalização

e expansão das tecnologias *Web* levou a uma progressiva utilização do *Voice*XML para o desenvolvimento de serviços telefónicos.

A linguagem *Voice*XML constituiu um nível de abstracção em XML para a programação de serviços através de *standards* para a área dos IVR, contemplando as seguintes funcionalidades:

> Geração de áudio sintetizado (TTS);

> Tocar ficheiros de áudio pré-gravados;

> Gravação de ficheiros de áudio;

> Reconhecimento de DTMF (garantindo a compatibilidade com os serviços telefónicos tradicionais);

> Reconhecimento de Fala (ASR);

> Transferência e fim de chamadas.

Disponibiliza também mecanismos lógicos para o controlo do fluxo a implementar. Facilita as actualizações ao serviço através da simples edição do documento VoiceXML, alojado no *Application & Content Server*. A alteração estará disponível na próxima chamada ao serviço, sem necessidade de recompilar ou instalar qualquer aplicação.

2.5 Funcionalidades

O serviço de atendimento automático possibilita a marcação por voz para as extensões dos vários colaboradores da empresa. O utilizador liga para o número do serviço através de um terminal fixo, móvel ou VoIP. O serviço atende a chamada e toca um anúncio de boas vindas solicitando o nome da pessoa com a qual o utilizador deseja falar. Sempre que a resposta deste é compreendida pelo sistema com um nível de confiança elevado, a chamada é transferida para o número de destino correspondente.

O serviço reconhece o que é dito pelo utilizador quer sejam frases completas ou simples palavras chave. Por exemplo:

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da PT Inovação. Com quem deseja falar?

<U> Quero falar com António Silva por favor

<S> António Silva. Um momento por favor.

OU

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da…[Barge-in]

<U> António Silva

<S> António Silva. Um momento por favor.

Como se pode observar no exemplo acima, o serviço permite a interrupção dos anúncios (*Barge-in*) por parte do utilizador, possibilitando assim uma interacção mais rápida.

Caso o nível de confiança no resultado do reconhecimento de fala seja inferior a um determinado limiar (pré-definido e configurável), o serviço solicita uma confirmação, por parte do utilizador, do nome que foi reconhecido. Se o nível de confiança for mesmo muito baixo, o serviço solicita a repetição do nome do colaborador. No exemplo seguinte ilustra-se o pedido de confirmação do nome do colaborador.

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da PT Inovação. Com quem deseja falar?

<U> António Silva

<S> Pretende falar com António Silva? Diga sim ou não!..

O serviço implementa ainda um mecanismo para desambiguar as situações em que existe mais do que uma entrada na base de dados do serviço que satisfaça o nome proferido pelo utilizador. Nesta situação é gerada e apresentada a lista dos possíveis colaboradores podendo o utilizador navegar e escolher o destinatário da sua chamada através de confirmação verbal.

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da PT Inovação. Com quem deseja falar?

<U> Pedro Silva

<S> Desculpe, disse Pedro Silva? Confirme dizendo sim ou não.

<U> Sim

<S> Pretende falar com João Pedro Silva. Diga sim ou não!

<U> Não

<S> Pretende falar com Pedro Silva Costa. Diga sim ou não!

<U> Sim

<S> Pedro Silva Costa. Um momento por favor.

Existe também um mecanismo para processamento de falhas quando o utilizador se mantém em silêncio ou em situações de erro no reconhecimento. Sempre que uma situação destas é detectada, é tocado um anúncio para avisar o utilizador da ocorrência e é solicitado novamente o nome do colaborador a contactar. A seguir ilustra-se a situação de silêncio e de palavras e expressões fora do vocabulário incluído no sistema (*Out-Of-Vocabulary errors*).

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da PT Inovação. Com quem deseja falar?

<U> (Silêncio)

<S> Desculpe, não ouvi nada. Por favor diga o primeiro e o último nome da pessoa com quem deseja falar.

<U> PT Inovação

<S> Desculpe, não percebi. Por favor diga o primeiro e último nome da pessoa com quem deseja falar.

No caso de detecção de falhas que o sistema não resolva (no reconhecimento de fala ou silêncio por parte do utilizador), a chamada é transferida para um operador humano.

O serviço disponibiliza igualmente um conjunto de comandos universais para flexibilizar a interacção do utilizador. Por comando universal

entende-se um comando que pode ser dito em qualquer ponto do fluxo. No exemplo seguinte ilustra-se a chamada à ajuda do sistema.

<S> Boa tarde. Bem-vindo ao sistema de atendimento automático da PT Inovação. Com quem deseja falar?

<U> Ajuda

<S> Está no serviço de atendimento automático da Portugal Telecom Inovação. Diga o primeiro e o último nome da pessoa com quem deseja falar. Por exemplo, António Silva..

A abordagem seguida para implementar o Gestor de Diálogo deste serviço baseia-se numa máquina de estados (*Finite State Machine*) que controla o diálogo com o utilizador, colocando as questões apropriadas para a recolha das respostas. O utilizador pode em qualquer altura interromper o serviço e empregar uma linguagem mais ou menos elaborada de acordo com a sua preferência. Desta forma a iniciativa do diálogo é mista, ou seja, partilhada pelo interlocutor e pelo sistema. O sistema ignora as respostas que não correspondem ao que foi perguntado, solicitando novamente a informação em falta.

2.6 Mecanismo para carregamento de dados

A informação relativa aos colaboradores, que inclui as extensões a marcar, está armazenada num directório local. Esta informação é obtida a partir do directório da empresa (no caso, a PT Inovação) num formato CSV de acordo com uma *template* cujo formato foi previamente definido para facilitar o processo de importação dos dados para o directório local do serviço.

A figura 3 ilustra o processo de importação e carregamento dos dados dos colaboradores para o serviço ORADIGA.

A informação, em formato CSV de acordo com a *template* pré-definida,

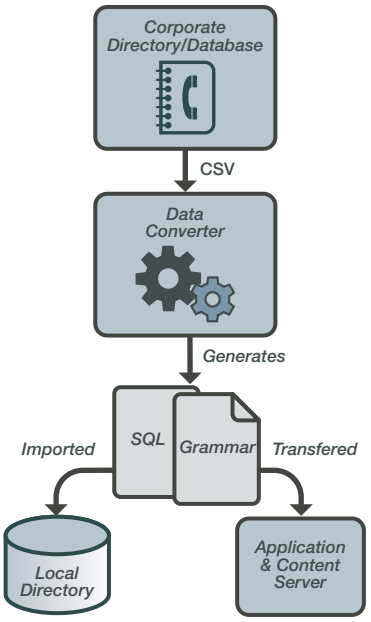

figura 3 – Carregamento de Dados

é processada por um conjunto de *scripts* que geram de forma automática:

> uma *script* em SQL para o carregamento dos dados para o directório local;

> uma gramática base com todos os nomes abreviados, que é transferida e publicada no *Application & Content Server* e utilizada pelo reconhecedor de fala.

2.7 Situação actual

O serviço ORADIGA está actualmente em fase experimental, encontrando-se a plataforma base instalada no Pólo do Porto da PT Inovação.

A integração da plataforma do serviço ORADIGA na PT Inovação, como se pode observar na figura 4, apresenta dois tipos de interfaces: uma interface baseada em TCP/IP que se destina a assegurar o acesso por VoIP e o acesso à *intranet* (para recolha de informação dos colaboradores); a outra interface é estabelecida com o PABX do PLP através de uma linha de extensão analógica.

A configuração actual do ORADIGA concentra num único servidor todos os módulos da solução. O serviço abrange todos os colaboradores da PT Inovação (na sede em Aveiro e nos pólos de Lisboa, Porto e São Paulo) e apresenta uma grande variedade de opções de acesso:

> Acesso PSTN/PLMN Externo – através da marcação de um número do tipo 22XXXXXXX;

> Acesso PSTN Interno – através da marcação de uma extensão (4 dígitos);

> Acesso por VoIP *gateway* – utilizando um telefone convencional e marcando uma extensão, sendo o tráfego de voz transportado na infra-estrutura IP/ATM da empresa;

> Acesso VoIP directo - através de um PC com um *softphone* (ou um telefone IP) registado na plataforma, marcando uma extensão SIP.

Assim que o serviço reconhece o nome do destinatário, as possibilidades de reencaminhamento da chamada também se desdobram. O nome de colaborador pode estar associado aos seguintes números de destino:

> **Extensão analógica –** utilizando a rede fixa convencional, pela interligação dos PABX's da sede da PT Inovação e respectivos pólos;

> **Extensão SIP (Telefone) –** podendo a chamada ser atendida no telefone convencional, contudo, o tráfego será transportado sobre IP (VoIP *gateway*);

> **Extensão SIP (SIP-*Phone*) –** sendo a chamada reencaminhada para um *softphone* instalado no PC do colaborador (ou telefone IP);

> **N.º de telemóvel –** com a chamada a ser reencaminhada para a rede GSM e entregue ao telemóvel do colaborador.

A informação que associa um colaborador ao seu actual número de destino encontra-se armazenada em Base de Dados, podendo ser

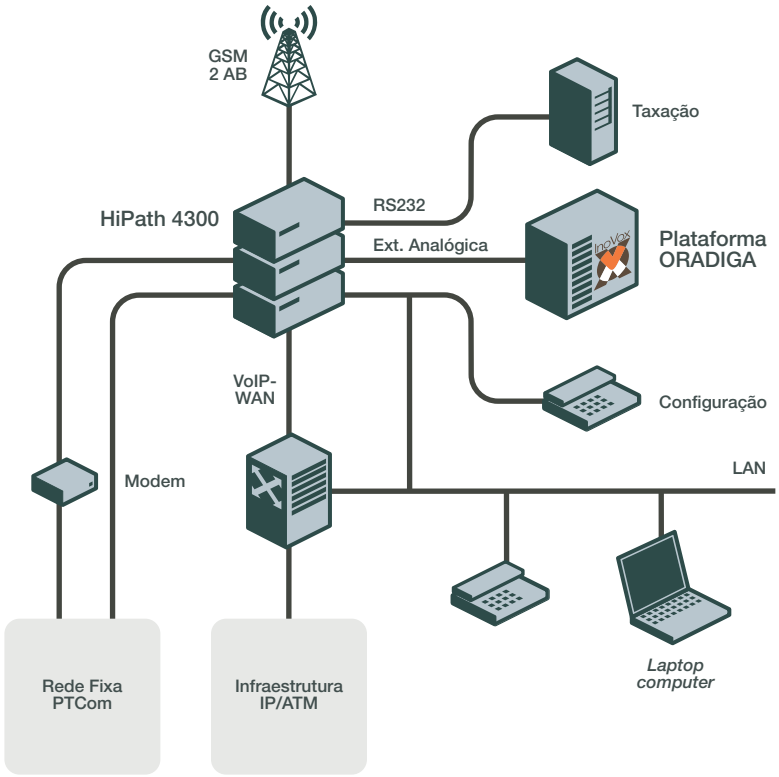

figura 4 – Integração do ORADIGA na PT Inovação

configurada mediante o desejo do colaborador em estar contactável no seu telefone, *softphone* ou telemóvel.

Como já foi referido, está a decorrer um piloto do serviço ORADIGA no Pólo do Porto. Os utilizadores são convidados a utilizar o serviço para contactar os colegas da PT Inovação e a apresentar a sua opinião sobre o mesmo. Aspectos como a facilidade de interacção com o serviço, a navegabilidade, a intuitividade da interface, a robustez do serviço em relação a falhas do sistema, entre outros, são avaliados pelos utilizadores. As sugestões apresentadas pelos utilizadores para a melhoria do serviço, em particular novas funcionalidades, poderão ser incorporadas em versões posteriores do ORADIGA. Para além dos aspectos subjectivos de avaliação também são recolhidos dados objectivos de desempenho do serviço,

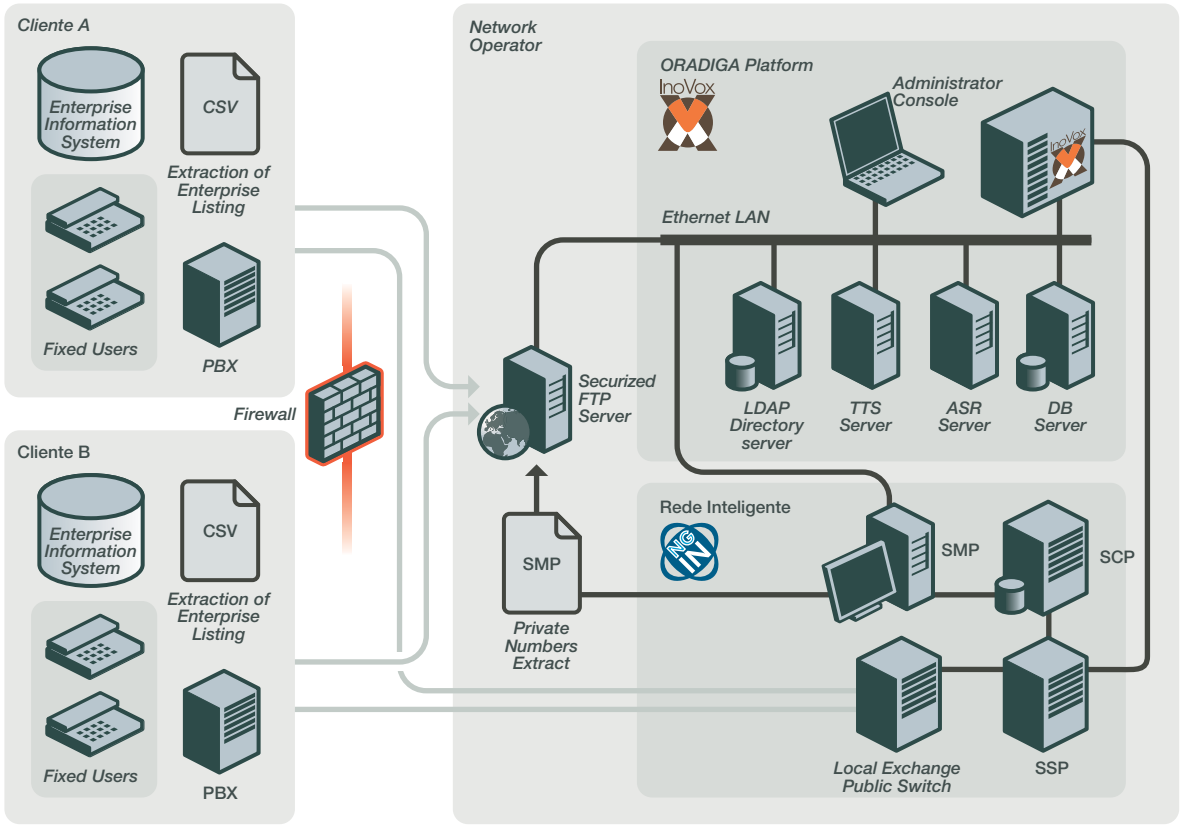

figura 5 – *Integração do ORADIGA num operador*

nomeadamente a percentagem de transacções com sucesso, a duração das interacções, o número e tipo de erros, entre outros.

A curto prazo, o serviço ORADIGA será estendido a toda a PT Inovação.

3. Evolução

A evolução da solução ORADIGA permitirá a flexibilidade necessária para suportar, em simultâneo numa mesma plataforma, o serviço de múltiplas empresas. Num cenário destes, um operador como a Portugal Telecom poderá comercializar a solução, fazendo o alojamento do serviço ORADIGA para pequenas e médias empresas, como ilustrado na figura 5. No limite, ligeiras alterações à solução apresentada poderão permitir a facilidade da marcação por fala para os clientes residenciais, sobretudo os mais idosos.

Futuros desenvolvimentos permitirão a associação do serviço com a agenda do empregado no *Microsoft Outlook* indicando a sua disponibilidade: ocupado, em férias, localizável noutro número de telefone (*follow-me*), com uma mensagem personalizada em caso de impossibilidade de atendimento e permitindo a gravação de uma mensagem do chamador. O serviço poderá ainda ser “personalizado” por um avatar, quando chamado de um terminal com capacidades multimédia, facilitando a criação/associação a uma imagem de marca da empresa.

A utilização da informação de departamento para eliminar ambiguidades de nomes, a marcação por nome e/ou departamento e a evolução do diálogo aproximando-o, cada vez mais, da "fala natural" são possíveis evoluções da solução ao nível da interface de utilizador.

4. Conclusões

Este artigo descreve o serviço de atendimento automático ORADIGA que, a curto prazo, será generalizado à PT Inovação.

O ORADIGA é um serviço baseado na tecnologia de reconhecimento de fala independente do orador, que permite aos clientes a realização de chamadas para os seus contactos através de simples comandos vocais. É intuitivo, de fácil utilização e evita a necessidade de memorizar números de telefone e extensões - apenas um número para toda a empresa qualquer que seja o acesso. Funciona 24 horas por dia, pode gerir um grande número de chamadas telefónicas e liberta os agentes humanos para um nível superior de atendimento aos clientes.

A adopção e divulgação deste tipo de serviços de fala, ao tornar a utilização da tecnologia mais acessível e mais amigável para todos, representa mais um passo para a construção do futuro reflectido na Iniciativa PT*Next* [8] do Grupo PT – " *Um futuro que permite só à distância da voz interagir com tudo o que o rodeia*".

Referências

[1] http://www.w3.org/TR/voicexml20/
[2] http://www.w3.org/Voice/2004/ssml-ir/
[3] http://www.w3.org/TR/speech-grammar/
[4] http://www.speech.cs.cmu.edu/openvxi/
[5] http://www.eclipse.org/
[6] http://www.nuance.com/recognizer/.
[7] http://www.nuance.com/realspeak/
[8] http://ptnext.telecom.pt

Curriculum Vitae

Luís Almeida, licenciatura em Engenharia Electrotécnica/Telecomunicações pela FEUP, Porto, ingressou na PT em 1992 passando a integrar a PT Inovação desde 1999. Tem participado em diversos projectos Europeus estando presentemente a exercer a sua actividade no Departamento de Serviços e Plataformas de Voz, nomeadamente no desenvolvimento de serviços telefónicos interactivos de fala para a plataforma InoVox, para as empresas do Grupo PT, como a TMN e a Vivo.

Alexandre Sapage, licenciado em Engenharia Electrotécnica e Computadores no ramo de Telecomunicações, pela FEUP, Porto. Na PT Inovação, desde 2004, participa no desenvolvimento da plataforma InoVox e na evolução da mesma no âmbito da arquitectura *IP Multimedia Subsystem* (IMS), mais especificamente na área das tecnologias de voz como *Speaker Verification*, *Voice*XML e telefonia IP.

Joaquim Azevedo, mestrado e licenciatura em Engenharia Electrónica e Telecomunicações, pela Universidade de Aveiro. Responsável, no Pólo do Porto da PT Inovação, pelo desenvolvimento e evolução da plataforma InoVox. Tem participado em diversos projectos Europeus na área das tecnologias da fala.

Glória Branco, Engenheira Electrotécnica, com pós-graduação em Análise de Sistemas, ingressou na PT em 1986. Desde 1993 no CET/PT Inovação, trabalhou em projectos nas áreas das Redes Inteligentes e dos Sistemas de Gestão de Redes. Actualmente no PLP, a sua actividade centra-se no estudo da Interacção Homem-Máquina, em particular na usabilidade e acessibilidade de produtos e serviços no âmbito do processamento da linguagem.

Nuno Beires, mestrado em Engenharia de Computadores e Telecomunicações pelo IST, Lisboa. Licenciatura em Engenharia Electrotécnica/Telecomunicações pela FEUP, Porto. Responsável pelo Departamento “Serviços e Plataformas de Voz”, do Pólo do Porto da PT Inovação, que tem desenvolvido vários sistemas e serviços interactivos de fala para as empresas do Grupo PT.

05

Georeferenciação e Telemática na Base da Gestão Remota de Recursos



Miguel Serrano,
Teresa Soares,
Pedro Salvado

palavras-chave:
Georeferenciação,
Telemática, Comunicações,
Gestão Remota, Monitorização.

A associação de informação de localização a eventos e poder-se utilizar essa informação como um valor acrescentado a serviços existentes ou criar serviços inteiramente dedicados a essas temáticas, é algo que tem vindo a conhecer sempre novas oportunidades de mercado de alguns anos para cá. A PT Inovção detém um considerável *know-how* em várias vertentes dos serviços baseados em localização, com particular destaque para a implementação de serviços e plataformas que conciliam tecnologias de localização, comunicações sem fios e telemática para o mercado da gestão remota de recursos.

A importância destas capacidades próprias da PT Inovação é de grande interesse para a nossa autonomia na concepção, desenvolvimento e oferta de serviços ajustados às necessidades deste mercado em permanente evolução.

1. Introdução

A possibilidade de se combinar a telemática com as tecnologias de georeferenciação permitiu o desenvolvimento de um novo tipo de sistema de gestão, focalizado para a gestão dos recursos remotos. Uma das variantes mais comum destes sistemas são os designados Sistemas de Gestão de Frotas (SGF) que possibilitam aos operadores efectuarem uma contínua monitorização da actividade dos seus recursos, associando a esse desempenho a componente geográfica. Deste modo, passa a ser possível uma optimização de esforços nos mais diversos tipos de sistemas de transporte, traduzindo-se numa melhor qualidade de serviço, ao mesmo tempo que se garante uma melhoria de resultados operacionais e de segurança de pessoas e bens.

A conjugação de diferentes tecnologias no âmbito da telemática, intimamente relacionadas com as tecnologias de comunicações e de posicionamento, permite um elevado grau de liberdade na especificação de diferentes tipos de SGF, em função do tipo de actividade das frotas e dos serviços a que servem de apoio e complemento.

2. Georeferenciação

A georeferenciação é a capacidade de associar a bens, eventos ou informação um referencial geográfico. Esta associação permite que a caracterização daquele conjunto de informação potencie inúmeras utilizações em que o carácter espacial poderá ser uma mais valia na percepção e interpretação de fenómenos ou acções.

As vantagens da georeferenciação associada a diferentes tipos de utilização começou a ter a sua importância evidenciada nos anos oitenta do século anterior graças aos Sistemas de Informação Geográfica (SIG). Paralelamente, o aparecimento do sistema de localização por satélite GPS, aproximadamente na mesma época, e sua rápida expansão na comunidade civil, permitiu facilitar em muito a tarefa de atribuir um referencial geográfico absoluto a qualquer tipo de recurso, evento ou informação (figura 1).

O potencial de utilização da informação georeferenciada conheceu um crescimento exponencial a partir do momento em que os SIG se simplificaram e se tornaram mais acessíveis, numa evolução em muito associada à evolução da capacidade de processamento do suporte informático.

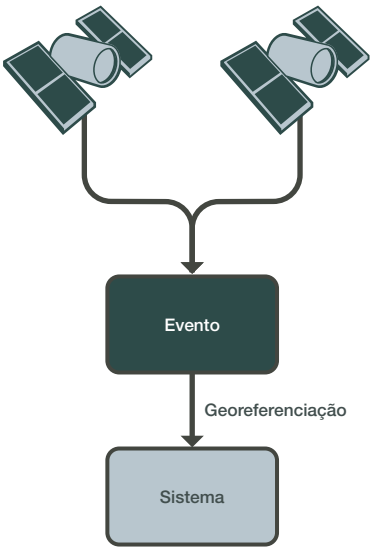

figura 1 – Conceito da georeferenciação de eventos

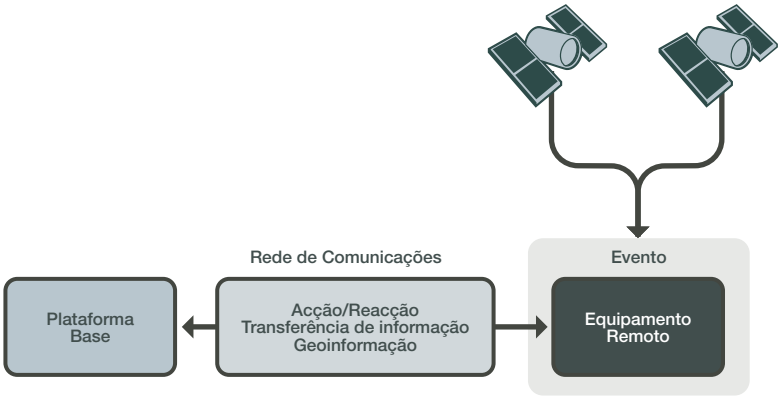

figura 2 – Telemática e georeferênciação

2.1 Tecnologias de localização

Até recentemente, a georeferenciação de informação ou eventos era feita com recurso a técnicas complexas e pouco expeditas ou então com recurso a cartas geográficas. Em qualquer dos casos, não era possível a autoreferenciação do evento. Com o advento do posicionamento por satélite, genericamente designado por GPS, passou a ser possível georeferenciar de um modo simples, económico e eficiente qualquer coisa, para além de permitir a autoreferenciação do evento por modo automático.

Posteriormente, com a rápida expansão das redes de comunicações móveis com o protocolo GSM, a estrutura dessa rede celular passou também a permitir outras técnicas de posicionamento, se bem que menos precisas e com um menor potencial de utilização.

2.1.1 Tecnologias de posicionamento celular

Existem diversas tecnologias de posicionamento com recurso à rede celular, embora os organismos de normalização considerem apenas um conjunto mais restrito de técnicas, de que se podem evidenciar: Posicionamento por Identificação de Célula (*Cell ID*); *Time Difference of Arrival* (TDOA); *Enhanced Observed Time Difference* (EOTD). Cada uma destas tecnologias, por sua vez, possuem variantes que procuram a optimização de resultados com recurso a implementações mais ou menos complexas que por vezes se reflectem em actualizações em infra-estruturas de rede, o que se traduz em custos excessivos.

2.1.2 Tecnologias de Posicionamento por satélite

Em termos das tecnologias de posicionamento por satélite salienta-se o Sistema de Posicionamento Global GPS e o Sistema de Posicionamento por Satélite Galileo, permitindo uma precisão bastante elevada.

O sistema de posicionamento por satélite Galileo é o equivalente europeu para o GPS, com algumas melhorias e um diferente modelo de implementação. Na realidade, o sistema ainda está na fase de desenvolvimento, estando a sua disponibilização no mercado prevista para 2008.

2.1.3 Tecnologias mistas

As tecnologias mistas de localização procuram combinar as vantagens das tecnologias baseadas em satélite e celulares, procurando deste modo minimizar as suas desvantagens. Existem diferentes aproximações para este problema, mas talvez a mais bem sucedida, e com melhores condições de implementação e evolução, é a designada *Assisted GPS* (A-GPS).

3. Telemática

A telemática, no essencial, é um vasto conjunto de tecnologias que permitem a interacção remota entre equipamentos ou plataformas, com recurso a diferentes tecnologias de comunicações, promovendo a troca de dados, acesso a informação e/ou acções de controlo (figura 2).

A diversidade do conceito pode ser representado por dois tipos de telemática bem diversos: num dos extremos está o exemplo da Internet, que permite a interligação de diferentes computadores e noutro extremo estão os casos de diferentes tipos de equipamentos de controlo e monitorização à distância (meteorologia, equipamentos de tráfego, sistemas de portagem, etc.).

Desde o seu início, a telemática baseia-se essencialmente nas comunicações sem fios, utilizando para o efeito diferentes gamas de frequências rádio consoante o tipo de acções remotas a empreender ou a distância a que se tem de operar

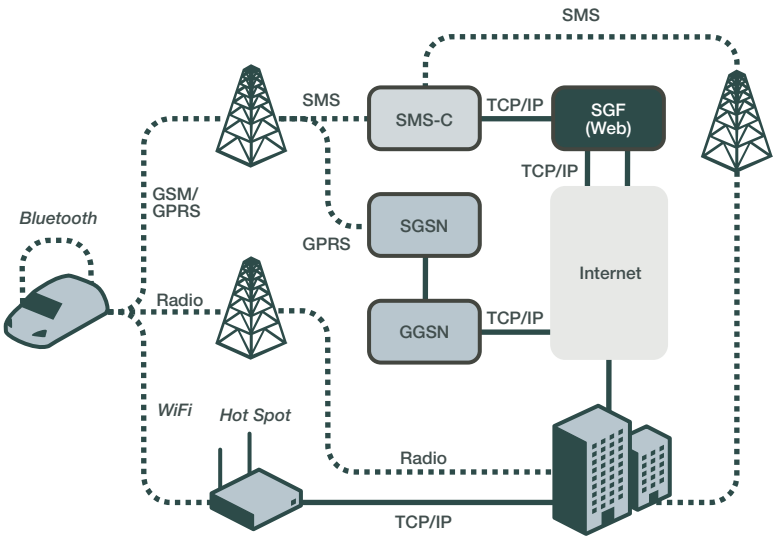

figura 3 – Diversidade de tecnologias de comunicação em telemática

	GSM			Radio			
	GSM	**SMS**	**GPRS**	**Analógico**	**Digital**	**WIFI**	**Bluetooth**
Tipo	Voz/Dados	Dados	Dados	Voz/Dados	Voz/Dados	Dados	Dados
Âmbito	Internacional	Internacional	Internacional	Regional/ Nacional	Regional/ Internacional	Local	Local
Vantagens	> Comunicação voz > Disponibilidade > Diversidade de terminais	> Disponibilidade > Simplicidade > Diversidade de terminais	> Capacidades em dados > Baixos custos > Diversidade de terminais	> Baixos custos de exploração (rede própria) > Disponibilidade > Facilidade de implementação > Diversidade de terminais	> Capacidades em dados > Baixos custos	> Baixos custos > Capacidades em dados	> Baixos custos > Capacidades em dados
Desvantagens	> Custos e limitações em dados	> Custos e limitações em dados	> Disponibilidade variável > Maior complexidade	> Custos elevados de implementação (rede própria) > Custos elevados de exploração (rede alugada /serviços) > Grande diversidade de utilizadores	> Custos elevados de implementação (rede própria) > Dificuldades várias na disponibilização (rede nacional/ rede pública?)	> Âmbito local	> Âmbito local

figura 4 – Compração de tecnologias de comunicação sem fios

com os equipamentos. De um modo geral, existe sempre uma plataforma base que coordena as acções de envio/recepção de dados de diversos equipamentos remotos. Estes equipamentos remotos podem ser estáticos (por exemplo, estações meteorológicas ou equipamentos de apoio à segurança rodoviária) ou equipamentos móveis (a bordo de viaturas, por exemplo). Neste último caso, a necessária ligação entre a plataforma móvel e o equipamento remoto é mais instável, sendo necessário criar mecanismos que assegurem a integridade dos processos entre aqueles componentes.

3.1 Papel das telecomunicações

As comunicações constituem um componente fulcral em qualquer sistema de interacção remota entre equipamentos, pois garantem a necessária integridade de processos. No caso da telemática associada a equipamentos remotos móveis, como é o caso da telemática que suporta os Sistemas de Gestão de Frotas, esse aspecto é ainda mais evidente dado o carácter perturbador nas comunicações, provocado pela mobilidade dos equipamentos alvo.

A evolução das redes móveis celulares GSM veio garantir a necessária qualidade de serviço que permitisse uma operacionalidade da telemática móvel em contínuo, tornando possível a permanente e dinâmica ligação entre os serviços centrais e os equipamentos remotos associados a veículos (figura 3).

A opção pelo sistema de comunicações mais conveniente vai sempre depender de uma série de factores, de onde se pode destacar: a disponibilidade da tecnologia, os custos de utilização, a versatilidade e capacidade de evolução, o desempenho perante o carácter móvel dos equipamentos alvo, o suporte garantido pelo fornecedor dos serviços (figura 4).

4. Gestão remota de recursos

A conjugação de todas as técnicas associadas à telemática e georeferenciação, constituem a base de um Sistema de Gestão de Frotas que, com os seus três principais eixos estruturais constituídos pelos Serviços, as Comunicações e os Equipamentos Remotos, passaram a permitir a resposta em tempo real a uma série de questões fundamentais que permitem caracterizar as acções essenciais numa gestão remota de recursos móveis (figura 5). Estas acções essenciais têm um carácter espacial (Onde?), temporal (Quando?), quantitativo (Quanto?), processual (Como? Que acção? O quê?) e operacional (Quem? O quê?).

4.1 Arquitectura genérica

Um sistema de gestão de frotas pode apresentar diferentes tipos de arquitectura, embora no essencial consista sempre de um Sistema Base que comunica com os Sistemas Remotos a bordo das viaturas, fazendo uso da componente de comunicações.

figura 5 – Versatilidade de aplicações da telemática com a georeferênciação

Cada um destes componentes tem as seguintes particularidades:

> **Sistema base:** é a componente aplicacional que suporta igualmente a interface que permite ao utilizador efectuar todas as acções de monitorização, gestão e configuração das unidades móveis, que embarcam o equipamento móvel. Este Sistema Base pode ser mais ou menos sofisticado na versatilidade das funcionalidades disponibilizadas, dependendo para tal das características dos restantes componentes. O modo como é disponibilizado o acesso ao sistema pode ser mediante uma arquitectura tipo *standalone*, em que toda a componente aplicacional é instalada em máquinas do cliente (monoposto ou em rede) ou em arquitectura cliente servidor, em que o acesso ao Sistema Base é fornecido como um serviço acedido via portal *Web*, em que cada cliente tem um acesso personalizado e seguro;

> **Sistema de bordo:** é constituído pelo equipamento ou unidade de bordo, que integra um *modem* para as comunicações, o receptor GPS, um processador (na maioria dos casos) e um conjunto de interfaces que permitem a integração de equipamentos de terceiros. O sistema de bordo é instalado na viatura associado ao sistema eléctrico e de ignição da mesma, permitindo que a sua utilização se faça mesmo que a viatura esteja desligada. O processador da unidade permite alguma inteligência e capacidade de decisão perante eventos que podem ser programados remotamente pela interface do Sistema Base;

> **Georeferenciação:** a componente de georeferenciação compreende não só a capacidade própria do Sistema Remoto se auto-localizar como também a componente associada à base de dados geográfica e servidor de mapas que suporta a análise espacial do Sistema Base. Como já foi referido, esta componente directamente herdada dos SIG é responsável pela representação gráfica da informação com recurso a mapas digitais. A qualidade desta representação assim como a variedade de análises espaciais possíveis (proximidade, cálculo de rotas, pesquisa de locais) são directamente dependentes da qualidade da base de dados geográfica utilizada (nível de detalhe, estruturação de dados) e do servidor de mapas que a suporta;

> **Interfaces:** é a existência de interfaces num Sistema de Gestão Remota que permite o enriquecimento da versão base do sistema através da integração de outros sistemas existentes (a nível do Sistema Base) e de equipamentos a bordo das viaturas (a nível do Sistema de Bordo). Uma vez que um Sistema de Gestão de Frota, principalmente se considerado numa arquitectura *standalone*, não pode ser considerado como um sistema à parte na orgânica de uma empresa, torna-se fundamental que sejam criados mecanismos de integração. Do mesmo modo, a bordo da viatura é fundamental que o Sistema de Bordo possa integrar informação gerada por equipamentos externos, de modo a esses dados serem acedidos e manipulados do lado do serviço central (figura 6).

4.2 Vantagens e condicionantes

Os Sistemas de Gestão de Frotas têm-se vindo a revelar como uma ferramenta essencial para a optimização da operacionalização de qualquer frota de veículos que se dedique a uma actividade profissional, como é o caso do ramo da distribuição, transportes, recolhas, serviços de emergências, serviços de táxis, comerciais, etc. Esta optimização é conseguida através de uma actualizada gestão dos recursos móveis de uma empresa, potenciando a sua utilização em função da sua distribuição geográfica, conhecimento de *status*, optimização de processos, diminuição de custos de exploração, ao mesmo tempo que a monitorização permite averiguar do cumprimento de tarefas e detecção de anomalias. Estas mais valias reflectem-se em termos da produtividade e da maior segurança a nível dos recursos (materiais e humanos) de produtos transportados (figura 7).

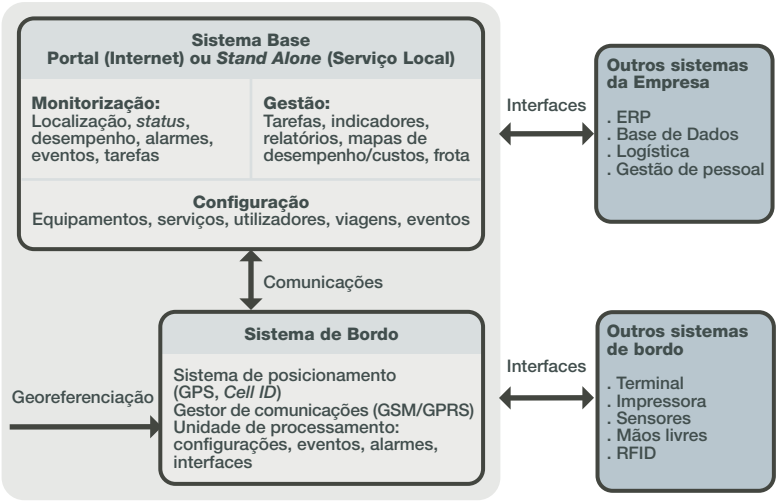

figura 6 – Arquitectura genérica de um sistema de gestão remota de recursos

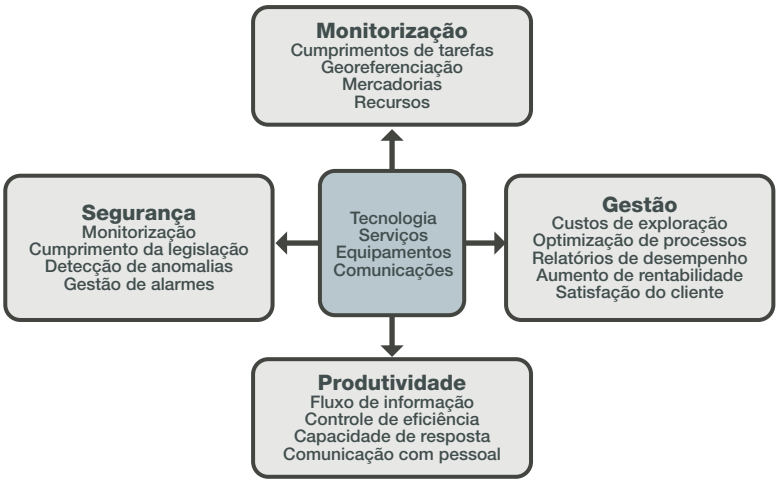

figura 7 – Mais valias de um sistema de adopção de um sistema de gestão remota de recursos

Estes factores levaram organismos internacionais ligados à legislação dos transportes rodoviários a considerarem a obrigatoriedade de adopção de um sistema deste tipo aos operadores de determinados tipos de transporte (matérias perigosas, por exemplo) e a tendência é que esta obrigatoriedade se venha a estender a mais mercados. No entanto, dada a grande variedade de ofertas que o mercado actualmente disponibiliza, também estes sistemas passarão a ter de respeitar regras de normalização de modo a ser possível uma actuação coordenada dos diversos intervenientes.

As questões relacionadas com a privacidade das pessoas não são directamente consideradas neste tipo de sistemas, pelo menos até ao momento, uma vez que o que está em causa é a monitorização de bens materiais e processos e não propriamente de pessoas, ao contrário do que se passa com sistemas de localização pessoal.

4.3 Gestão operacional e logística

Os SGF, para além dos aspectos relacionados com a actividade e movimentação espacial das viaturas, podem também incluir componentes associadas à logística e à operacionalização da frota. Os sistemas mais sofisticados compreendem todas estas componentes, embora seja de considerar que aquelas dependem em certa medida de diferentes modelos de negócio.

A componente da logística compreende a inclusão de todos os processos de criação de tarefas de distribuição, associação de documentação de entregas e facturação, actualização da base de dados directa ou remotamente, etc.

A nível da gestão operacional, esta compreende toda a gestão da parte dos serviços de manutenção da frota, actualização de contratos (seguros, ALD, renovação de frota, etc.) e ainda a gestão da disponibilidade de viaturas, no caso das frotas partilhadas.

5. A PT Inovação na gestão remota de recursos

A experiência da PT Inovação na gestão remota de recursos e em particular na área da gestão de frotas, já vem desde 1999, com o início do desenvolvimento da plataforma SLV (Sistema de Localização de Veículos) e do In@bus, concebido em especial para a gestão de transportes urbanos de passageiros. Posteriormente, foi criada a plataforma Navegante, como sistema de apoio a emergências e mais recentemente enveredou-se igualmente para a produção do próprio sistema remoto.

5.1 Plataforma SLV

O Serviço de Localização de Veículos (SLV) é uma plataforma de gestão remota de veículos, que pode adoptar uma arquitectura *standalone* e Internet. É nesta última versão que a plataforma que suporta o serviço comercial *Frotalink* desde a sua inauguração, em 2001, em parceria com a TMN. Foi uma das primeiras soluções deste tipo a nível europeu. Trata-se de um sistema de gestão e monitorização de veículos, cobrindo todo o território do continente europeu, com o acesso a ser providenciado através de um *login* pessoal e seguro.

Utiliza tecnologia GPS, com recurso alternativo a *Cell ID*, para a localização dos veículos e GSM/GPRS na componente de comunicações entre as unidades remotas e o Sistema Base. Actualmente, o serviço está desenvolvido e implementado em ambiente Linux, com a adopção de ferramentas *open-source* e em respeito pelas normas vigentes (protocolos de partilha e acesso de dados)(figura 8).

Foi recentemente lançada a última versão do sistema, beneficiando de uma série de novidades e evoluções, de que se pode destacar a adopção das comunicações também por GPRS, em complemento do SMS, a utilização do posicionamento celular como alternativa para situações de indisponibilidade de GPS, uma nova interface gráfica, novo servidor de mapas com base de dados geográfica mais completa e um acréscimo de novas funcionalidades, resultantes do potencial gerado pelo novo servidor de mapas e utilização de transmissão GPRS.

De igual importância é o desenvolvimento da componente *Frotalink Mobile*, uma aplicação para correr em PDA que permite a localização do utilizador e fornecimento de indicações de navegação, que é suportado na plataforma SLV.

A plataforma SLV foi igualmente utilizada para suportar a componente de localização e navegação no Projecto Esquadra Século XXI, actualmente em vigor na Esquadra do Estoril, no âmbito do Projecto-piloto.

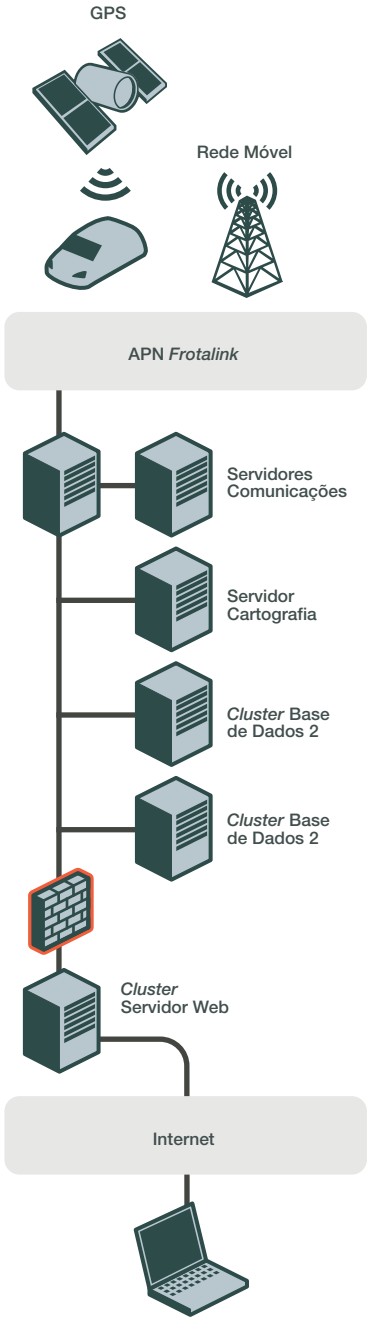

figura 8 – Diagrama operacional genérico do Frotalink

5.2 Plataforma In@bus

O motivo que levou ao desenvolvimento da plataforma In@bus, implementada em Aveiro no sistema Sisgeta, foi a constatação de que as empresas de Transportes Rodoviários têm vindo a sentir dificuldades crescentes em se adaptarem às contínuas alterações e condicionantes da rede viária, em particular nas áreas urbanas, o que contrasta com as maiores exigências da sociedade. A concorrência do transporte privado e a pouca flexibilidade ou ineficiência dos transportes urbanos, geraram as condições para que só um sistema de gestão integrada possa ajudar a solucionar os problemas.

A plataforma In@bus constitui precisamente uma solução integrada para a gestão de frotas de transportes rodoviários, públicos e/ou privados, permitindo a sua monitorização e gestão remota em tempo real, ao mesmo tempo que providencia informação actualizada aos utentes, com recurso a diferentes suportes.

Utiliza tecnologia GPS para obter o posicionamento e GSM/GPRS ou rádio *Trunking* para as comunicações entre os autocarros e o Sistema Base.

No âmbito da prestação de informação aos utentes, esta informação pode ser constituída por actualizações horárias nas paragens da rede de transporte urbano, assim como acesso via Internet a um site específico com informação diversificada no âmbito da utilização da rede de transportes, tal como localização de autocarros, percursos, horários e optimização de percursos utilizando os transportes públicos.

O sistema é desenvolvido em ambiente *Windows* e utiliza uma arquitectura modular com estrutura cliente-servidor, permitindo que seja construído à medida dos requisitos do Cliente, garantindo a sua expansibilidade em fases posteriores. Esta arquitectura permite que o sistema seja implementado de um modo integrado com outros sistemas da Empresa cliente, permitindo sessões multi-utilizadores nos seus diferentes componentes.

A concepção do In@bus obedeceu à lógica de estruturar os componentes em módulos que agrupassem funcionalidades relacionadas entre si, com a vantagem de simplificar os processos de segurança na utilização do sistema, através de diferentes perfis de utilizadores.

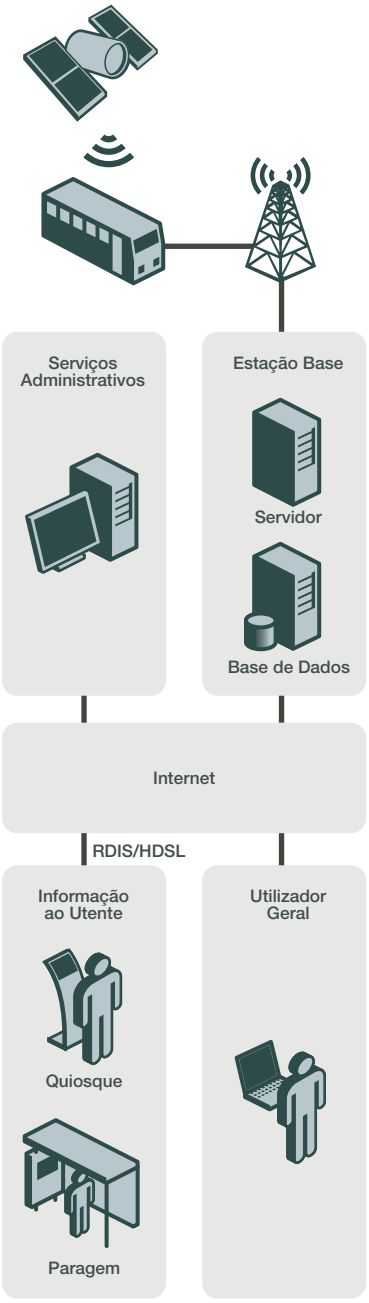

figura 9 – Diagrama operacional genérico do In@bus

5.3 Plataforma Navegante

O Sistema de Apoio a Emergências Navegante constitui uma solução integrada para o apoio a situações de emergência que exijam navegação em tempo real para o local da ocorrência. O sistema consiste num Sistema Central, que é uma base localizada numa entidade coordenadora, que permite a localização, monitorização e gestão dos veículos e respectivos *status* de operacionalidade em tempo real e nas Unidades Remotas. Estas Unidades são os veículos de emergência com equipamentos instalados a

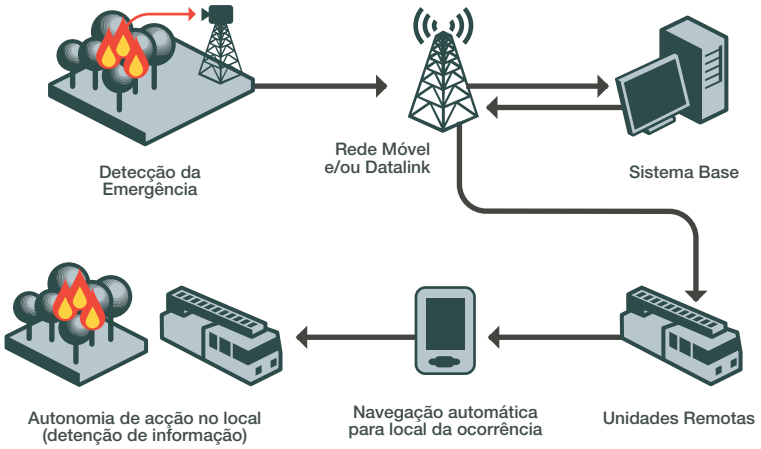

figura 10 – Diagrama operacional genérico do Navegante

bordo que permitem, entre outras tarefas, a sua navegação automática para o local da emergência, assim como para outros pontos de importância na assistência a que procedem (ex: hospitais, pontos de água, etc.).

O objectivo principal do sistema é optimizar o tempo de resposta das unidades de emergência, através da disponibilização imediata do melhor percurso para chegar ao local em função do tipo de viatura e do tipo de vias de comunicação. Ao mesmo tempo, permite no Sistema Central efectuar o acompanhamento e gestão remota dos efectivos afectos a uma ou mais ocorrências, optimizando o desempenho das equipas no terreno (figura10).

Este sistema encontra-se implementado no Parque Natural do Alvão desempenhando acções de gestão e de optimização de resposta a situações de incêndio e acidentes, tendo estado também instalado no INEM, em versão piloto, integrado no processo de gestão e optimização de resposta das VMER (Viatura Médica de Emergência e Reanimação). Actualmente decorre no P.N. do Alvão um projecto piloto com a integração do Navegante com um sistema automático de detecção de incêndios florestais, procurando-se deste modo completar o ciclo detecção, intervenção e gestão da emergência.

6 - Conclusões

O potencial da integração da telemática com a georeferenciação é em pleno demonstrado com o sucesso que as soluções de SGF têm vindo a registar a nível mundial, permitindo que os operadores de transportes possam ter ferramentas eficazes para uma gestão optimizada dos seus recursos.

Deste modo, não surpreende que as perspectivas de crescimento no mercado dos SGF sejam promissoras, não só devido às próprias características dos sistemas e benefícios óbvios para os seus utilizadores, como também devido à crescente regulamentação do sector de transportes na área de influência da Comunidade Europeia. Esta regulamentação visa uma maior interoperabilidade entres os diversos actores do mercado (fornecedores de soluções e serviços SGF, equipamentos de bordo, tecnologias de comunicações e de localização, operadores de transportes, fabricantes de material circulante, etc.) e a sua ligação a outros tipos de serviços.

A PT Inovação encontra-se particularmente bem posicionada em Portugal para ser um participante activo no desenvolvimento e concepção de soluções SGF, devido à sua experiência neste ramo e também por ter uma capacidade de intervenção e criação diversificada, desde a área das comunicações até à produção de *hardware* específico, em cumprimento das normas vigentes.

Carlos Miguel Pereira Serrano, nascido em 1962, licenciado em Engenharia Geológica em 1988 pela Universidade de Aveiro, com Pós-Graduação em 'Produção Integrada de Geoinformação' em 1990 e Grau de Mestrado em 'Processamento Digital de Imagens de Satélite' em 1992, ambos no Instituto ITC, em Enschede, Holanda. Actualmente, é responsável de Unidade Funcional Aplicações de Localização na área de negócios Aplicações para a Sociedade da Informação (ASI). Anteriormente colaborou nas Áreas de Negócio Sistemas de Suporte à Gestão de Redes (SSG) e Serviços de Redes Móveis (SRM), em ambas colaborando nas tarefas de gestão de projectos e na concepção e integração de soluções. Em actividades anteriores, exerceu funções na Universidade de Aveiro, como professor convidado do Curso de Engª Geológica, formador em Pós Graduação em Sistemas de Informação Geográfica e projectos na área da informação geográfica e análise de imagens de satélite.

Maria Teresa Soares Costa, licenciada em Engenharia Electrónica e Telecomunicações pela Universidade de Aveiro em 1982. Ingressou no CET em 1982, tendo desempenhado diferentes cargos e funções na área do desenvolvimento de Sistemas de Comutação Digital. Em 1999, após a constituição da PT Inovação, foi responsável por uma Unidade Funcional na área de Serviços e Plataformas para Redes Móveis, exercendo funções ao nível da gestão de projectos e na concepção e integração de soluções, com especial ênfase para a área dos sistemas e soluções de Localização. Actualmente, é responsável pela área de negócios Aplicações para a Sociedade da Informação (ASI). Desde 2003, acumula funções de Gestora de Cliente TMN.

António Pedro Harrisson Salvado, licenciado em Engenharia Electrónica e de Telecomunicações pela Universidade de Aveiro, em 1983, desde então a trabalhar no Centro de Estudos de Telecomunicações em Aveiro (agora PT Inovação, SA). Entre 1991 e 1993 responsável pela Unidade Funcional de Protocolos integrada no departamento de desenvolvimento de Produtos de Comutação. Entre 1992 e 1996 responsável pelo Bloco Funcional – Sistema de Comutação, pela orientação do projecto nacional ELD na sua fase de desenvolvimento comercial, e pelo projecto SICOM – Sistema Integrado de Comunicações Militares. Entre 1996 e Julho de 1998 responsável pelo Bloco Funcional – Redes e Protocolos no âmbito da Área Funcional Gestão de Redes. De Julho de 1998 a Dezembro do mesmo ano, Director de Gestão de Redes e Serviços. De Janeiro a Junho de 1999 Responsável pela Harmonização de Processos e Qualidade na DID/CET. Desde Julho de 1999, e com a criação da PT Inovação, SA, Director de Serviços e Redes Móveis.

06

DiNO (Download Innovation) Plataforma de Gestão e Entrega de Conteúdos




Helder Biscaia, João Ricardo Silva, Bruno Cabral, Victor Batista (Present Technologies)

palavras-chave:
DiNO, Gestão de Conteúdos, MIDP, OMA *Download*, DRM, SMS, MMS, OTA.

Com a demanda por novos serviços de dados, os fornecedores de conteúdos e as operadoras são obrigados a aumentar a complexidade dos serviços e dos conteúdos a fornecer, de forma a suportar a vasta gama de terminais móveis existente. Um destes serviços é a gestão e entrega de conteúdos para terminais móveis. No entanto, dada a variedade, capacidade, compatibilidade e características dos terminais existentes no mercado, é fundamental que a plataforma que forneça este tipo de serviço contemple uma série de funcionalidades complexas que permitam que os clientes não tenham de estar familiarizados com pormenores técnicos de forma a receberem o conteúdo desejado.

Este artigo apresenta a plataforma DiNO e descreve algumas das suas funcionalidades cruciais que têm permitido que esta seja utilizada na Vivo, como base para os serviços.

1. Introdução

As operadoras de telecomunicações móveis estão continuamente à procura de novos serviços que as possam destacar das suas concorrentes e assim aumentar as suas receitas. Um tipo de serviços que tem contribuído para estes objectivos e que tem agradado simultaneamente às operadoras e aos próprios clientes é o que permite a pesquisa e entrega de conteúdos de valor acrescentado directamente ao terminal do cliente.

A plataforma de gestão e entrega de conteúdos da PT Inovação vai ao encontro da necessidade de harmonizar e unificar o processo de gestão, armazenamento e entrega de conteúdos. Assim, esta plataforma disponibiliza conteúdos para *download* directo ou indirecto para os clientes da operadora. O acesso à plataforma é garantido através da disponibilização de interfaces genéricos; actualmente, os clientes podem adquirir conteúdos através de um dos vários canais disponíveis, entre os quais se destacam: WAP, *Web*, SMS, IVR e plataformas externas.

A plataforma é capaz de gerir vários tipos de conteúdos entre os quais se destacam os seguintes:

> Imagens;

> Sons (toques monofónicos, toques polifónicos, *real tones*);

> *Clips* de áudio, incluindo músicas MP3 ou outros formatos;

> *Clips* de vídeo;

> Conteúdos áudio e vídeo para entrega via *streaming* RTSP;

> Jogos Java, Symbian, Smartphone.

A entrega de conteúdos por métodos de *streaming* requer a utilização de plataformas suplementares e específicas para este fim. Mesmo assim, não existem restrições quanto aos formatos suportados. Apenas a título de exemplo, a plataforma consegue gerir os seguintes formatos:

> Sons/*clips* de áudio: AMR-NB, AMR-WB, AAC, WAV, MP3;

> *Clips* de vídeo: MP4, Windows-Media, 3GPP, H.263, RealMedia, Quicktime.

Actualmente, a plataforma constitui o *core* de alguns serviços de grande sucesso como é o serviço de Toques & Imagens e Jogos Java na TMN e os serviços inovadores da Vivo conhecidos como VivoPlay e VivoDownloads.

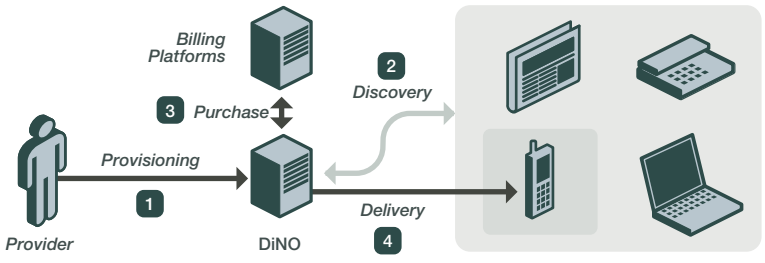

figura 1 – Gestão de conteúdos (etapas)

2. Funcionamento

Descrito de uma forma muito simples, a plataforma DiNO funciona como um repositório genérico de conteúdos que simultaneamente gere o acesso e entrega destes aos respectivos utilizadores. Assim, existem várias etapas pelas quais um conteúdo tem de passar e que um cliente tem de efectuar para que a aquisição e a entrega do mesmo seja concluída com sucesso. Estas etapas estão ilustradas na figura 1 e são descritas em maior pormenor nas secções posteriores.

2.1 *Provisioning*

O aprovisionamento de conteúdos é efectuado directamente pelo fornecedor de conteúdos. Existem dois mecanismos que permitem a inserção de conteúdos; via *browser* ou processos *batch*. Durante o processo de aprovisionamento (*provisioning*) é necessário incluir informação adicional sobre o conteúdo como: tipo de conteúdo, categorias, datas de disponibilização, mecanismos DRM a aplicar, preços, etc. O processo de aprovisionamento passa pela identificação ou inserção do binário ou *stream* que deve ser entregue ao cliente. Isto é fundamental para que o conteúdo a entregar seja efectivamente compatível com o terminal e que, no caso de *streaming*, seja entregue o conteúdo que melhor se adeque às características da rede de acesso.

2.2 *Discovery*

O termo *Discovery* refere-se ao processo que é utilizado pelo cliente para descobrir quais os conteúdos que mais lhe interessam e facilitar o processo de aquisição destes mesmos. Actualmente a plataforma DiNO está munida com capacidades que lhe permite disponibilizar interfaces WAP e *Web* aos clientes. Mais, a plataforma tem mecanismos de exportação do catálogo que permitem que o processo de *discovery* seja intermediado, por exemplo, através de IVR ou outros meios de divulgação: revistas, televisão, jornais, etc.

2.3 *Purchase*

Purchase designa o processo de compra de um determinado conteúdo. Na realidade, um cliente adquire apenas uma licença que o autoriza a visualizar ou a fazer *download* do conteúdo em causa. A plataforma DiNO está preparada com mecanismos de *online* e *offline charging,* o que permite um maior controlo à operadora. Sendo o DiNO uma plataforma da PT Inovação é óbvio que existem mecanismos de interacção com o serviço de OSA *charging* do NGIN; no entanto, a plataforma DiNO tem a capacidade de dialogar com outros sistemas externos para efeitos de *charging* e *billing*.

2.4 *Delivery*

O processo de *Delivery* (ou entrega do conteúdo) é um processo bastante complexo e essencial para que o cliente receba o conteúdo pretendido, compatível com o seu telemóvel e transferido via um protocolo que o terminal reconheça. O processo de *Delivery* também está munido com funcionalidades que permitem a recepção de eventuais notificações provenientes do terminal. Estas podem confirmar a recepção e instalação do conteúdo, ou em caso de erro, a razão que originou a falha. Por fim, é este processo que dá indicação à plataforma DiNO que a licença atribuída a um cliente para um determinado conteúdo deve ser retirada do sistema.

3. Arquitectura

Conhecido o funcionamento geral da plataforma DiNO, descrito na secção anterior, apresenta-se na figura 2 um diagrama com a arquitectura funcional da plataforma.

A plataforma DiNO é constituída por diversos módulos independentes que lhe conferem uma escalabilidade e flexibilidade notável, estando por isso preparada para as constantes evoluções na área da gestão de conteúdos.

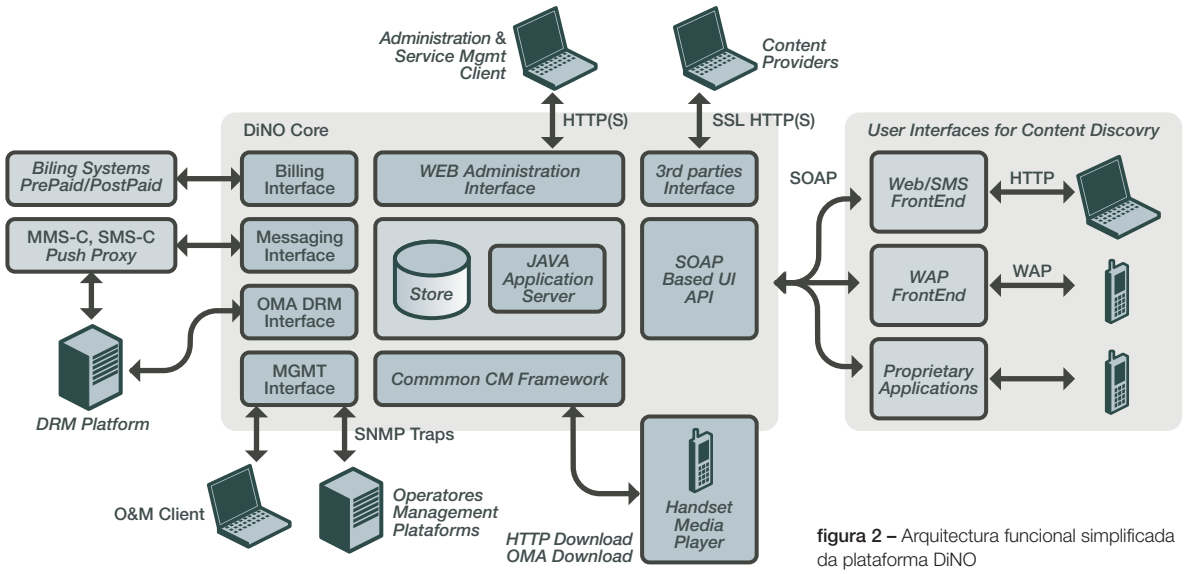

figura 2 – Arquitectura funcional simplificada da plataforma DiNO

O *core* do DiNO fornece todos os mecanismos base que permitem, de uma forma genérica e transversal:

> Autenticação e autorização do acesso dos clientes às interfaces de *discovery*;

> Disponibilização de interfaces ou API (internas e externas) para:

  > *Discovery* de conteúdos e navegação inteligente na árvore de categorias;

  > *Delivery* de conteúdos, das mais diversas formas;

  > Administração da plataforma e do catálogo de conteúdos;

  > Geração de indicadores e dados estatísticos;

  > Geração de eventos de *billing*;

  > Processamento automatizado de categorias de sistema (categorias de novidades, top, destaques, ...);

  > Aprovisionamento de utilizadores;

  > Gestão de licenças de utilização dos conteúdos;

  > Gestão do álbum pessoal dos clientes;

  > Gestão do *revenue share*;

  > Manipulação do repositório (*file system*) dos conteúdos;

  > Comunicação com plataformas externas (*service monitoring, messaging centers*, DRM, LDAP, *mobile portals*, ...)

  > Funcionalidades OMA e DRM

  > [...]

> *TransactionManager* e *BillingInterfaces*: gestão de compras e ofertas, regras de negócio, campanhas, tarifação (*online*/*offline*, interna/externa), entrega de licenças, segurança.

Sobre o *core* assentam diversos componentes que tornam possíveis todas as operações que se esperam de uma plataforma de gestão e entrega de conteúdos:

> *Administration & Service Mgmt Client*: interface *Web* de administração da plataforma e dos serviços/conteúdos nela disponibilizados;

> *Content Provisioning*: interfaces para a automatização do processo de aprovisionamento e publicação de conteúdos por parte do fornecedor ou operadora;

> *Content Discovery*: interfaces para *discovery* de conteúdos pelo *end-user, Web*, WAP, SMS, IVR, etc. Podem ser internas à plataforma ou externas (via SOAP-*based* API);

> *CM & Delivery Framework:* para a gestão integrada e *delivery* inteligente de conteúdos;

> *Operators Mgmt Platforms*: sistemas para monitoria dos restantes componentes e serviços, garantindo uma elevada qualidade e alta disponibilidade dos mesmos;

> DRM *Platform*: sistemas de gestão de DRM;

> *Billing Platforms*: plataformas para *billing online* e gestão de subscrições associadas às transacções no DiNO.

4. Funcionalidades Adicionais

A plataforma DiNO está munida com um conjunto diverso e alargado de funcionalidades que assumem uma importância enorme. Identificar e descrever todas as funcionalidades da plataforma está fora do escopo deste artigo. Assim, de seguida, são apresentadas apenas algumas destas funcionalidades características desta plataforma:

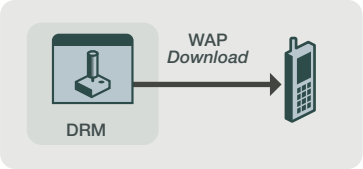

figura 3 – *Forward Lock*

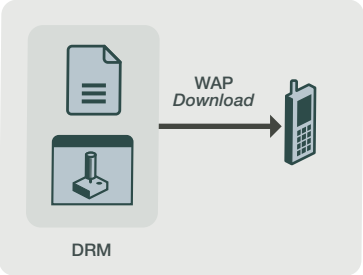

figura 4 – *Combined Delivery*

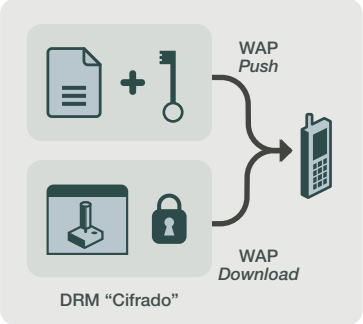

figura 5 – *Separate Delivery*

4.1 Protecção dos conteúdos

Digital Rights Management

A plataforma DiNO está munida com mecanismos internos de *Digital Rights Management*. No entanto, na eventualidade do cliente possuir uma plataforma genérica de DRM, então é possível que a plataforma DiNO utilize as funcionalidades disponibilizadas por esta. Como solução DRM de referência, o DiNO utiliza soluções da CoreMedia AG para *licensing* e *packaging servers* 100% OMA DRM compatíveis, incluindo a nova versão 2.

Assim, a plataforma DiNO é totalmente compatível com o OMA DRM

1.0 e com os seguintes casos de uso especificados pelo *Open Mobile Alliance* [1]: *Forward Lock, Combined Delivery* e *Separate Delivery.*

Forward Lock
Forward Lock é o mecanismo mais simples de OMA DRM (figura 3). Conteúdos entregues com esta forma de protecção contêm informação adicional que é utilizada pelo terminal para que este não permita que o conteúdo seja reenviado para outros utilizadores ou dispositivos (p. ex. por IrDA ou MMS) e que seja dele extraído por qualquer meio. Assim, o conteúdo solicitado pelo utilizador é enviado para o seu terminal (não cifrado) e pode ser executado tantas vezes quantas as pretendidas.

Combined Delivery
O mecanismo de *Combined Delivery* (figura 4) é semelhante ao *Forward Lock* já que o conteúdo continua a ser enviado num formato não cifrado. No entanto, conteúdos enviados com este tipo de protecção contem informação adicional sobre a forma como o conteúdo deve ou não ser utilizado. Esta informação adicional é conhecida como a licença de uso e é da responsabilidade do terminal possibilitar o uso do conteúdo de acordo com a respectiva licença e regras de utilização.

Separate Delivery
Em OMA DRM 1.0, a forma mais segura de proteger os conteúdos é conhecida por *Separate Delivery*. Com este tipo de protecção o conteúdo é cifrado antes de ser enviado para o terminal do utilizador. Na verdade, o conteúdo cifrado pode ser reenviado para qualquer terminal, o que permite a possibilidade de recorrer a um maior leque de modelos de negócio. Para que o terminal consiga decifrar o conteúdo necessita de ter em sua posse uma licença de utilização onde esteja incluída a chave que permite que o conteúdo seja decifrado. No entanto, é da responsabilidade do terminal impedir que o conteúdo seja utilizado de uma forma que não esteja de acordo com a respectiva licença e regras de utilização.

4.2 Acesso por Licença
O direito de aceder a um conteúdo na plataforma DiNO é gerido internamente através de um mecanismo de licenças. Uma licença confere o direito a um utilizador com determinado MSISDN de aceder a um conteúdo específico. Uma licença é criada depois de comprado o conteúdo. Devido às variadas utilizações pretendidas para a plataforma DiNO, o mecanismo de licenças é flexível nos direitos conferidos para uma única compra:

> **Licença por Número de *Downloads*** – Uma licença deste tipo permite N acessos ao conteúdo para a mesma compra. Geralmente aplicável a conteúdos do tipo Jogos ou Toques&Imagens;

> **Licença por intervalo de tempo –** Este tipo de licença permite um número infinito de acessos ao conteúdo, por um intervalo de tempo iniciado no instante da compra. Este tipo de licença é útil para acesso a conteúdo que é *streamed*.

Outra característica do DiNO é que permite vender um conteúdo em diferentes vertentes. Este conceito é designado de Oferta. Uma oferta é composta por:

> O conteúdo a vender;

> Tipo de licença;

> Preço.

Desta forma, é possível disponibilizar um jogo Java, por exemplo, da seguinte forma:

> Oferta 1: Jogo xpto, *Download* 1 vez, 3 eur;

> Oferta 2: Jogo xpto, *Download* 5 vezes, 10 eur.

A plataforma DiNO tem mecanismos que "protegem" o cliente e impedem que a licença para uma determinada oferta seja adquirida se já possuir uma licença de uma outra oferta para o mesmo conteúdo.

De realçar que as licenças internas do DiNO são independentes do DRM, isto é, uma licença define regras de acesso ao conteúdo na plataforma. Regras de DRM definem regras de utilização do conteúdo no terminal onde este foi instalado.

4.3 *One Time* URL
O OMA DRM 1.0 e o mecanismo de gestão de licenças do DiNO permitem a protecção contra acessos directos aos conteúdos da plataforma que utilizam os métodos de entrega internos.

Nos casos em que o DiNO integra com plataformas externas para a entrega de conteúdos, como é o caso dos conteúdos para *streaming*, é implementado um mecanismo de protecção contra acessos directos, com a introdução de um *timestamp* na criação do *url* de acesso, para limitar a sua validade, e de um parâmetro cifrado com base no *url* para o conteúdo e numa chave privada para garantir a autenticidade do *url* de acesso na plataforma externa.

5. Tecnologia e Inovações
5.1 Base de Dados vs. Sistema de Ficheiros
Uma plataforma de gestão e entrega de conteúdos, capaz de disponibilizar dezenas de tipos diferentes de *media* (totalizando em média cerca de 20.000 novos conteúdos por ano[1] e cerca de 100.000 binários, não contabilizando os

[1] Número médio de conteúdos aprovisionados na plataforma por ano pelas diferentes operadoras.

previews), terá que garantir que a performance não é afectada. De notar que a existência de cerca de 300 terminais diferentes no mercado implicará aproximadamente 1 milhão de diferentes associações binário-terminal.

Assim, em DB é guardada apenas a informação sobre o conteúdo, e não o conteúdo em si: metadados, binários do conteúdo e associação com os terminais.

Neste contexto foi desenvolvido um mecanismo no *core* do DiNO que, a partir de informação mínima na DB ou em *cache* (ver capítulo sobre Desempenho do *Discovery* e JBossCache), permite a pesquisa instantânea de binários no *file system*.

Esse mecanismo, usado em larga escala pelos componentes de *provisioning*, *discovery* e *delivery* de conteúdos, assenta num algoritmo de *file hashing*.

5.2 Desempenho do *Discovery*

O *discovery* de conteúdos, ao contrário do que parece, pode revelar-se muito complexo. Existem diversas variáveis em jogo:

> **Interfaces de *discovery*:** existem *n* interfaces de *discovery* distintas (WAP, *Web*, etc). As categorias e os conteúdos podem estar acessíveis (visíveis) em todas as interfaces ou apenas num *sub-set*. É tipico disponibilizarem-se determinados conteúdos numa interface de *discovery* específica, fomentando a utilização desta;

> **Tipos de conteúdo**: as categorias podem estar associadas a um ou mais tipos de conteúdo; o conteúdo está sempre associado a um e um só;

> **Categorias e conteúdos:** adicionalmente, as categorias e os conteúdos têm um estado (activo, por publicar, inactivo, apagado, ...);

> **Meio de *delivery*:** um conteúdo pode estar associado apenas a determinado meio de *delivery*. Exemplo: OMA *Download*. Se o terminal não suportar OMA *Download* não deverá ver o conteúdo;

> **Binários do conteúdo:** um conteúdo terá vários binários, cada um destes com um estado (activo, por publicar, inactivo, apagado, ...). Normalmente, para o mesmo conteúdo, o fornecedor tem vários binários, de acordo com o terminal ou conjunto de terminais que os suportam. Exemplo: binário do jogo "FIFA 2007" para os Nokia N-series é diferente do binário para os SonyEricsson;

> **Redes e famílias de rede:** como complemento do ponto anterior, para o mesmo conteúdo/terminal pode existir mais do que um binário. Isto aplica-se nas situações em que o conteúdo entregue varia de acordo com a rede do cliente. Exemplo: determinado vídeo entregue por *streaming* para o Nokia N-80 tem uma qualidade distinta que depende da rede onde o cliente está registado: GPRS, UMTS, HSDPA.

A vantagem da operadora optar pelos *discoveries* internos à plataforma, para além das questões como a segurança, a flexibilidade e uma interacção mais directa com o *core*, é a garantia dada no que diz respeito à sua performance: independentemente do número de parâmetros com os quais o módulo de *discovery* tem que lidar, o processo de *discovery* (cada *hit* por parte do cliente) concluirá na ordem das dezenas de milisegundos, um valor que não poderá nunca ser ultrapassado, principalmente se se tratar de um *discovery* WAP.

Para além do tempo de resposta aos pedidos HTTP do cliente, há vários princípios que são seguidos de forma a maximizar a experiência de utilização do portal:

> O cliente nunca vê categorias vazias;

> O cliente apenas vê conteúdos suportados pelo seu terminal (de acordo com as características na DB de *User-Agent Profiles* do DiNO), associados à interface de *discovery* e à rede de acesso em causa e para o tipo de conteúdo seleccionado.

Em termos macro, tal é conseguido através de um processo responsável por:

> Criar e gerir uma estrutura em *cache*, dos dados necessários para um *discovery* eficiente, minimizando o número de pedidos ao *backend* (DB) da plataforma. Esta estrutura é refrescada frequentemente: qualquer alteração que tenha havido na DB (novo conteúdo aprovisionado) é detectada através da análise de deltas e a estrutura consequentemente actualizada. A *cache* é progressiva e conteúdos raramente ou nunca acedidos são removidos da memória;

> Manter uma matriz auxiliar *light-weight*, que permite analisar em tempo praticamente nulo que conteúdos deverão ser mostrados em determinado nível da árvore de navegação, qualquer que seja a combinação das variáveis identificadas em cima (terminal que acede, categoria e tipo de conteúdo seleccionado, interface de *discovery* acedida, formas de *delivery* suportadas, rede de acesso e binários associados).

Esta última estrutura auxiliar consiste basicamente num conjunto de *bitsets*, um por cada variável a considerar no processo de *discovery*, formando assim uma matriz multi-dimensional.

Consequentemente, o *discovery* dos conteúdos é conseguido através de simples operações *bitwise* sobre essa matriz, que em memória são imediatas. Os cálculos são feitos *on-the-fly* (a cada pedido do cliente), com base nos valores dos parâmetros acima, não havendo necessidade de qualquer consulta à DB, o que é significativo num cenário com cerca de 1 milhão de associações entre binários e terminais.

5.3 JBossCache

Como complemento ao mecanismo referido no ponto anterior, a plataforma tem a capacidade de interagir com um *framework* do servidor aplicacional – JBossCache – criando uma sinergia adicional entre o portal (interfaces de *discovery*) e a interface de administração da plataforma. Qualquer alteração ao nível da árvore de categorias ou conteúdos feita pela operadora irá reflectir-se, de imediato, no portal do lado do cliente.

Com JBossCache é possível centralizar o processo descrito acima, permitindo minimizar os recursos necessários para manter as estruturas em memória (poderão estar distribuídas) e a carga na DB.

6. O Futuro da Plataforma

Actualmente, encontra-se em exploração a versão 3.1 da plataforma DiNO. No entanto, a plataforma está em contínua evolução e no futuro pode estar munida com funcionalidades adicionais que incluem, mas não limitadas a:

> Gestão de *Ring Back Tones* e outros tipos de media emergentes;

> Gestão de Concursos e subscrições;

> Perfis de utilizador e personalização;

> Conceito de comunidades para partilha de conteúdos;

> Integração com plataformas de *mobile gaming*, *mobile advertisement*, gestão de *high-scores;*

> Utilização de *voucheurs*;

> *Cross-selling* para incentivar aquisições de conteúdos adicionais.

Referências

[1] http://www.openmobilealliance.org
[2] "Plataforma de Gestão e Entrega de Conteúdos" DiNO Whitepaper v2.0, Maio 2006.
[3] DiNO - Manual de Utilização, v2.0, Julho 2006

Helder Biscaia, detém uma licenciatura em *Electrical Computer and Information Science* pela Universidade de Guelph no Canadá e um mestrado em Engenharia Electrónica e Telecomunicações pela Universidade de Aveiro. Ingressou na PT Inovação em 1991 onde participou em vários projectos Europeus nos programas RACE e ACTS. Em 2003, juntou-se ao departamento de Serviços de Redes Móveis e é hoje responsável pela unidade de Plataformas de Conteúdos e Internet móvel onde são desenvolvidas plataformas de serviços de valor acrescentado entre as quais se destacam a M-WiPEP, DiNO, Serviços de *Streaming* e MM-Box.

Bruno Cabral, licenciado em Engenharia Informática e Computação pela Faculdade de Engenharia da Universidade do Porto, participa, desde 2003, em diversos projectos na área de Serviços e Redes Móveis da PT Inovação, desempenhando funções sobretudo de engenharia de *software*, como a arquitectura e o desenvolvimento orientados à Internet móvel. Assumiu em 2005 a responsabilidade técnica pela plataforma MWiPEP (*Multimedia Wireless Internet Portal*), que suporta actualmente o portal da maior operadora de telecomunicações móveis em Portugal. Contribui igualmente no design, desenvolvimento e evolução de plataformas de valor acrescentado, destacando-se as plataformas de gestão e entrega de conteúdos MMBox e DiNO.

João Ricardo Silva, licenciado em engenharia Electrónica e Telecomunicações pela Universidade de Aveiro no ano de 2001 e MSc em Telecomunicações pela Queen Mary University of London em 2004. Inicio da actividade na PT Inovação em 2001 com estágio profissional na área de "Plataformas de *Billing* para redes UMTS *all* IP". Desde 2002 desenvolve funções na área de serviços para redes móveis, nomeadamente na concepção e desenvolvimento da plataforma M-WiPEP, serviços de *streaming* video para redes móveis 2G/3G e plataforma DiNO de gestão de conteúdos.

Victor Nelson Marques Batista, licenciado em Engenharia Informática pelo Departamento de Engenharia Informática (DEI) da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, em 1998. De 1998 a 2000 trabalhou como investigador no Grupo de Sistemas Confiáveis (DSG) do mesmo departamento. Em Abril de 2000 foi sócio fundador da empresa Present Technologies, onde se encontra actualmente. Em 2003 concluiu uma pós-graduação em Engenharia Informática na FCTUC. No ano de 2005 tornou-se certificado como "JBoss Certified Developer", na plataforma J2EE JBoss. As principais áreas de actividade são as tecnologias móveis e os sistemas distribuídos, onde tem vindo a desenvolver a maioria dos projectos. Destes, são de destacar as plataformas MWiPEP e DiNO, desenvolvidas no âmbito da PT Inovação.

07

Comunicações Multimédia Personalizadas

Alexandre Miguel Mendonça; Bruno Miguel Duarte; Eduardo Miguel Martins; Luis Barreiro; Neutel Rodrigues; Paulo Alexandre Freire; Paulo Chainho; Toni Barata

palavras-chave:
Personalização, IMS, Presença, Contexto, SIP, XMPP

As emergentes Comunicações Multimédia Personalizadas contribuem para enriquecer a experiência do utilizador, oferecendo a este facilidades que não estavam presentes nos serviços de telefonia tradicionais. Isto deve-se ao facto de permitir a personalização de regras que determinam o comportamento das comunicações em função de diferentes condições. Este artigo apresenta um inovador facilitador Comunicações Multimedia Personalizadas, disponibilizado pela plataforma XMAS, o servidor aplicacional SIP para as redes IMS da PT Inovação.

1. Introdução

Graças às redes *All-IP* IMS, é hoje possível colocar no mercado aplicações capazes de se adaptarem cada vez mais às necessidades específicas de informação e comunicação de cada utilizador ou grupo de utilizadores (comunidade): Comunicações Multimédia Personalizadas. Este tipo de aplicações permitem às empresas do Grupo PT, que têm melhor conhecimento do mercado onde actuam, aumentar a fidelização do cliente, aumentar as suas margens e combater com maior eficácia os novos intervenientes que operam numa escala global como é o caso do Skype, Google ou MSN. Este artigo apresenta um inovador facilitador de Comunicações Multimédia Personalizadas disponibilizado pela plataforma XMAS, o servidor aplicacional SIP para as redes IMS da PT Inovação.

Na secção 2 é feito o enquadramento das Aplicações Personalizadas, identificando os principais aspectos a ter em conta; na secção 3 são descritas as funcionalidades das Aplicações Personalizadas desenvolvidas pela PT Inovação, e a sua solução técnica descrita na secção 4. A secção 5 conclui o artigo.

2. Enquadramento

A personalização das aplicações de informação e comunicação, é um vasto e complexo problema composto por vários aspectos. No limite, cada utilizador quer ter (ou ele próprio criar) aplicações feitas à medida das suas necessidades pessoais ou da comunidade onde está inserido, e.g., a sua família ou a empresa onde trabalha. Distinguem-se os seguintes aspectos da personalização:

> **Comunicações Personalizadas –** o utilizador pretende comunicar da forma mais apropriada, de acordo com o seu perfil e contexto. Por exemplo, adaptar as prioridades das comunicações de acordo com a sua actividade (e.g., ocupado no trabalho, ocupado em casa);

> **Informação Personalizada –** o utilizador pretende ser informado da forma mais apropriada sobre conteúdos, serviços, pessoas, etc. de acordo com o seu perfil e contexto. Exemplo, o utilizador é notificado acerca da existência de uma boa oportunidade para contratar trabalhadores em *outsourcing* para a sua empresa;

> **Calendarização Personalizada –** o utilizador quer gerir a sua agenda da forma mais apropriada de acordo com o seu perfil e contexto. Por exemplo, adaptar as prioridades de calendarização de acordo com o contexto do utilizador, e de acordo com o contexto de outros utilizadores envolvidos;

> **Entretenimento Personalizado –** o utilizador pretende ser entretido de acordo com o seu perfil e contexto, de modo a aproveitar o melhor possível o seu tempo livre. Por exemplo, se o utilizador estiver aborrecido e não estiver bem disposto, o serviço deverá ter a iniciativa de, e.g., tocar conteúdos alegres (música, clips do Gato Fedorento) de acordo com as suas preferências.

Neste artigo apenas são focados os aspectos da Comunicação Personalizada.

3. Aplicações do tipo *Personal Communication Manager* (PCM)

As aplicações do tipo *Personal Communication Manager*, são aplicações de Comunicação Personalizada que permitem ao seu utilizador definir o comportamento mais desejável das suas comunicações, através de regras configuráveis por si próprio. As regras são definidas por condições e acções. A acção duma regra é executada quando as condições dessa mesma regra são satisfeitas. Deste modo, o utilizador pode definir inúmeras regras, de acordo com as suas necessidades particulares e.g., encaminhar chamadas para a secretária quando está ocupado no trabalho, encaminhar chamadas de trabalho para o *voicemail* quando está fora do horário de trabalho, rejeitar chamadas não desejáveis, etc. As regras podem ser aplicadas tanto para chamadas de entrada, como para chamadas de saída. Em particular, é possivel ter funcionalidades de *Call Barring* para, por exemplo, determinados domínios ou gamas de numeração (e.g., serviços de valor acrescentado).

Mas uma das características mais inovadoras destas aplicações, é a possibilidade de usar informação de contexto do utilizador (usando informação de presença do utilizador) para determinar qual a acção mais apropriada a ser executada. Deste modo, é possível ter regras, como por exemplo "se estiver triste encaminhar chamadas para *voicemail*" ou "se estiver muito feliz tocar um *sketch* do Gato Fedorento como *Multimedia Ringing Tone*, ou "quando estou a ver um jogo do Benfica não quero ser incomodado e as chamadas são atendidas com um vídeo do Nuno Gomes a marcar um golo ao Baía."

3.1 Gestão de Regras

Na versão actual do serviço, o utilizador pode configurar as suas regras através de uma página *Web*. O utilizador pode criar vários grupos de Regras, e seleccionar o grupo de regras activo. Deste modo apenas pode estar um grupo de regras activo de cada vez (figura 1).

figura 1 – Regras Activas

figura 2 – Criação de Regra

Para cada grupo de regras o utilizador pode adicionar e remover regras. Para criar uma regra o utilizador dá um nome à regra, selecciona uma ou mais condições e selecciona a acção a executar (figura 2).

3.2 Criação Administrativa de Serviços Pessoais

A aplicação *Personal Communication Manager* não é uma aplicação para o utilizador final, mas antes uma aplicação para os administradores criarem classes de serviços de acordo com as necessidades mais urgentes do mercado. Para o efeito, o administrador tem ferramentas para criar novas classes de serviço definindo para cada uma delas quais as condições e acções disponíveis.

Alguns exemplos de classes de serviço que podem ser criados pelos administradores são:

> **Serviço básico de *voicemail*** onde define condições estáticas para a acção de encaminhar chamada para depósito (i.e., o subscritor não pode configurar as condições e acções);

> **Serviço *Gold* de *voicemail*** onde o subscritor pode configurar algumas condições para depósito;

> **Serviço de Barramento de Chamadas** onde o subscritor pode definir quais as gamas de endereços ou domínios para os quais os seus utilizadores (no caso duma família podem ser os filhos, no caso duma empresa podem ser os colaboradores) não podem efectuar chamadas;

> **Serviço de *Ringback Tone Multimedia*** onde o utilizador pode definir diferentes toques multimédia de acordo com diferentes condições.

Para cada classe de serviço o administrador pode aplicar os modos de tarifação mais apropriados, de modo a gerar novas fontes de receitas através da disponibilização de funcionalidades de personalização e exclusividade a quem esteja disposto a pagar por isso.

Outro aspecto importante, é a distinção entre subscritor (quem paga) e utilizador (quem usufrui). Exemplos: pais e filhos; gestor de empresa e colaboradores. Deste modo o subscritor, e de acordo com os previlégios dados pelo administrador que criou o serviço, pode ele próprio criar serviços para os seus utilizadores. Por exemplo, o subscritor pode definir duma forma estática quais os serviços utilizáveis pelos utilizadores (chamadas para serviços de valor acrescentado), e quais as condições e acções disponíveis para os utilizadores configurarem.

3.3 Condições e Acções Disponíveis

Actualmente, estão disponíveis as seguintes condições:

> **Condição dia/hora –** o utilizador pode definir uma regra que conduza a acções diferentes consoante a hora do dia ou a data. Por exemplo, durante o período de trabalho pode desejar que a chamada seja rejeitada ou reencaminhada para uma caixa de correio;

> **Condição estado da chamada –** consoante o chamado esteja desligado, não atenda, esteja ocupado ou rejeite a chamada pode ser desencadeada uma acção diferente em cada um dos casos;

> **Condição identidade do chamador –** de acordo com a identidade de quem chama (e.g., se fizer parte da minha "lista negra") são executadas diferentes acções, e.g., a chamada é rejeitada com indicação de que não deve voltar a tentar de novo;

> **Condição identidade do chamado –** de acordo com a identidade do chamado, e.g., chamadas para serviços de valor acrescentado, são executadas diferentes acções, por exemplo a chamada é rejeitada tocando o anúncio "acesso ao serviço proibido pelo subscritor";

> **Condição estado de presença –** de acordo com o estado de presença que o chamado assinala (ocupado, ausente, a almoçar, não perturbar, etc.), são desencadeadas diversas acções diferentes, por exemplo depósito de uma mensagem ou reencaminhamento para outro contacto;

> **Incondicional –** nesta situação a regra é sempre executada independentemente das condições satisfeitas.

As acções disponíveis são:

> **Encaminhar para depósito de mensagem** de voz (*voicemail*) ou vídeo (*videomail*) por diferentes razões incluindo por ocupado, não atendimento, desligado, etc.;

> **Encaminhar chamada para outra pessoa** ou para outro terminal alternativo (e.g., o meu telemóvel);

> **Rejeitar chamada;**

> **Tocar mensagem** de áudio e vídeo e desligar;

>**Tocar como *ringing backtone*** um determinado *clip* de áudio ou vídeo.

3.4 Aplicações de Conferência Personalizadas

Outro exemplo de Aplicação IMS Personalizada é o serviço de Conferência Multimodal XMAS (figura 3). Esta aplicação permite ter sessões de conversação de voz, texto e vídeo entre três ou mais pessoas. Estas sessões podem ser controladas através de páginas *Web*, i.e., disponibiliza funcionalidades de "*floor-control*" incluindo

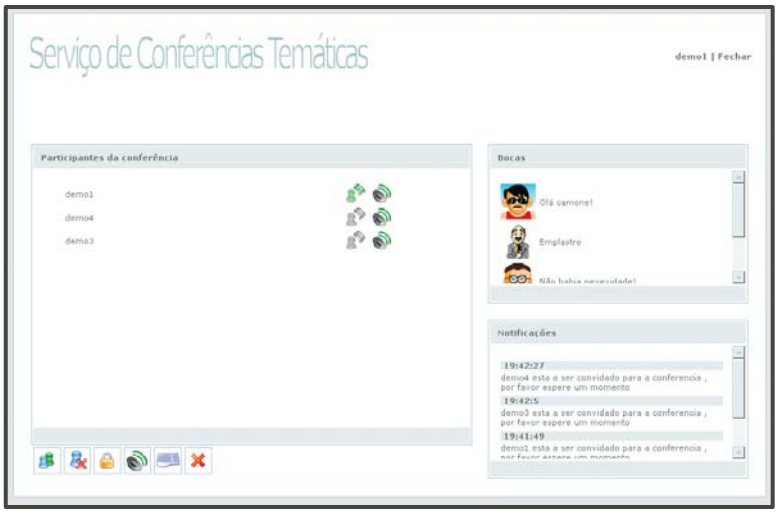

figura 3 – "*Floor Control*" das Sessões de Conferência

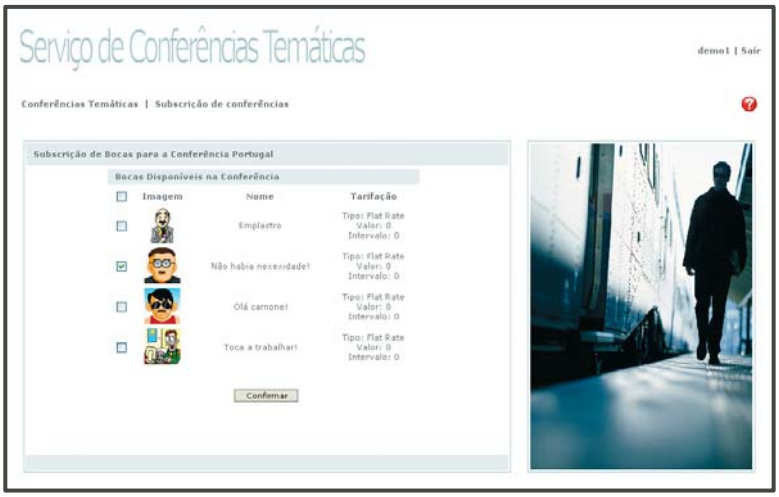

figura 4 – Subscrição de *Emoticons* Áudio para Salas de Conferência

identificação de orador, silenciar participante, dar a palavra ao participante, expulsão de participante, convite a novos participantes.

Mas a característica mais inovadora do serviço é a possibilidade de cada participante tocar para todos os restantes participantes, clips de áudio do tipo *emoticon*, de modo a tornar a sua comunicação mais expressiva tal como é habitual nas comunicações por *Chat* (ver funcionalidade Bocas do serviço SAPO Messenger do Portal SAPO). Cada participante pode ter o seu conjunto distinto de *emoticons*. Para o efeito o utilizador pode subscrever mais ou menos *emoticons* a partir da página de personalização do serviçofigurara 4). Cada *emoticon* pode ter um custo associado de acordo com o grau de exclusividade pretendido. Deste modo, as funcionalidades de personalização também podem ser usadas como geradoras de novas receitas.

4. Solução Tecnológica: XMAS

As aplicações descritas na secção 3 foram desenvolvidas de acordo com as emergentes arquitecturas de referência IMS, e são executadas no Servidor XMAS – *eXtensible Multimedia Aplication Server* – o servidor de Aplicações SIP AS da PT Inovaçãofigurara 5).

O Servidor XMAS usa as tecnologias mais apropriadas para as emergentes plataformas de execução de serviços (SDP – *Service Delivery Plataform*) e é constituido pelas seguintes camadas funcionais horizontais:

> **Ambiente de Execução de Serviço (*Container*)** – disponibiliza todas as funcionalidades necessárias para gerir o ciclo de vida dos diferentes componentes constituintes das aplicações. Existem três tipos de ambientes de execução, cada um especializado em três tipos de componentes:

1. SLEE: Ambiente de Execução de Lógica Serviço Assíncrona (orientada para eventos), especializado para componentes de controlo de serviços de telecomunicações e.g., controlo de chamada;
2. EJB: Ambiente de Execução de Lógica Serviço Síncrona. Especializado em componentes tradicionais dos sistemas de informação e.g., aprovisionamento;
3. *Servlet*: Ambiente de Execução de Lógica de Apresentação. Especializado em componentes para interface do utilizador *Web*.

> ***Framework* –** Disponibiliza funcionalidades transversais, incluindo do aprovisionamento (de serviços, aplicações, utilizadores), descoberta e orquestração de serviços, monitorização, segurança, etc.;

> **Conectores ou *Resource Adaptors* –** Componentes para suporte às interfaces para elementos de rede. A versão actual suporta as seguintes interfaces de rede: ISC (SIP), Sh (*Diameter* HSS), Rf (*Diameter offline charging*), Ro (*Diameter online charging*), *Media Server* (INOVOX-IP), Bases de Dados relacionais, XMPP (Mensagens Instantâneas e Presença);

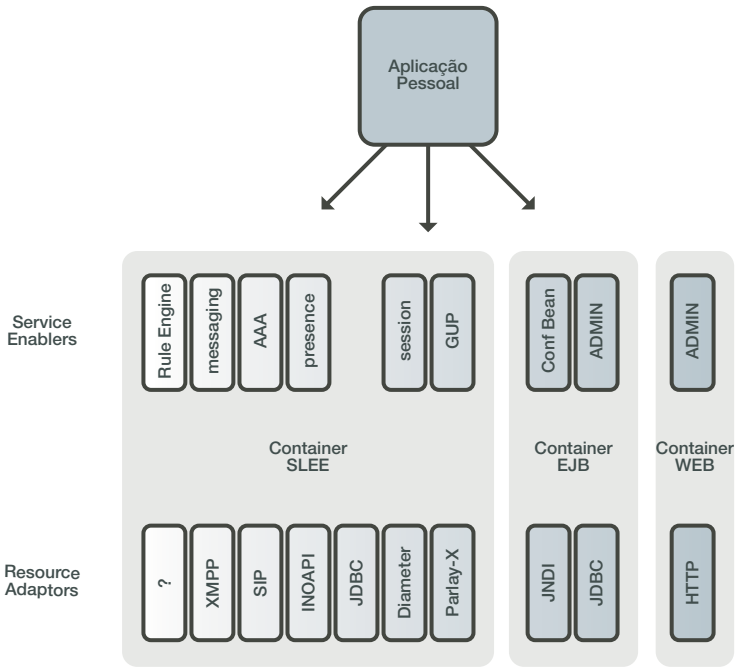

figura 5 – Arquitectura XMAS

> **Facilitadores ou *Service Enablers* –** Componentes para disponibilização de funcionalidades de mais alto nível, com abstração sobre as componentes Conectores. A versão actual suporta os seguintes facilitadores: *Proxy*, B2BUA (sessão), *Redirect*, AAA, Presença, *Generic User Profile*, *Charging*, *Messaging*. Do ponto de vista da personalização destacam-se os facilitadores GUP e Presença;

> **Aplicações –** As aplicações são compostas por componentes específicas da aplicação e por *Service Enablers*. Actualmente são disponibilizadas as seguintes aplicações: *Personal Communication Manager*, Conferência Multimodal, PC2Phone com diferentes modos de tarifação (incluindo pré-pagos, *flat-rate* e chamadas patrocinadas), *Unified Messaging*, Video Portal, *Multimedia Ringback tone*.

5. Conclusões

Este artigo descreveu muito brevemente o trabalho desenvolvido pela PT Inovação na área das Aplicações Personalizadas IMS. Foram descritas as principais características e funcionalidades destas aplicações desenvolvidas sobre o servidor Aplicacional XMAS, o servidor SIP AS da PT Inovação para as redes IMS.

Actualmente, está em curso o desenvolvimento da 2.ª versão destas aplicações através da experimentação e integração de potentes motores de regras como é o caso do JBOSS *Rules*.

Com este tipo de tecnologia pretende-se atingir um novo patamar na personalização das aplicações através da inferência do perfil de utilizador, por observação do comportamento desse mesmo utilizador e da combinação com outros perfis, incluindo perfis de serviço e conteúdo. Esta informação mais estática será cruzada com informação mais dinâmica relacionada com o contexto do utilizador ou conteúdo/ serviço, de modo a adaptar duma forma dinâmica a execução do serviço às necessidades do utilizador.

Finalmente, o trabalho em curso pretende cobrir os restantes aspectos da personalização identificados na secção 2: Informação, Calendarização e Entretenimento.

Referências

[1] XMAS_co_180105_v01 - Definição de Requisitos e Arquitectura
[2] OMA-AD-Service-Environment - OMA Service Environment, Version 1.0.3
[3] 3GPP TR 23.941, 3GPP Generic User Profile (GUP); Data Description Method (DDM)

Alexandre Miguel Mendonça, licenciado em Engenharia Informática pela Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa. Especialista em Bases de Dados e mecanismos de persistência em JAIN SLEE. Responsável pelo desenho e gestão do modelo de informação do XMAS, pelo *Resource Adaptor* de Bases de Dados Relacionais, do *Enabler* de GUP (*Generic User Profile*) e pelas Interfaces de Utilizador das Aplicações XMAS.

Bruno Duarte, finalista de Mestrado em Engenharia Informática pela Faculdade de Ciências da Universidade de Lisboa. Responsável pelas aplicações *Personal Communication Manager* e pelo *Enabler* de motor de regras do XMAS.

Eduardo Martins, licenciado em Engenharia Informática pela Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa. Especialista em JAIN SLEE, Presença e Contexto. Responsável pela especificação e desenvolvimento de componentes do XMAS, nomeadamente o *Resource Adaptor* de XMPP e os *Service Enablers* de Presença e Mensagens. Participante activo no desenvolvimento dos projectos *Open Source* Mobicents e OPUCE.

Luís Barreiro, licenciado em Engenharia Electrotécnica e de Computadores pelo Instituto Superior Técnico. Especialista no Protocolo SIP e na Arquitectura 3GPP IMS. Experiência com diversas linguagens de programação (C, JAVA, etc.) e em JAIN SLEE. Responsável pelo *Service Session Enabler* (incluindo *Call Control*) do XMAS.

Neutel Rodrigues, técnico de Electrónica e Telecomunicações. Especialista no desenvolvimento de aplicações Multimédia IMS em ambiente JAIN SLEE e Voz sobre IP. Responsável por Aplicações de Conferência Multimédia do XMAS.

Paulo Alexandre Freire, técnico especializado em telecomunicações e redes, electrotecnia e electrónica. CCNA. Frequência universitária no IST no curso de LEEC. Especialista em terminais VoIP e IMS.

Paulo Chainho, mestre em Telecomunicações pelo IST da Universidade Técnica de Lisboa. Gestor de Projectos, Responsável pelo departamento de Protocolos e Platformas de Aplicação no Pólo de Lisboa da PT Inovação e pelo produto XMAS (*eXtensible Multimedia Application Server*), o SIP *Application Server* da PT Inovação. Entusiasta do *Open Source*. Grande experiência em projectos Internacionais de Investigação e Desenvolvimento incluíndo Eurescom, ETSIN SPAN e EU IST. Consultoria na introdução de plataformas de serviços RPG e 3G no Grupo PT.

Toni Barata, bacharel em Engenharia de Sistemas de Telecomunicações e Electrónica pelo ISEL/IPL. Especialista em VoIP e protocolo SIP (Cinco anos de experiência), na configuração e administração de sistemas Linux, centrada na distibuição da Red Hat. Responsável pela Qualidade (*Quality Assurance*) do XMAS.

01

Optimização Automática de Redes Rádio UMTS


Álvaro Gomes, Pedro D'Orey (IT)

palavras-chave:
UMTS, UTRAN, WCDMA, RRM, Planeamento Rádio, optimização, KPI

Devido às características intrínsecas do WCDMA e à grande variedade de serviços que a rede UMTS transporta, o ambiente rádio é muito mais dinâmico que o existente no sistema GSM, tornando a gestão de recursos rádio (RRM) mais complexa. O *automated tuning* de parâmetros de RRM utilizando técnicas de controlo inteligente foi introduzido recentemente. Diferentes estudos demonstram o elevado impacto que esta técnica tem no desempenho da rede rádio UMTS devido à sua adaptabilidade às variações de tráfego. Neste artigo são descritos os conceitos principais relacionados com o planeamento e optimização de redes, ajuste de parâmetros de RRM e funcionamento de um sistema automático de *tuning* (ATS).

1. Introdução

O projecto de uma rede de acesso rádio é um processo complexo que pode ser dividido em duas fases principais. Na fase inicial os objectivos (capacidade, cobertura e QoS) são definidos, a rede é dimensionada e o planeamento rádio é iniciado. Neste processo são utilizadas sofisticadas ferramentas de planeamento e optimização, que recorrem a complexas funções de custo, dado que existem diversos *trade-offs*. Finalmente, os parâmetros da estação base, nomeadamente os correspondentes aos algoritmos de gestão dos recursos rádio (RRM), são fixados.

Numa segunda fase, e depois da colocação ao serviço da rede, a performance e as características de qualidade desta são regularmente monitorizadas e poderá haver necessidade de proceder a optimizações, devido às flutuações de tráfego, tipo de serviço, mobilidade dos utilizadores, etc. Assim podem ser ajustados parâmetros *hard* (p.ex., *tilt* da antena) e/ou *soft* (mecanismos RRM) de modo a melhorar o desempenho. Este é um processo manual que deve ser realizado periodicamente. O processo de *auto-tuning* (*automated tuning*) tem como objectivo ajustar dinamicamente estes parâmetros (*soft* apenas) de um modo contínuo sem intervenção humana, sendo apenas necessário definir no início da optimização os critérios de qualidade.

Uma representação esquemática do conceito de *auto-tuning* é apresentada na figura 1.

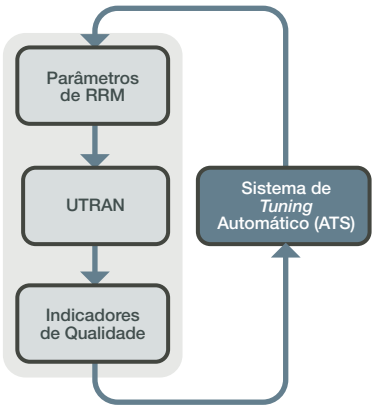

figura 1 - Ciclo de optimização automática dos parâmetros de RRM.

2. Processo de Optimização

A optimização de uma rede UMTS tem como objectivo melhorar a qualidade tal como ela é vista pelos utilizadores, e por outro lado, garantir que os recursos rádio são utilizados eficientemente. A optimização é um ciclo realimentado, em que se procede ao ajuste manual dos parâmetros, e que termina quando os critérios de qualidade propostos pelo operador são atingidos. Este ciclo está representado esquematicamente na figura 2.

O ciclo inicia-se com a definição de qualidade (global extremo-a-

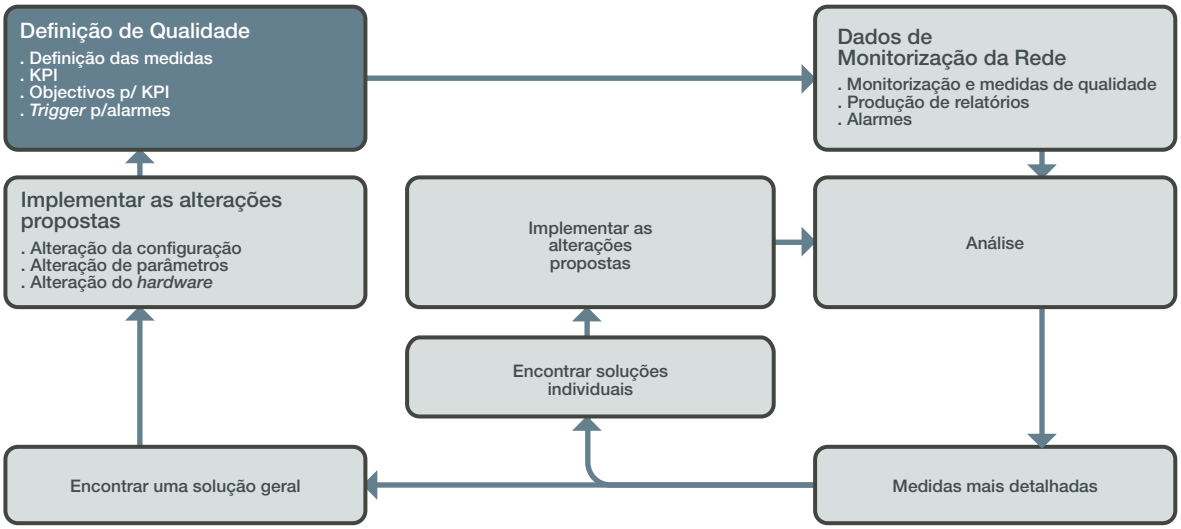

figura 2 - Ciclo de Optimização.

extremo, critérios para cada tipo de serviço e valor limite para os parâmetros chave (KPI). Na fase seguinte, os dados de performance da rede são recolhidos (através de NMS, *drive tests*, etc.) e relatórios com informação estatística dos KPI são produzidos. Com base nestes dados, e conhecendo a configuração actual da rede, é iniciado o processo de análise. A comparação entre os KPI e os valores alvo indicam as áreas problemáticas da rede e onde esta deverá ser optimizada.

O *tuning* da rede pode incluir a alteração de parâmetros *soft* (parâmetros RRM) ou *hard (hardware*). Alterações típicas dos parâmetros incluem, por exemplo, janela de *handover*, potência do piloto ou nível máximo do factor de carga. Por vezes são necessárias alterações a nível do *hardware* instalado: *tilt* da antena, número de sectores, introdução de nova portadora, ou de novos sites, etc. A introdução de novo *hardware* só deve ser ponderada quando não é possível atingir os objectivos através da optimização de parâmetros. Esta solução é mais dispendiosa (aquisição/instalação HW e manutenção), e não é linear, devido às características da tecnologia CDMA, que a capacidade ou cobertura sejam significativamente melhoradas com este tipo de alteração.

São então realizadas iterativamente correcções, através do *tuning* de parâmetros individuais que afectam a qualidade reportada, até que a qualidade alvo seja atingida. Por fim, além do ajuste de parâmetros individuais, uma solução geral deve ser encontrada. O ciclo volta ao início e verifica-se se as alterações introduzidas permitiram atingir os objectivos de qualidade definidos pelo operador.

Após a optimização ter sido realizada e as alterações terem sido introduzidas na rede, é necessário comparar a performance global com o caso pre-optimização. Tem que se avaliar o impacto noutros subsistemas da célula e nas células adjacentes pois a optimização pode melhorar um subconjunto da área funcional mas pode degradar a qualidade noutros subconjuntos.

A optimização é um ciclo contínuo de monitorização e correcções que pode ser feito célula a célula. Exige um conhecimento técnico profundo do sistema, da sua parametrização, interacção entre subsistemas e das características específicas da tecnologia CDMA.

3. Ajuste automático dos parâmetros de RRM

3.1 Conceitos

Os algoritmos de RRM são responsáveis por garantir a qualidade de serviço (QoS), manter a cobertura planeada e oferecer elevada capacidade. A família dos algoritmos de RRM pode ser dividida em controlo de *handover*, controlo de potência, controlo de admissão, controlo de carga e funcionalidades de escalonamento de pacotes [2]. Estas funções têm diversos parâmetros associados que permitem o controlo e eficiente utilização dos recursos rádio. Os mecanismos de ajuste automático têm como objectivo ajustar dinamicamente certos parâmetros de RRM como, por exemplo, o nível máximo do factor de carga, que controla a admissão de novos utilizadores e a QoS na rede.

A temática da automatização tem tido interesse crescente por parte de fabricantes, operadores e comunidade académica e tem levado a intensa investigação. Alguns algoritmos foram desenvolvidos baseados basicamente no ajuste de um único parâmetro ou subconjuntos de valores de um parâmetro [1]. O *tuning* de parâmetros tais como: janelas do algoritmo de *handover* [3], nível máximo da interferência recebida

(factor de carga UL) [5], nível máximo do factor de carga do DL e potência máxima de transmissão por ligação [6], Eb/No no UL/DL para tráfego de dados [7] e potência de transmissão do piloto [8][9], foram propostos em diversos artigos e livros [1][4]. Estes estudos demonstraram a aplicabilidade do conceito e concluíram que é obtido um significativo incremento da capacidade da rede quando comparada com a configuração estática. Por outro lado, o estudo [13] obteve resultados interessantes mas focou a optimização na QoS em vez da performance.

Os parâmetros podem ser optimizados célula por célula ou para um grupo de células com características semelhantes (*cluster*). Em diversos estudos, nomeadamente em [4], ficou provado que a optimização célula a célula pode conduzir a ganhos adicionais de performance quando comparados com a optimização do *cluster*, dado que valores óptimos são obtidos para cada uma das células. Por outro lado, o estudo [10] afirma que será atingida maior estabilidade com a solução baseada no *cluster*.

Outro tópico interessante é a optimização multiparâmetro, na qual diversos parâmetros chave da rede rádio UMTS são optimizados simultaneamente com um método de controlo automático. Diversos estudos propuseram e demonstraram a aplicabilidade deste tipo de optimização [4][10].

A concepção de controladores para proceder ao ajuste de parâmetros RRM é uma das principais tarefas adjacentes ao processo de auto-*tuning*. Em trabalhos iniciais, nomeadamente [5][6], procedeu-se à implementação de controladores baseados num conjunto simples de regras principalmente devido à simplicidade da implementação. Um sistema *rule based* representa o conhecimento num conjunto de regras que indicam a acção a tomar em diferentes situações. Este é constituído por um conjunto de condições IF-THEN, um conjunto de factos, e um interpretador que controla a aplicação das regras. Para determinados factos pode não estar associada inequivocamente uma acção, sendo necessário recorrer a uma função de custo.

A concepção de controladores é uma tarefa complexa e demorada, dado que a performance destes varia de acordo com as características de tráfego da rede. Assim, de modo a obter elevada performance da rede, é necessário optimizar o sistema de controlo de acordo com as condições de utilização. São utilizadas duas técnicas na optimização de controladores: *offline* e *online*.

A primeira é executada num computador antes da introdução na rede em funcionamento. Um dos métodos utilizado é denominado *Particle Swarm* (PS), ver referência [11]. Por outro lado, a optimização *online* tem como objectivo optimizar um controlador numa rede em funcionamento [3], sendo o *Reinforcement Learning* (RL) um destes métodos [14][15].

Para ambientes em que coexistam diversas tecnologias de acesso (GSM, WCDMA, WLAN, etc.) a utilização de uma gestão comum permite distribuir eficientemente os recursos das redes de modo a optimizar capacidade e qualidade. Assim o conceito de *auto-tuning* pode ser estendido a *Common-RRM* para redes com diversas tecnologias de acesso (RAT), de modo a melhorar a cooperação entre sistemas.

3.2 O Sistema

No âmbito do projecto europeu IST – AROMA[1] [12], está a ser projectado um sistema para fazer o *tuning* de parâmetros RRM de redes rádio

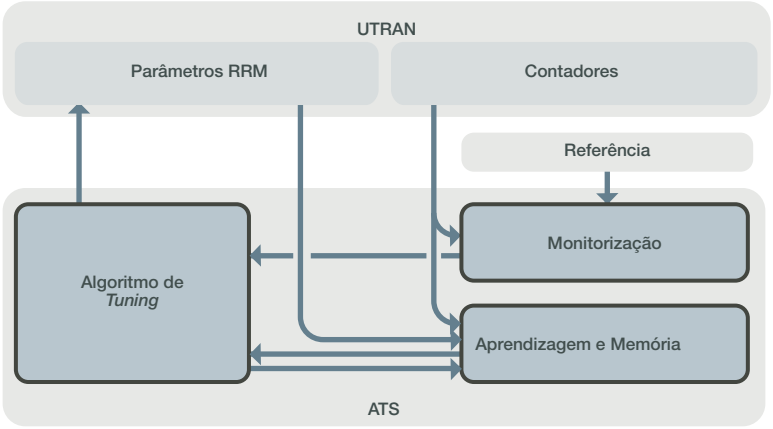

figura 3 – Diagrama funcional do ATS

[1] *Advanced Resource Management Solutions for Future All IP Heterogeneous Mobile Radio Environments*

UMTS. Uma representação conceptual do Sistema de Ajuste Automático (ATS) proposto é apresentado na figura 3. Este é constituído por três blocos principais (Algoritmo de *Tuning*, Aprendizagem e Memória, e Monitorização) e por duas interfaces (UTRAN e Referência). A UTRAN fornece os parâmetros de RRM e contadores, que podem ser agrupados num só parâmetro (KPI) para fornecer uma melhor compreensão do real estado da rede. O bloco Referência fornece a visão do operador da rede no que diz respeito à qualidade e performance.

O ATS cria um circuito de realimentação entre as medidas da rede e os parâmetros de RRM. A rede é constantemente monitorizada, os parâmetros seleccionados são colocados em memória para análise estatística e são comparados com o valor de referência. Quando uma determinada célula não cumpre os critérios, o algoritmo de *tuning* é iniciado e, possivelmente, os parâmetros de RRM são actualizados. Deste modo, o processo de ajuste da rede torna-se um processo automático.

O ATS é constituído por três blocos, sendo a função de cada um deles descrita de seguida:

a) Monitorização – Monitoriza os contadores ou KPI. Quando uma destas medidas tiver um valor superior ao de referência (por exemplo, KPIref) este bloco despoleta o processo de auto-*tuning* através da passagem de um alarme/evento para o bloco "Algoritmo de *Tuning*";

b) Aprendizagem e Memória – Este bloco pode ser visto como uma base de dados que acumula conhecimento da rede (parâmetros RRM e contadores), que serão utilizados quando se fizer a previsão da acção a tomar. Ao monitorizar constantemente os parâmetros de RRM, contadores e/ou KPI, uma base de dados com informação estatística é criada (memória). O bloco deve reter informação da relação entre variáveis, conhecimento que vai adquirindo ao longo da vida da célula e que é diferente de célula para célula, conforme as características específicas da célula, mobilidade dos utilizadores, serviços utilizados, etc.

Esta entidade é também responsável por determinar tendências e comportamentos típicos relacionados com aspectos de tráfego e cobertura rádio através de uma análise contínua no tempo (aprendizagem). Deste modo, extrapola a caracterização de cada célula, o que permite a simulação (previsão de resultados) de uma nova configuração, mesmo antes da sua implementação na rede.

Por fim, este bloco poderá identificar células com características similares e criar *clusters*. Esta funcionalidade é essencial quando se pretende optimizar uma rede com um elevado número de células;

c) Algoritmo de Tuning – O terceiro bloco também poderia ter sido denominado como algoritmo de controlo inteligente. Ele recebe um alarme do bloco de monitorização, e com a informação fornecida pelo bloco "Aprendizagem e Memória", decide as acções a tomar (sub-sistema, parâmetro e valor). A acção implica usualmente a mudança de um ou mais parâmetros de RRM. Caso se verifique que o objectivo da optimização não pode ser atingido, ou que a relação custo/ benefício não é vantajosa para o operador, a acção deverá ser terminada, e, por exemplo, outro sub-sistema deverá ser escolhido.

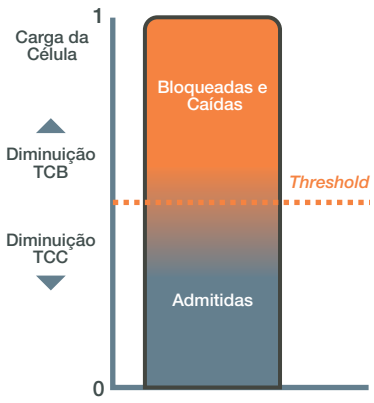

figura 4 – Controlo de admissão/carga

O algoritmo executa iterativamente as seguintes tarefas:

1. Identifica o(s) sub-sistema(s) implicados no processo de *tuning*: controlo de potência, controlo de carga, controlo de admissão, controlo de *handover*, etc.;

2. Identifica o(s) parâmetro(s) mais favoráveis a corrigir;

3. Calcula e atribui valores aos parâmetros;

4. Calcula o impacto nos restantes sub-sistemas e nas células vizinhas através da interacção com o bloco de aprendizagem e o valor final de cada parâmetro;

5. Actualiza a parametrização.

4. *Tuning* do Factor de Carga

Numa primeira fase o sistema ATS proposto no âmbito do projecto AROMA será dedicado ao *tuning* de um parâmetro do algoritmo de controlo de admissão: o nível máximo do factor de carga no *Uplink*.

4.1 Algoritmo de Controlo de Admissão

O algoritmo de RRM "Controlo de Admissão" assegura que a área de cobertura planeada e que a qualidade das ligações existentes é mantida após a aceitação de novas ligações. O procedimento controla os pedidos de novas ligações, isto é, se um novo *Radio Access Bearer* (RAB) pode ser estabelecido ou não. O algoritmo estima o incremento de carga que

o estabelecimento de um novo RAB irá causar. São utilizadas medidas da interferência recebida pela estação base no UL e potência total de transmissão da estação base no caso do DL. Para cada uma das direcções (directa e inversa), o valor da nova carga estimada será comparada com um limite definido na fase de planeamento (*threshold* do factor de carga). A ligação poderá ser aceite nessa direcção caso o valor estimado seja inferior ao valor do nível máximo, de outro modo será rejeitada. Esta tarefa tem que ser executada tanto no UL bem como no DL, dado que o RAB apenas será aceite se o algoritmo de controlo de admissão do UL e do DL o admitirem. Caso contrário o pedido de estabelecimento de um nova chamada será negado.

4.2 Parâmetros e QoS Alvo

Dois parâmetros importantes para o *auto-tuning* do algoritmo de controlo de admissão são o *threshold* do factor de carga para o UL e da potência de transmissão da estação base para o DL. O nível de carga máximo é dedicado ao controlo da congestão (interferência) da rede. Este foi introduzido para evitar efeitos de saturação da rede devido a ligações com ineficaz controlo de potência ou que deterioram a QoS.

Diversos indicadores de qualidade podem ser obtidos da rede para realizar o *auto-tuning* do algoritmo de controlo de admissão. Podemos referir por exemplo:

> Taxa de Chamadas Caídas (TCC);

> Taxa de Chamadas Bloqueadas (TCB).

4.3 Mecanismo de *tuning* do CA no UL

A escolha do valor para o factor de carga tem em conta indicadores de qualidade de serviço (chamadas caídas e bloqueadas) e cobertura/capacidade. O aumento do nível máximo do factor de carga provoca interferência adicional na rede, que tem como consequência um aumento da taxa de chamadas caídas/má qualidade, perda de cobertura e diminuição do número de chamadas bloqueadas.

Por outro lado, diminuir o *threshold* provoca uma maior probabilidade de bloqueio para chamadas de voz e maior probabilidade de espera (*queuing*) para dados. Deste modo pode-se afirmar que existe um compromisso entre os indicadores de qualidade e o nível a que é colocado o factor de carga da célula.

4.4 Impacto nas células vizinhas

A optimização da rede será realizada célula a célula. A alteração de um parâmetro numa determinada célula (teste) poderá ter impacto nas células vizinhas. Devido aos efeitos de acoplamento típicos de CDMA, um problema numa determinada célula de teste pode propagar-se para as células vizinhas e afectar a sua performance.

Na sua configuração básica todas as células de uma rede UMTS partilham a mesma frequência. O aumento de tráfego na célula de teste causa interferência adicional e variações de carga nas células da rede. O aumento do nível de interferência obriga a que os terminais aumentem também a sua potência de transmissão, criando assim um ciclo de *feedback*. O modo como a interferência se propaga entre as células depende de diversos factores nomeadamente: diagrama de radiação da antena, topologia do terreno, distribuição dos utilizadores e distância entre células.

Assim, antes de alterar os parâmetros de RRM de uma célula, deve ser analisado o impacto nas células vizinhas. A quantificação e caracterização dos fluxos de interferência entre células é uma das funções do bloco de Aprendizagem e Memória do ATS.

5. Vantagens

A implementação de um ATS tem para o operador de uma rede de comunicações móveis as seguintes vantagens:

> **Redução dos custos operacionais –** A optimização da rede é executada automatica e dinamicamente, reduzindo as despesas com a optimização manual e, consequentemente, o número de trabalhadores especializados necessários;

> **Redução do número de equipamentos instalado –** Com o *tuning* automático podem-se atingir níveis de optimização superiores aos obtidos com optimização manual, o que melhora a performance da rede e, consequentemente, o número de equipamentos necessário (ex.: Nós-B). Reduzem-se, assim, os investimentos em equipamento e os gastos em O&M (Operação e Manutenção);

> **Melhoria da capacidade e qualidade da rede (aumento das receitas) –** Os parâmetros da rede são dinamicamente optimizados de acordo com as condições reais da rede (tráfego, serviços, mobilidade, etc.), maximizando em tempo real a capacidade e qualidade da rede. Portanto, as receitas do operador podem aumentar substancialmente;

> **Aumento do nível de abstracção na gestão da rede –** Dado que com o *auto-tuning* o operador apenas precisa de escolher certos parâmetros de qualidade, em vez de parâmetros individuais do sistema, deixa de ser necessário um conhecimento tão profundo deste e um nível superior de abstracção da rede de gestão pode ser atingido. Assim, o operador pode focar-se mais no

negócio, qualidade e serviços para os seus utilizadores.

6. Conclusões

Este artigo apresentou de forma sucinta os conceitos principais relacionados com mecanismos de ajuste automático de parâmetros de RRM de uma rede UMTS. Apresentou o sistema ATS proposto no âmbito do projecto europeu IST-AROMA, bem como um exemplo ilustrativo do *tuning* de um parâmetro do algoritmo de Controlo de Admissão.

Esta área de investigação é uma das mais promissoras e com excelentes resultados comprovados em diversos estudos. Estes mecanismos conduzem a um aumento na eficiência operacional devido ao aumento do nível de gestão da rede. Por outro lado com a introdução da automação, a performance da rede é melhorada e os recursos rádio são utilizados mais eficientemente.

Referências
[1] Laiho, J,, Wacker, A, and Novosad, T., Radio Network Planning and Optimisation for UMTS, Second Edition, John Wiley & Sons, 2006
[2] Holma, H. and Toskala,A., WCDMA for UMTS - Radio Access for Third Generation Mobile Communications, Third Edition, John Wiley & Sons, 2004
[3] Nawrocki, M., Dohler, Mischa and Aghvami, A., Understanding UMTS Radio Network – Modelling, Planning and Automated Optimisation, John Wiley & Sons, 2006
[4] Höglund, A. and Valkealahti,K., "Automated Optimization of Key WCDMA Parameters", Wireless Communications and Mobile Computing, vol. 5, issue 3, pp. 257-271, 2005.
[5] Höglund, A., Pöllönen, J., Valkealahti, K. and Laiho, J. "Quality-based Auto-tuning of Cell Uplink Load Level Targets in WCDMA", Proc. 57th IEEE Vehicular Technology Conference Spring, vol. 4, pp. 2847 -2851, 2003.
[6] Höglund, A. and Valkealaht, K. "Quality-based Tuning of Cell Downlink Load Target and Link Power Maxima in WCDMA", Proc. 56th IEEE Vehicular Technology Conference Fall, vol. 4, pp. 2248-2252, 2002.
[7] Hämäläinen, A., Valkealahti, K., Höglund, A. and Laakso, J. "Auto-tuning of Service-specific Requirement of Received EbNo in WCDMA", Proc. 56th IEEE Vehicular Technology Conference Fall, vol. 4, pp. 2253-2257, 2002.
[8] Valkealahti, K. Höglund, A. Parkkinen,J. and Flanagan, A. "WCDMA Common Pilot Power Control with Cost Function Minimization", Proc. 56th IEEE Vehicular Technology Conference Fall, vol. 4, pp. 2244-2247, 2002.
[9] Valkealahti, K. Höglund, A., Parkkinen,J. Hämäläinen, A. "WCDMA Common Pilot Power Control for Load and Coverage Balancing", Proc.13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), vol. 3, pp 1412-1416, 2002.
[10] Valkealahti, K., Höglund, A and Novosad, T., "UMTS Radio Network Multi-parameter Control", Proc.14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, vol. 1, pp. 616-621, 2003.
[11] Dubreil, H., Altman, Z., Diascorn, V., Picard, J. and Clerc, M., "Particle Swarm Optimization of Fuzzy logic Controller for high quality RRM Auto-tuning of UMTS Networks", Vehicular Technology Conference Spring, vol. 3, pp.1865 -1869,2005
[12] IST-AROMA website, http://www.aroma-ist.upc.edu
[13] Soldani, D. and Valkealahti, K., "Genetic approach to QoS optimization for WCDMA mobile networks", vol. 4, pp. 2269 - 2273 2005
[14] Glorennec, P.Y., "Reinforcement Learning: an overview", Proceeding of the ESIT 2000, Aachen, Germany, 2000.
[15] Nasri, R., Altman, Z., Dubreil, H., Nouir, Z, "WCDMA Downlink Load Sharing with Dynamic Control of Soft Handover Parameters", Vehicular Technology Conference Spring 2006, vol. 2, pp. 942 – 946, 2006.

Curriculum Vitae

Álvaro Gomes, é licenciado em Electrónica e Telecomunicações pela Universidade de Aveiro pertence ao quadros da Portugal Telecom desde 1981, trabalha actualmente no departamento ETC (Experimentação Tecnológica e Difusão do Conhecimento) da PTIN.

Desde 2000 que o seu trabalho de investigação está focado nas redes móveis 3G e B3G (Beyond 3G) nomeadamente interface rádio e gestão de recursos rádio (RRM), tendo participado nos seguintes projectos IST: SHUFFLE(2000), SEACORN(2002), EVEREST(2004), WINNER(2005) e AROMA(2006).

Na área da consultoria desenvolveu as seguintes actividades: expansão e reengenharia da rede GSM da Cabo Verde Telecom (2003), avaliação da qualidade rede UMTS na área de Aveiro para a TMN (2004) e avaliação do estado da arte das tecnologias de acesso sem fios (BWA) para a PTC (2005). Menção honrosa do concurso "Fórum de Ideias" com o projecto "FootPrint" (2006). È formador nos cursos da PTIN para o UMTS.

Pedro M. d'Orey, licenciado em Engenharia Electrotécnica e de Computadores pela Universidade do Porto. Iniciou a sua actividade profissional como Engenheiro de Telecomunicações na Metro do Porto na área de redes fixas e rádio. Actualmente encontra-se a trabalhar como bolseiro de investigação no Instituto de Telecomunicações (Aveiro), estando integrado no projecto europeu IST-AROMA. No âmbito deste projecto tem como principal função o desenvolvimento e avaliação de performance de mecanismos de optimização automática de redes UMTS. Os seus interesses de investigação prendem-se com a gestão de recursos rádio e técnicas de optimização inteligente.

Agradecimentos
O trabalho aqui apresentado foi co-financiado pelo projecto europeu IST-AROMA [12].

02

Daidalos, Implementação de um Sistema B3G


Telma Mota, Francisco Fontes

palavras-chave:
QoS, SIP, Multimédia, Mobilidade, IMS, AAA, Daidalos

O projecto IST Daidalos tem como objectivo principal a definição, desenvolvimento e validação de um sistema B3G ("*beyond* 3G"). Encontra-se, actualmente, em período de transição (entre a primeira e segunda fases) em que os resultados obtidos permitirão a especificação da nova arquitectura e identificação de novos requisitos. A PT Inovação, conjuntamente com o pólo de Aveiro do Instituto de Telecomunicações, foi e continuará a ser por excelência o *site* de integração do projecto e, como tal, existe já uma infra-estrutura de suporte a serviços multimédia (VoIP, Videotelefonia e *Streaming*), a funcionar sobre uma rede de núcleo IPv6 e acesso heterogéneo (WLAN, TD-CDMA, …). Inclui funcionalidades diversas, das quais se podem destacar a autenticação e autorização de utilizadores, controlo de serviços, reserva de QoS e mobilidade de sessão e de terminal. Neste artigo, descreve-se essencialmente a funcionalidade de que dispomos actualmente neste protótipo, mas aproveitamos também para introduzir alguns dos novos desafios, já identificados pelo projecto no início da segunda fase.

1. Introdução

A definição do IMS (IP *Multimedia Subsystem*) pelo 3GPP [1] veio permitir a aproximação entre as redes móveis e a Internet. Os actuais sistemas 3G prevêem a sua integração, permitindo a disponibilização de diversos serviços, sobre tecnologia IP, incluindo voz e vídeo.

Numa fase em que os sistemas 3G ainda se encontram em processo de implantação no mercado mundial, projectos como o Daidalos, procuram demonstrar o que poderá ser o futuro pós 3G, encontrando soluções técnicas e propondo modelos de negócio que poderão tornar-se realidade num espaço de tempo até 10 anos. Nesse cenário, o operador deverá oferecer aos seus clientes serviços essenciais, como a mobilidade, segurança e qualidade nas comunicações, suportados sobre uma tecnologia IP que teve na sua génese a necessidade de comunicação robusta entre máquinas. As actuais soluções técnicas, incluindo protocolos e arquitecturas, são limitadas, sendo portanto necessário apontar linhas de evolução.

Deste modo, o trabalho neste projecto permite à PT Inovação acompanhar os recentes desenvolvimentos e contribuir na especificação, desenvolvimento, experimentação e avaliação de um sistema B3G.

2. Enquadramento

O projecto IST Daidalos – projecto co-financiado pela Comissão europeia – propõe-se desenhar e desenvolver um sistema "pós 3G" (B3G – *Beyond* 3G), para ambientes móveis e sem fios. A arquitectura Daidalos deve suportar o acesso seguro a serviços pervasivos e personalizados, sobre infra-estruturas de rede heterogéneas. O objectivo é poder garantir aos utilizadores o acesso e a continuidade dos seus serviços, a partir de qualquer rede de acesso, tendo em conta o contexto e as suas preferências.

O projecto entende a mobilidade como sendo crucial para os clientes; considera a mobilidade a alavanca principal das telecomunicações nas mais diversas áreas, como o negócio, a educação e o lazer. No entanto, as mudanças rápidas de tecnologia e a diversidade nas redes de acesso e nos próprios terminais, tornam complexa a tarefa de garantir aos utilizadores continuidade e qualidade na utilização dos serviços.

A actividade do projecto é orientada por cinco conceitos chave que estão na base da definição e do desenvolvimento da arquitectura Daidalos:

> **MARQS** (Mobilidade, AAA – Autenticação, Autorização e *Accounting*, Gestão de Recursos de Rede, QoS e Segurança);

> **VID** (*Virtual Identity*) que garante independência dos utilizadores em relação aos terminais, segurança e privacidade entre domínios;

> USP (*Ubiquitous and Seamless Pervasiveness*) que permite a personalização e a adaptação dos serviços a diferentes domínios e tecnologias;

> **SIB** (*Seamless Integration of Broadcast*), que visa a integração de componentes *broadcast*, ao

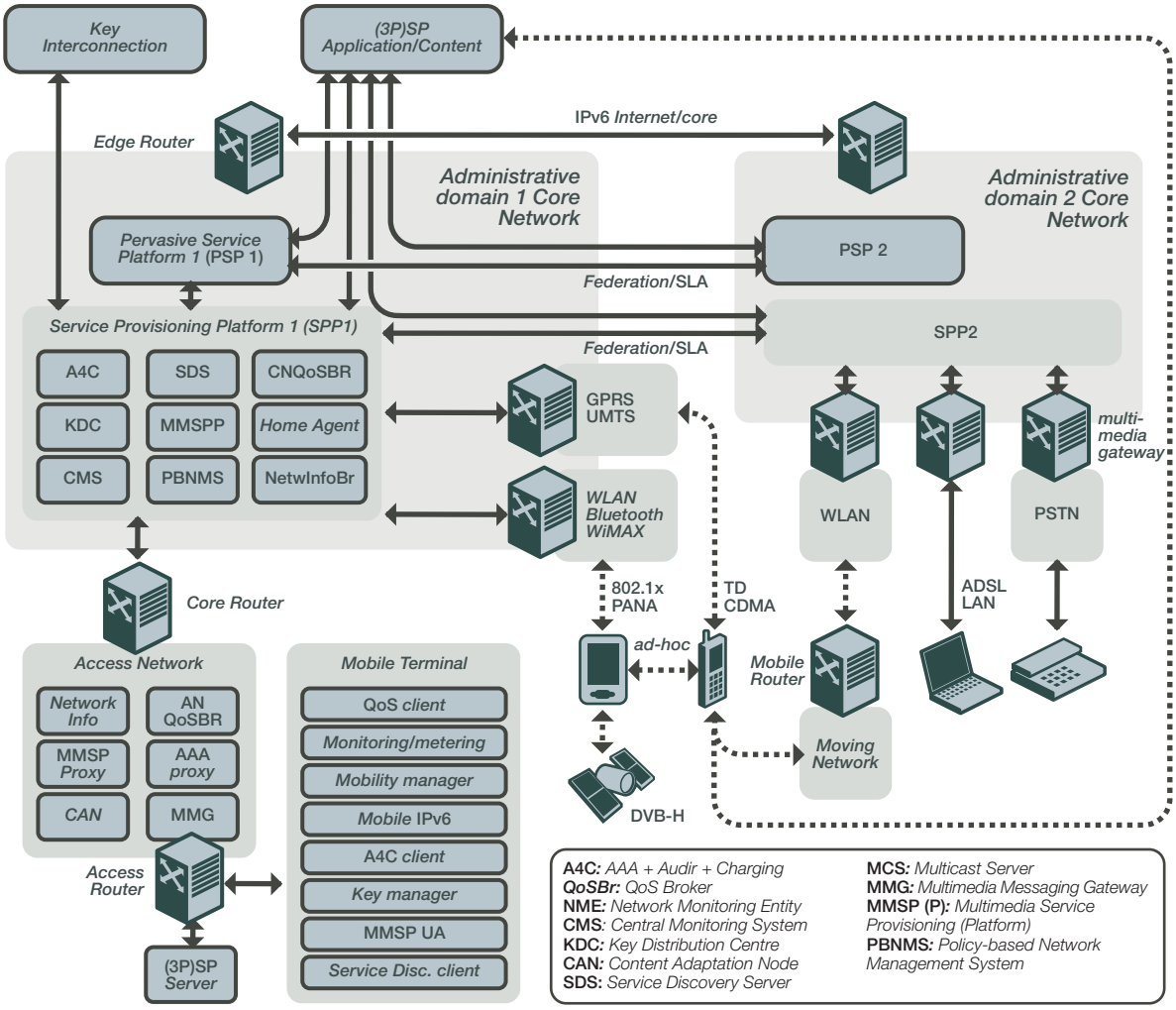

figura 1 - Arquitectura Daidalos

nível da tecnologia (DVB-S/T/H) e ao nível dos serviços (TV, *carouseles*, *datacast*,…);

> **Federação**, que permita aos operadores de rede e fornecedores de serviços a troca inter-domínio de utilizadores e serviços, seguindo um modelo de negócio dinâmico e flexível.

2.1 Arquitectura Daidalos

A arquitectura global do Daidalos encontra-se representada na figura 1. Esta distribui-se por servidores de serviços, terminais, redes de acesso e de núcleo (*core*), sendo composta pelas seguintes plataformas:

> **PSP** (*Pervasive Service Platform*): plataforma de serviços genéricos que oferece interfaces abertas a terceiros. É composta por um conjunto de servidores distribuídos (*Preference Manager, Context Manager, Service Discovery, Service Composition*,…), que através de interfaces *web*, oferece às aplicações desenvolvidas por terceiras partes (e.g. fornecedores de serviços) a possibilidade de aceder a informação personalizada dos utilizadores e de contexto, apresentando ainda outras funcionalidades como a descoberta e composição de serviços;

> **SPP** (*Service Provisioning Platform*): plataforma nuclear do Daidalos que inclui funcionalidades diversas como autenticação e autorização de utilizadores, gestão de identidades, estabelecimento de sessões multimédia, reserva de recursos (QoS) de rede, suporte a mobilidade nas suas várias vertentes, monitoria e contabilização;

> **AN** (*Access Network*): plataforma que se situa nas redes de acesso e dá suporte a algumas das funcionalidades descritas no SPP, como a autenticação e autorização de utilizadores, o estabelecimento de sessões multimédia, mobilidade e reserva de recursos;

> **MT** (*Mobile Terminal*): como a arquitectura Daidalos é por natureza distribuída, todas as funcionalidades anteriormente descritas terão a sua correspondente componente no terminal.

Podemos também verificar pela figura, o enfoque do Daidalos nos aspectos de inter-domínio e multi-tecnologia. A continuidade dos serviços, ou seja, a mobilidade de utilizadores, sessões e terminais é extensivamente estudada e implementada no Daidalos, tendo em conta este ambiente heterogéneo.

2.2 Multimédia e Qualidade de Serviço

A PT Inovação está especialmente envolvida nos subsistemas de *core*

AS: *Application Server*
SDS: *Service Discovery Server*
A4C: *Authentication Authorisation Accounting Auditing and Charging*
AG: *Accounting Gateway*
HSS: *Home Subscriber Service*
MMSPB: *Multimedia Service Provisioning Broker*
MMSPP: *Multimedia Service Provisioning Proxy*
MMSPUA: *Multimedia Service Provisioning User Agent*
ANQoSB: *Access Network QoS Broker*
CS: *Conferencing Server*
SS: *Streaming Server*

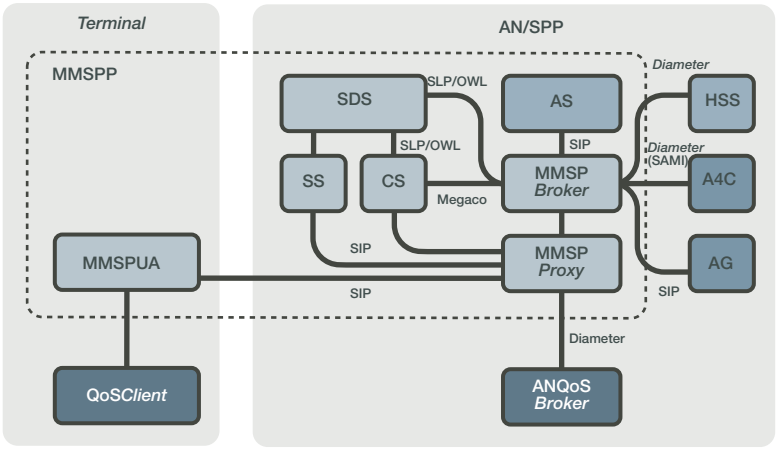

figura 2 - Plataforma Multimédia (MMSPP)

SPP (*Service Provisioning Platform*) e de AN (*Access Network*), posicionando-se claramente na área de interesse dos operadores. Desenvolve trabalho nos subsistemas de controlo de sessões multimédia e de reserva de recursos na rede, oferecendo a possibilidade aos operadores de estabelecer sessões de áudio e vídeo em redes IP (IPv6) com qualidade de serviço assegurada. Este trabalho surge na continuidade da iniciativa SHipNET [4].

2.2.1 Plataforma Multimédia

Esta plataforma (MMSPP – *MultiMedia Service Provisioning Platform*) segue as normas do 3GPP/IMS (organismo normalizador das redes móveis de próxima geração) e usa como base o componente da arquitectura IMS que controla o estabelecimento de sessões SIP, o CSCF (*Call Session Control Function*), desenvolvido no âmbito na iniciativa SHipNET. Devido aos requisitos Daidalos, este componente será enriquecido a vários níveis, nomeadamente no que se refere à autenticação e autorização de utilizadores, gestão das suas identidades virtuais (VID), capacidade de comunicar com outros servidores para localizar e adaptar conteúdos (e.g. *transcoders*) e, sobretudo, à possibilidade de fazer pedidos de reservas de recursos para assegurar QoS, inter-domínio extremo-a-extremo.

Na figura 2, podemos ver a arquitectura onde se insere o subsistema de controlo de sessões multimédia.

A plataforma multimédia inclui os seguintes componentes:

> **AS** (*Application Server*): serve, tal como previsto no IMS, a execução de todos os serviços e aplicações. Como complemento do trabalho efectuado na fase I do projecto, na fase II ligar-se-á o MMSPB ao servidor de aplicações X-TMAS[4] da PT Inovação. Será também necessário portar o código do AS Daidalos para o AS da PT Inovação. Essa operação permitirá mobilidade parcial de sessões multimédia, ou seja, será possível escolher, por exemplo, entre transferir apenas um dos fluxos, áudio ou vídeo, para um novo terminal, ou ambos;

> **MMSPB** (*MultiMedia Service Provisioning Broker*): é essencialmente um servidor SIP e tem como base funcional o I-CSCF (*Interrogating Call Session Control Function*) e o S-CSCF (*Serving Call Session Control Function*) desenvolvidos pela PT Inovação. O S-CSCF garante a intercomunicação com outros domínios IMS e trata dos registos dos utilizadores. Identifica também privilégios de utilizador/serviço e é responsável pela selecção e acesso ao AS. O I-SCCF é opcional entre o *Proxy* e o *Serving* de outro domínio; apenas existe se o "*home domain*" quiser esconder detalhes do S-CSCF (por exemplo, o seu endereço);

> **MMSPP** (*MultiMedia Service Provisioning Proxy*): é um SIP *proxy* e, portanto, o ponto de contacto inicial dos terminais. Restrições de rede e reservas de recursos devem ser aqui aplicadas, antes de fazer seguir os pedidos para o S-CSCF. Tem como base o P-CSCF (*Proxy Call Session Control Function*) da PT Inovação;

> **MMSPUA** (*MultiMedia Service Provisioning User Agent*): representa o utilizador SIP no seu terminal. Tem como base um cliente *open source* (SIP *Communicator* [2]) adaptado para funcionar como cliente IMS;

> **SDS** (*Service Discovery Platform*): componente que permite encontrar serviços, entre os quais os prestados por certos recursos na rede, como por exemplo os oferecidos por adaptadores de média (i.e. *transcoders*), servidores de conferência ou servidores de *streaming*;

> **CS** (*Conferencing Server*): é o um servidor de conferências e também *transcoder*;

> **SS** (*Streaming Server*): servidor de *streaming*.

A plataforma multimédia interage com outros componentes da arquitectura Daidalos, de onde se destacam os seguintes:

> **A4C** (*Authentication, Authorisation, Accounting, Auditing and Charging*): este componente encarrega-se de autenticar e autorizar os utilizadores e guardar parte dos perfis dos utilizadores (contratos e SLA);

A4C: *Authentication Authorisation Accounting Auditing and Charging*
AG: *Accounting Gateway*
MMSPP: *Multimedia Service Provisioning Proxy*
MMSPUA: *Multimedia Service Provisioning User Agent*
ANQoSB: *Access Network QoS Broker*
CNQoSB: *Core Network QoS Broker*
QoSClient: *QoS Client*
ARM: *Advanced Router Mechanism*
PBNMS: *Policy-Based Network Management System*
CMS: *Central Monitoring System*
NME: *Network Monitoring Entity*

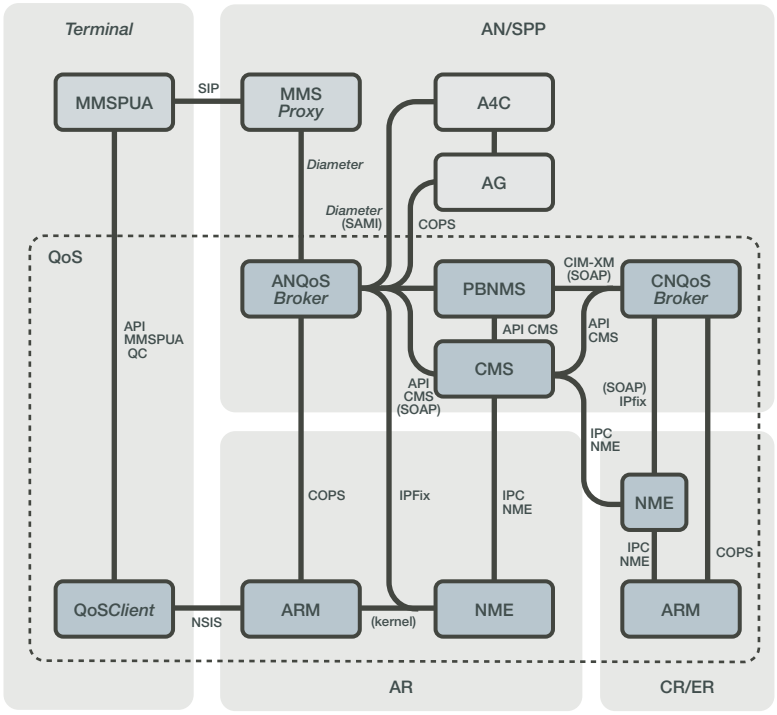

figura 3 - Plataforma de QoS

> **HSS** (*Home Subscriber Service*): é o componente IMS que conceptualmente desempenha algumas das funções executadas pelo A4C do Daidalos. Deste modo, no contexto do projecto, o HSS funcionará como complemento do A4C; no âmbito Daidalos será usado apenas para guardar e ceder informação IMS de utilizador ao MMSPB, ou seja, informação relativa aos pontos de *trigger* para o AS e qual o S-CSCF atribuído a cada utilizador;

> **AG** (*Accounting Gateway*): encarrega-se de recolher dados das sessões SIP, para se poder fazer a taxação, com base no tempo de duração das sessões ou em eventos;

> **ANQoSB** (*Access Network QoS Broker*): no contexto multimédia, este componente serve para reservar recursos na rede, por sessão, a pedido do MMSPP. A explicação mais detalhada deste componente encontra-se no próximo capítulo.

2.2.2 Plataforma de QoS

A plataforma de QoS segue várias normas, entre as quais se destaca o TISPAN [3], organismo normalizador que usa a arquitectura do 3GPP/IMS e adapta-a à rede fixa. O TISPAN introduziu um novo subsistema na arquitectura, a que chamou RACS (*Resource and Admission Control Subsystem*) que serve para controlar os recursos na rede e consequentemente a qualidade de serviço. Os componentes Daidalos QoSB seguem as normas TISPAN e implementam, para além disso, requisitos específicos do Daidalos. Estes componentes participam também na arquitectura global SHipNET [4], da PT Inovação.

A arquitectura onde se insere a plataforma de QoS está representada na figura 3.

Esta plataforma inclui os seguintes componentes:

> **ANQoSB** (*Access Network QoS Broker*): este componente assegura, de acordo com o modelo *DiffServ*, que a qualidade de serviço seja aprovisionada nos *routers* da rede de acesso. Participa também no processo de *handover* (mudança do ponto de ligação do terminal à rede) garantindo que os recursos são reservados correctamente no novo ponto de acesso. Para os serviços *legacy* (por exemplo HTTP e FTP), o controlo de admissão é executado pelos ANQoSB a pedido do ARM (ver descrição na continuação do texto) e pode ser baseado num DSCP (*DiffServ Code Point*) pré-configurado ou em sinalização explícita (por exemplo RSVP ou NSIS); no caso dos serviços SIP é apresentado um exemplo no próximo capítulo;

> **CNQoSB** (*Core Network QoS Broker*): esta entidade é responsável pela gestão de recursos na rede de núcleo e também serve de *proxy* na mobilidade inter-domínio. Por razões de escalabilidade o CNQoSB gere agregados de fluxos e não fluxo a fluxo como o ANQoSB;

> **QoSClient** (QoS *Client*): reside no terminal e tem como responsabilidade iniciar e manter as reservas pedidas pelo MMSPUA ou por outras aplicações que residam no terminal; pode também informar outros componentes sobre os recursos disponíveis na rede de acesso e marcar os pacotes à saída do terminal com os valores de DSCP apropriados;

> **ARM** (*Advanced Router Mechanism*): este componente deve residir nos *routers* de acesso e oferecer a todos os outros

componentes de operador (e.g. QoSB, PBNMS), uma interface genérica que lhes permita abstrair da heterogeneidade de tecnologias de rede. O ANQoSB interage com o ARM, sempre que precisa de configurar políticas de controlo de admissão nos *routers* de acesso. Para os serviços *legacy* o ARM pergunta ao QoSB o que fazer com os pacotes que recebe. Note-se que todo o tráfego a cursar sobre a rede deve ser admitido/autorizado pelo ANQoSB implícita ou explicitamente;

> **PBNMS** (*Policy-Based Network Management System*): armazena as políticas (de *operator*) de configuração da rede e a definição de parâmetros de QoS. O ANQoSB recebe estas políticas do PBNMS e gere os recursos e a largura de banda de acordo com essas políticas;

> **CMS** (*Central Monitoring System*): existirá um por domínio a monitorizar em tempo real, testando e recolhendo estatísticas. Os QoSB podem usar esta informação para tomar decisões sobre a atribuição de recursos (CAC – *Call Admission Control*). O CMS controla os NME e recolhe medidas dos vários NME distribuídos na rede;

> **NME** (*Network Monitoring Entity*): efectua medidas de QoS activas ou passivas, ou seja, intrusivas ou não intrusivas. O CMS programa nos NME quais as medidas a efectuar e com que frequências e, os NME, por sua vez, assim que recolhem as medidas reportam-nas ao CMS.

2.3 Protocolos e Serviços

Protocolos

O Daidalos usa protocolos definidos essencialmente pelo 3GPP/IMS e IETF. Alguns destes protocolos sofrerão melhoramentos ao longo do projecto, dependendo dos requisitos internos de mobilidade, segurança, etc. Os mais importantes, na perspectiva das sessões multimédia e correspondente reserva de recursos, são:

> **SIP** (*Session Initiation Protocol*) e **SDP** (*Session Description Protocol*): à luz do IMS, o Daidalos usa o protocolo SIP para controlar sessões multimédia e o SDP para caracterizar as capacidades dos terminais, definindo o tipo de sessão que se quer estabelecer (e.g. áudio, vídeo, *codecs*);

> **RTP** (*Real Time Protocol*) e **RTCP** (*Real Time Control Protocol*) para o transporte e controlo de fluxos IP multimédia;

> **COPS** (*Common Open Policy Service*) para o controlo de recursos (QoS) nos *routers* (i.e. ARM);

> **DIAMETER** para a autenticação, autorização e contabilização (actividades relacionadas com o A4C e HSS) e também para a reserva de recursos (interfaces entre QoSB e entre o MMSPB e o ANQoSB);

> **NSIS** (*Next Steps in Signalling*) **/RSVP** (*Resource Reservation Protocol*) e o modelo *DiffServ* para assegurar QoS extremo-a-extremo.

Serviços

Para mostrar e validar o potencial desta arquitectura, o projecto Daidalos tem como referência dois cenários:

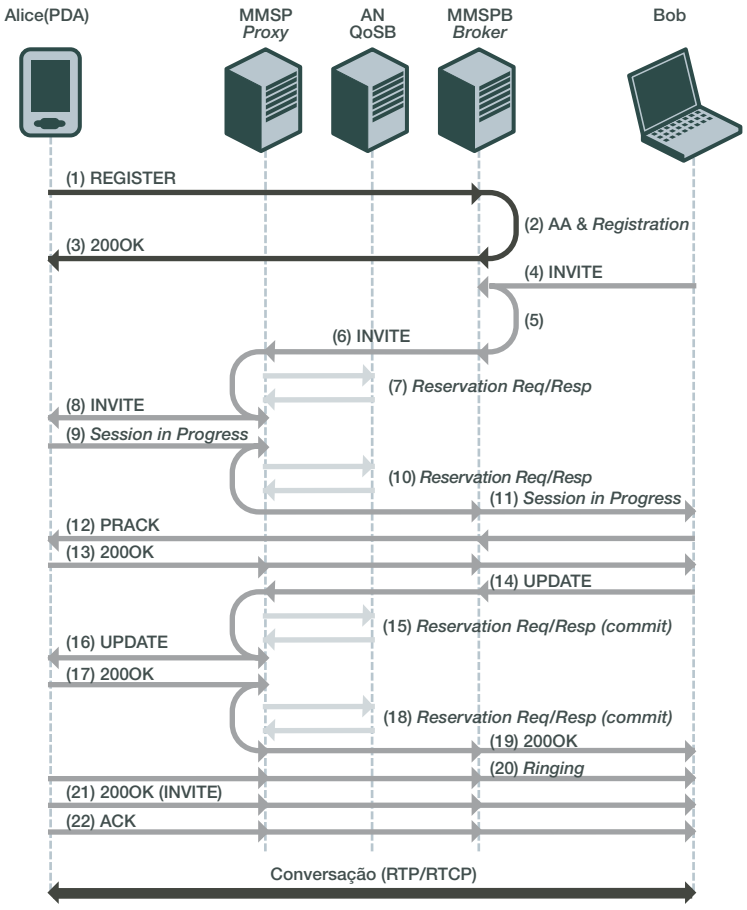

figura 4 - Sessão Multimédia e Reserva de Recursos

> **Universidade Móvel:** Suporte de mobilidade em diversas instituições académicas. O público alvo são estudantes, professores e outros eventuais utilizadores que se relacionem com universidades de toda a Europa e que tenham necessidade de utilizar os seus serviços em diversos ambientes e contextos. A ideia é oferecer mais flexibilidade e mobilidade aos utilizadores, promovendo a capacidade de personalização e adaptação dos serviços aos mais diversos contextos. Este conceito poderá ser alargado às chamadas "empresas móveis";

> **Automóvel:** Este cenário tem como objectivo facilitar o lazer nos automóveis. Os serviços devem manter a sua continuidade com o automóvel em movimento bem como na aproximação e afastamento dos utilizadores destes. Este cenário poderá também ser utilizado em aviões, comboios e outros ambientes onde a mobilidade seja crucial.

Destes serviços globais, destacamos para este artigo um cenário que realça o desempenho dos componentes desenvolvidos pela PT Inovação, nomeadamente o MMSPB/P e os QoSB.

2.4 Cenário

O cenário aqui descrito ocorrerá várias vezes durante a execução dos serviços globais Daidalos anteriormente referidos. Será executado sempre que se estabelecer uma sessão multimédia (por exemplo VoIP, *streaming* de vídeo, Vídeotelefonia) sendo para tal necessário reservar recursos na rede.

Na figura 4, encontra-se representado o diagrama (simplificado) de mensagens trocadas entre as principais entidades envolvidas neste cenário:

> **1-3:** Considerando dois utilizadores, a Alice e o Bob, ambos terão de pedir inicialmente acesso à rede, usando para isso uma chave+certificado (ID*Token*) (processo não mostrado na figura 3) e, de seguida, registar os seus endereços SIP no MMSPB (SIP *Server*). Na figura 4, para simplificar, representamos apenas o processo de registo SIP da Alice. Tanto o pedido de acesso à rede como a autenticação dos utilizadores SIP, são efectuados pelo A4C que em troca devolverá o perfil do utilizador; a parte do perfil relativa à rede é entregue ao ANQoSB e a parte do perfil relativa aos serviços ao MMSPB. Estes perfis determinam os direitos de cada utilizador; contêm parâmetros de (QoS) de rede (*delay, jitter, loss, DiffServ Class*) e parâmetros de serviço (tipo de média e *codecs* correspondentes), respectivamente. Servem para autorizar o utilizador a iniciar uma determinada sessão multimédia e efectivar a correspondente reserva de recursos;

> **4-8:** de seguida o Bob (já registado) envia um INVITE para iniciar a sessão multimédia com a Alice. Nesta mensagem, é também enviado o SDP, o qual caracteriza a sessão a estabelecer, e as capacidades do terminal do Bob (por exemplo, sessão áudio com *codecs* PCMa, PCMu e GSM). No decorrer deste processo o MMSPB verificará (a partir do seu perfil) se o Bob tem direitos para utilizar o serviço invocado (mensagem 5) e o MMSPP, por sua vez, verificará a disponibilidade dos recursos de rede interagindo com o ANQoSB (mensagem 7);

> **9-11:** a Alice responde ao pedido do Bob aceitando-o mas restringindo o leque de oferta de *codecs* do Bob àqueles que o seu terminal pode suportar (por exemplo, áudio, com *codec* PCMa ou PCMu; qualquer um destes *codecs* poderá ser usado pelo Bob na transmissão do *media*). Entretanto, o MMSPP procede à actualização da reserva interagindo com o ANQoSB (mensagem 10), sendo que neste momento, o MMSPP já dispõem de informação relativa aos endereços IP e respectivos portos de recepção do média, para ambos os utilizadores;

> **12-13:** troca de mensagens para decidir, entre os possíveis qual o *codec* a utilizar na sessão (e.g. PCMa);

> **14-19:** actualização do estado do terminal relativamente às reservas. Basicamente, sinaliza que tudo está preparado no terminal para iniciar a sessão, no que diz respeito à reserva de recursos. Deste modo, o MMSPP aproveita para fazer o "*commit*" da reserva (mensagens 15 e 18) junto do ANQoSB;

> **20-21:** o *Ringing* é gerado e, assim que a Alice atende, é dado o OK (mensagem 21) ao convite do Bob;

> **22:** por último, o Bob envia um ACK, terminando assim a fase de sinalização. A sessão RTP/RTCP pode então ser iniciada com QoS assegurada.

3. Conclusões

A actividade desenvolvida no projecto Daidalos, permitiu à PT Inovação contribuir significativamente para a demonstração, dentro e fora do projecto, do que poderá ser um futuro sistema B3G, através da plataforma instalada em Aveiro. Para além do projecto, este trabalho reflecte-se na iniciativa interna SHipNET, que tem por objectivo garantir a continuidade do trabalho da PT Inovação, na área de desenvolvimento para o operador de plataformas de serviços e controlo de rede. Nesse âmbito, conseguiu-se com o projecto Daidalos o desenvolvimento e experimentação de protótipos de componentes controladores do estabelecimento de sessões multimédia e que permitem gerir a qualidade de serviço na rede.

A fase II do projecto, iniciada em Janeiro e que deverá concluir-se no final de 2008, permitirá evoluir as soluções da fase I, melhorando processos de mobilidade, segurança e controlo de qualidade, entre outros.

Referências

[1] http://www.3gpp.org
[2] http://www.sip-communicator.org
[3] http://portal.etsi.org/
[4] Arquitectura IMS, Saber & Fazer 2005

Telma Mota, licenciatura e mestrado em Engenharia Electrotécnica e de Computadores pela Universidade do Porto, em 1988 e 1994, respectivamente. Ingressou na empresa TLP SA em 1988, onde realizou trabalho de planeamento e dimensionamento de redes de comutação digital, actividades na área de teletráfego e esteve também ligada ao planeamento, avaliação e dimensionamento de Redes Inteligentes. Desde 1994, que integra a PT Inovação quase sempre ligada à área de gestão de redes e serviços. Participou em diversos projectos Europeus (Eurescom e IST); actualmente, é responsável pela área de Serviços e Plataformas e está envolvida nos projectos Daidalos, Muse e OPUCE. Centra o seu interesse na área de análise e definição de arquitecturas de rede e modulação de sistemas.

Francisco Fontes, licenciado em Engenharia Electrotécnica e de Computadores, Ramo de Electrónica e Telecomunicações, pelo Instituto Superior Técnico (Set/91) e Doutorado pela Universidade Politécnica de Madrid (Nov/00) na área de gestão distribuída de redes de telecomunicações. Em Set/91 iniciou a sua actividade profissional na PT Inovação (CET) tendo-se especializado em tecnologias de rede de banda larga, especialmente nas tecnologias ATM e IP. Actualmente os seus interesses situam-se nas áreas de Qualidade de Serviço em redes multi-serviço e na sua evolução para arquitecturas RPG All-IP.

03

Controlo de Qualidade num Cenário VoD numa Rede xDSL e MPLS/Ethernet

Francisco Fontes, Ricardo Azevedo, André Oliveira (IT), Gonçalo de Lemos (IT), Jorge Loura (IT), João Ferreira (PTCom), Pedro Bandeira (PTCom)

palavras-chave:
VoD, CAC, RTSP, *DiffServ*, Opnet++, RACS, *Ethernet*, MPLS

O controlo de admissão e a monitorização do nível de desempenho, são indispensáveis em redes *all-IP*, quando os serviços e aplicações a utilizar requerem QoS, partilhando uma mesma infra-estrutura física ou lógica. Nessa situação, é necessária a autorização da utilização de determinados recursos de forma a garantir a obtenção da qualidade esperada, sem degradação dos serviços já activos. Este tipo de actividade torna-se especialmente importante quando aplicado às secções da rede de acesso e agregação/distribuição da rede, onde os recursos são mais escassos e o seu sobredimensionamento não é possível ou é muito dispendioso. A necessidade deste tipo de controlo é reconhecida pela sua inclusão em recentes trabalhos de normalização na área das redes de próxima geração. O subsistema TISPAN/RACS (*Resources and Admission Control Subsystem*), em conjunto com o TISPAN/NASS (*Network Attachment Sub-system*), pretende fazer esse tipo de controlo, através de interfaces apropriadas com componentes derivados do IMS.

Neste contexto, surgem soluções *triple-play* (voz, vídeo e dados) onde se incluem serviços VoD (*Video on Demand*), disponibilizados sobre redes de agregação/distribuição utilizando tecnologia xDSL, MPLS e *Ethernet*. Nesta situação é necessário validar a disponibilidade de recursos (transporte e plataformas de serviço) e autorizar, ou não, em tempo real, cada pedido de acesso ao serviço.

A presente publicação pretende descrever o contexto técnico actual para a disponibilização de serviços VoD sobre tecnologias xDSL/ *Ethernet*/MPLS, incluindo requisitos e soluções aplicáveis nesse contexto para o controlo de admissão. Descrever-se-ão os estudos, simulações e desenvolvimento efectuados na área de *Call Admission Control* (CAC) e a sua aplicação no desenvolvimento de entidades de controlo do nível de desempenho da rede.

1. Introdução

A evolução observável nas redes de telecomunicações, considerando o mercado, os organismos de normalização e as áreas de pesquisa e investigação, indica a adopção generalizada da tecnologia IP (muito certamente na sua versão 6 [1]) para a disponibilização da generalidade dos actuais e futuros serviços de telecomunicações. Essa evolução está a ocorrer no âmbito das Redes de Próxima Geração (RPG ou NGN [2] – *Next Generation Networks* – na terminologia anglo-saxónica).

Essas redes apresentam uma organização por camadas, estabelecendo uma clara separação entre transporte, controlo e aplicações/serviços. Essa arquitectura proposta para as redes RPG apresenta diversas vantagens, permitindo a evolução de uma dessas camadas sem impacto nas restantes e o desenvolvimento de novos serviços convergentes, disponibilizados de forma independente da tecnologia de acesso, terminal utilizado e localização do utilizador. A convergência fixo-móvel é uma das facetas mais visíveis e que maiores potencialidades apresenta.

Neste âmbito, o ETSI [3] encontra-se a definir a arquitectura TISPAN [4]. Nesta, os dois sub-sistemas RACS (*Resources and Admission Control Sub-system*) e NASS (*Network Attachment Sub-system*) constituem o nível de controlo, servindo de mediadores entre os componentes de serviço e a rede (acesso e transporte). O primeiro deve garantir que qualquer serviço apenas se inicia se existirem os recursos necessários na rede (controlo de admissão). O segundo executa funções de autenticação, autorização, gestão de endereçamento IP e localização.

Os algoritmos de controlo de admissão devem maximizar a utilização dos recursos disponíveis, mantendo os níveis de qualidade estabelecidos (por exemplo, valores máximos de atraso e sua variação, e de perdas). O seu funcionamento, em especial no núcleo da rede, rege-se por análise estatística, apenas aplicável quando o número de fluxos é elevado. Já nos acessos, a análise terá de ser feita por fluxo individual levando a decisões mais conservadoras.

A par da tecnologia IP, outras duas tecnologias serão incontornáveis na implantação das RPG: a *Ethernet* [5] e o MPLS [6]. O primeiro deve o seu sucesso à sua simplicidade de operação, alto débito e baixo custo de implantação, o segundo à flexibilidade na disponibilização de novos serviços de rede e aumento que induz nas prestações e disponibilidade das redes que o adoptem. Estamos assim perante três elementos que, em sintonia, permitem a constituição de redes flexíveis e com elevadas prestações.

Também no segmento de acesso se observa um aumento significativo das prestações onde novas soluções (WiMAX e ADSL2+, por exemplo), resultantes da evolução tecnológica recente, vieram permitir a convergência de redes e serviços possibilitando a disponibilidade de alguns serviços que antes necessitavam de uma infra-estrutura própria.

Esse é o caso dos serviços de televisão (IPTV) e vídeo (VoD) que requerem elevados débitos para o transporte (*streaming*) dos conteúdos.

O cenário descrito nesta publicação analisa precisamente a disponibilização de um serviço VoD sobre uma moderna infra-estrutura multiserviço que explora as tecnologias *Ethernet*, MPLS e ADSL, com transporte IP fim a fim, seguindo um modelo em linha com as RPG. É aqui dada ênfase à componente de verificação de recursos que garante a disponibilização do serviço (visualização do vídeo pretendido) apenas se existirem condições de rede para tal. Essa verificação (execução de algoritmos de controlo de admissão) deve garantir que o serviço é prestado com a qualidade esperada e que os serviços já activos não se degradem com a admissão de novos fluxos VoD na rede.

1.1 Serviço *Video on Demand* (VoD)

O serviço de *Video on Demand* (VoD) tradicional tem como objectivo providenciar ao cliente, o acesso a conteúdos multimédia no momento em que este deseja, utilizando para o efeito, uma infra-estrutura de telecomunicações (figura 1).

A arquitectura do serviço VoD assenta numa lógica cliente-servidor onde o servidor de vídeo aguarda que o equipamento do cliente (*Set-Top Box* – STB) lhe envie um pedido referente a um conteúdo que se encontre disponível para o serviço VoD.

Este pedido é efectuado pela utilização do protocolo *Real-Time Streaming Protocol* (RTSP [7]), o qual, para uma sessão comum, apresentaria a troca de mensagens representada na figura 2.

Na sequência ilustrada existem pontos onde terão de ocorrer decisões, tais como verificação de

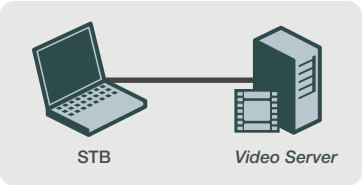

figura 1 – Serviço VoD tradicional

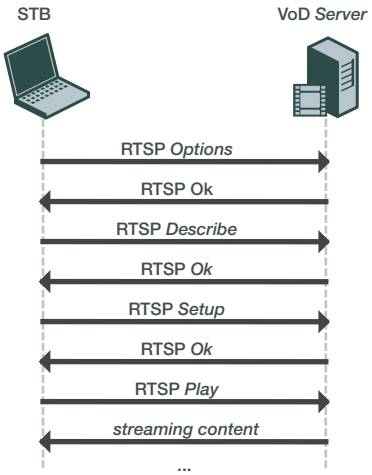

figura 2 – Estabelecimento de uma sessão RTSP

disponibilidade do conteúdo, autorização de acesso para o cliente em questão, disponibilidade da rede e dos servidores de vídeo, entre outros. É durante este processo que se deve proceder às acções de controlo e gestão da qualidade de serviço (QoS).

1.2 Mecanismos de QoS e pedidos de recursos

Com a adopção generalizada do IP para todo o tipo de serviços, em todo ou apenas em parte do percurso entre o fornecedor e os consumidores, o termo Qualidade de Serviço (QoS) generalizou-se. O problema de QoS consiste em desenvolver mecanismos que, sobre uma mesma infra-estrutura onde os recursos disponíveis são partilhados, permitam que serviços com requisitos bastante diversos possam coexistir e obter da rede o comportamento

	atraso	var. atraso	erros
Voz	+	++	–
Vídeo	–	–	+
Dados	– –	– – –	++

tabela 1 – Sensibilidade de serviços

desejado e que terá de ser, naturalmente, diferenciado. Na tabela 1 é indicada a sensibilidade de três tipos de componentes multimédia base a três efeitos das redes IP.

Esses mecanismos são diversos e existem a diversos níveis, sendo de especial relevância os disponibilizados pelas tecnologias *Ethernet* (IEEE 802.1q/p [8]), MPLS (mecanismos de engenharia de tráfego e *bits* CoS [6]) e pelo próprio IP (*Differentiated Services* [9]). Todos estes partilham o facto de marcarem cada pacote ou trama à entrada da rede numa de várias classes de transporte disponíveis. O processo de marcação pode ser complexo, considerando o tipo de serviço, o contrato de tráfego existente entre cliente e prestador do serviço e o perfil real do tráfego entregue à rede. A gestão da QoS faz-se assim por agregados e não por fluxo individual, sendo garantida a escalabilidade das soluções. No entanto, a definição de classes de transporte, por si só e de uma forma genérica, apenas garante que o tráfego agregado de uma classe obterá um tratamento diferenciado relativamente às restantes classes, considerando um ou vários de um conjunto de parâmetros (por exemplo atraso e perdas). A garantia, para cada uma dessas classes, de se operar abaixo de valores máximos para esses parâmetros, consegue-se com a adição de mecanismos de pedido de recursos e mecanismos de controlo de admissão. Este último poderá ter de ser complementado por uma monitorização adequada e completa da rede.

Os pedidos de recursos poderão ser efectuados de diversas maneiras, destacando-se três:

1. **Pedido explícito despoletado pelos terminais** e utilizando um protocolo apropriado; actualmente o RSVP e, no futuro, o NSIS são os melhores candidatos para esta função;

2. **Pedido explícito despoletado por elementos de controlo na rede**, derivado da análise de sinalização de nível de aplicação (por exemplo SIP ou RTSP);

3. **Pedido implícito**; pedidos implícitos adequam-se a fluxos provenientes de serviços em que não existe um estabelecimento prévio de uma sessão anterior ao envio dos pacotes com informação (por exemplo na utilização dos protocolos HTTP e FTP). Para estes, um conjunto de regras, associadas ao perfil do utilizador, às suas preferências e possivelmente ao contexto, e que têm de ser descarregadas para os elementos de rede e/ou terminais, determinam a forma de efectuar a marcação dos pacotes IP.

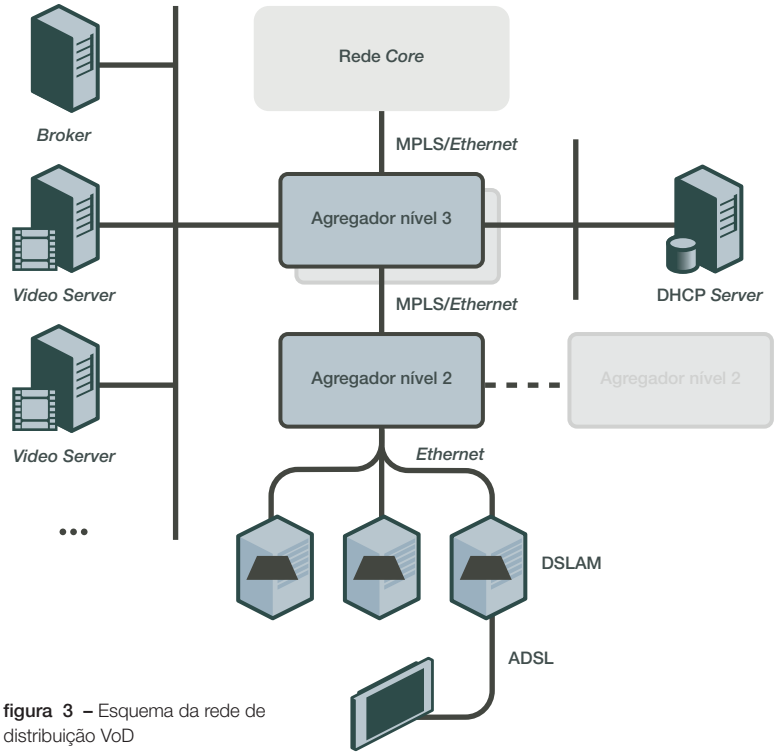

figura 3 – Esquema da rede de distribuição VoD

No contexto desta publicação, tem especialmente interesse a situação de pedidos explícitos derivados da análise de sinalização RTSP. O protocolo RTSP não inclui elementos de informação nas mensagens trocadas que permitam retirar directamente requisitos de QoS. É da análise, efectuada por elementos adicionais colocados no caminho da sinalização RTSP, que se pode inferir o nível de QoS pretendido e os recursos de rede necessários para a sua satisfação.

2. Descrição de uma rede de suporte ao serviço VoD

Os clientes do serviço VoD ir-se-ão ligar a uma rede de distribuição multi-serviço com capacidade de transportar débitos de informação correspondentes a fluxos de vídeo. Na residência do cliente terá de existir uma STB capaz de interagir com a plataforma de VoD (servidores e *middleware*), acompanhada de um *modem* ADSL com capacidade para usar várias ligações ATM (VC - *Virtual Circuits*) em simultâneo (separação de serviços).

A rede de distribuição, como se pode observar na figura 3, é composta por uma rede de acesso ADSL, que compreende vários agregadores de linha (DSLAM). Cada um destes DSLAM liga a um agregador de nível 2 através de *Ethernet*, com capacidade para 1 *Gbit*/s ou superior. Neste é agregado o tráfego relativo aos vários DSLAM a ele ligados e é efectuada a interface entre a rede de acesso e a rede de transporte MPLS. Esta é suportada por duas ligações ponto-a-ponto *Ethernet* em regime de partilha de carga, com capacidade superior a 1 *Gbit*/s, que irá fornecer mecanismos de transporte e redundância necessários para garantir um sistema de distribuição robusto e eficiente. Nesta ligação existirão VLAN [8] (*Virtual LAN*), tipicamente uma por cliente. Estas permitem a existência de um contexto sobre o qual se podem exercer políticas de operação (QoS, filtragem, contabilização).

Esta rede MPLS irá assegurar a conectividade através de uma VPLS (cada serviço e operador deverá ter a sua própria VPLS [6] - *Virtual Private LAN Service*) entre os elementos de agregação de nível 2 e os elementos de agregação de nível 3. Esta rede de acesso e distribuição, além de permitir o uso do serviço VoD por parte dos clientes, também servirá simultaneamente para fornecer serviço de Internet ou mesmo serviço de televisão sobre IP (IPTV).

Os servidores de conteúdos vídeo irão ficar ligados a alguns dos agregadores de nível 3 da rede num segmento também *Gigabit Ethernet*. A rede para o aprovisionamento do serviço VoD pode assim ser segmentada nos seguintes domínios:

> **Rede do cliente** (não representada na figura 3); delimitada pela STB e pelo equipamento de acesso à rede (*router* ou *modem* xDSL);

> **Rede de acesso**; delimitada pelo equipamento de acesso à rede e os DSLAM;

> **Rede de distribuição/ agregação**; delimitada pelos DSLAM e os elementos de agregação de nível 2;

> **Rede de transporte**; delimitada pelos elementos agregadores de nível 2 e agregadores de nível 3;

> **Rede do serviço**; delimitada pelos elementos agregadores de nível 3 e os próprios servidores VOD.

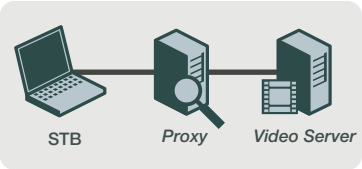

figura 4 – *Proxy Interceptor*

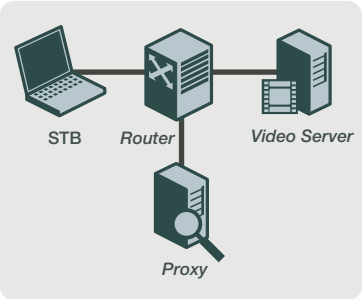

figura 5 – Desvio de mensagens por intermédio de um *router*

Devido à atribuição dinâmica de endereços de rede (endereço IP) aos clientes por um servidor DHCP, não é possível, de forma simples, obter uma correlação directa entre o endereço IP presente na negociação VoD (que é o único elemento que poderá identificar univocamente o cliente envolvido) e a localização (linha ADSL) do mesmo. No entanto essa localização é necessária para a determinação do percurso dos fluxos VoD na rede.

Para solucionar este problema, terá que se configurar a funcionalidade de *DHCP Relay* [10] dos DSLAM para inserir um campo específico nos pedidos de endereço (*DHCP Request*) feitos pelos clientes ao servidor DHCP: campo opcional com código 82, denominado de *Relay Agent Information*. Neste campo será inserido o código identificativo da linha telefónica (número telefónico) e do DSLAM de onde provém o pedido, possibilitando assim uma associação entre cada endereço IP, atribuído pelo servidor DHCP na rede, e a localização física desse mesmo endereço. Esta informação será obtida e analisada pelo elemento gestor dos recursos de rede (*Bandwidth Broker* – BB).

3. Integração do controlo de recursos num serviço VoD

A solução de controlo pretendida para o serviço VoD deverá integrar-se de forma transparente na plataforma. Isto deve-se ao facto de as plataformas VoD nem sempre contemplarem no seu processo de decisão uma verificação da disponibilidade de recursos de rede. A solução escolhida para a integração dessa funcionalidade, sem alterações nesses componentes, passa pela existência de um componente com comportamento semelhante a um *proxy interceptor*. Este tipo de elemento encontra-se no percurso entre o cliente e o servidor, interceptando e analisando mensagens (figura 4). Com este componente é possível inferir, de uma forma tão exacta quanto possível, requisitos de QoS associados aos conteúdos pretendidos pelo cliente, confrontá-los com a situação actual da rede e tomar uma decisão relativamente ao pedido.

Ao efectuar o pedido de visionamento de um determinado conteúdo, o cliente entra em contacto com o servidor, por meio do protocolo RTSP (ver secção 1.1). As mensagens trocadas são analisadas pelo *proxy* que verifica a disponibilidade de recursos na rede (por consulta ao BB) para satisfazer o pedido. Caso estes não sejam suficientes, o *proxy* assume o papel de servidor enviando uma mensagem de erro RTSP ao cliente; na situação de existir capacidade para satisfazer o pedido, o componente limita-se a servir de *proxy* transparente entre o servidor e o cliente, na troca de mensagens RTSP. Aplicando esta filosofia ao cenário real, efectua-se a ligação do *proxy* ao agregador de nível 3, o qual desvia todas as mensagens de RTSP para o *proxy* colocando-o no percurso entre o STB e o servidor de vídeofigurara 5).

Desta forma o *proxy* consegue obter os endereços IP do servidor de VoD e do STB, a duração e o *bitrate* do filme, informação que permite a elaboração de uma mensagem que é enviada ao gestor de recursos, por meio de uma interface do tipo Gq', normalizado pelo TISPAN.

4. Controlo de Admissão de Chamadas

O Controlo de Admissão de Chamadas ou *Call/Connection Admission Control* (CAC) é o processo de decisão sobre se um novo fluxo deve ser aceite ou rejeitado para transporte sobre uma rede, com o objectivo de assegurar a não degradação dos serviços em curso e a qualidade a receber por esse novo fluxo. Os algoritmos de controlo de admissão podem dividir-se em três grupos, conforme estes sejam baseados em (I) parâmetros, (II) medições ou (III) provas.

Os algoritmos de CAC baseados em parâmetros supõem o conhecimento dos modelos de tráfego das fontes. Com esta informação, quando um novo pedido é gerado, o controlo de admissão verifica o total de recursos consumidos com base nos pedidos aceites anteriormente, ao longo do percurso que o novo fluxo deverá percorrer, actualiza este valor com o volume de recursos requerido pelo novo fluxo e só aceita se o resultado obtido for inferior à largura de banda total disponível nos vários segmentos da rede envolvidos. O desempenho deste método depende essencialmente do grau de exactidão dos parâmetros de tráfego anunciados, já que não se efectuam medições ao tráfego existente na rede.

Os algoritmos de controlo de admissão baseados em medições da rede tomam as suas decisões a partir de informações obtidas por mecanismos de monitorização da rede. Assim, o controlo de admissão não tem de determinar os recursos reservados com base nos modelos de tráfego e no número de sessões activas, tendo apenas de se basear nos valores de utilização das ligações obtidos pelas medições. Este método tem a vantagem de possibilitar uma caracterização mais

simples dos fluxos, como por exemplo, a especificação da taxa média ou de pico do novo fluxo. Também requer capacidades computacionais menores à custa da necessidade de receber, com periodicidade adequada, o estado real da rede. De forma a garantir escalabilidade, o mecanismo de monitorização efectuará medições aos agregados de fluxos e não fluxo a fluxo. É com este método que as características estatísticas são mais facilmente obtidas, consequência da lei dos grandes números. O principal problema deste método está relacionado com o grau de exactidão das medições. Adicionalmente, este grupo de algoritmos, baseando-se nas medições recebidas, pode autonomamente ajustar um conjunto de parâmetros de forma a alterar o grau de agressividade. São, então, denominados de algoritmos adaptativos.

Nos algoritmos de controlo de admissão baseados em provas (*probe-based* ou *endpoint-based*), o terminal/aplicação envia um conjunto de pacotes prova pelo caminho utilizado pelos pacotes reais com o objectivo de testar as capacidades da rede. Com base no *feedback* fornecido pelo destino é o próprio terminal a decidir se o novo fluxo pode ser admitido, deixando de ser necessário um modelo centralizado para efectuar o controlo de admissão. Este método introduz atrasos no estabelecimento das sessões e tem problemas associados, causados pelo consumo de largura de banda por parte dos pacotes prova e pela negação de serviço quando pedidos simultâneos congestionam a rede e nenhum desses pedidos é aceite, havendo, embora, recursos disponíveis.

4.1 Estudos dos algoritmos de CAC

Foi efectuado um estudo, com recurso a um simulador de redes (OMNeT++ [11]) sobre algoritmos de controlo de admissão que integram características dos tipos indicados acima. O cenário de simulação construído é composto por uma entidade (BB – *Bandwidth Broker*) que executa o controlo de admissão e um conjunto de fontes ON-OFF que produzem sessões aleatoriamente (figura 6).

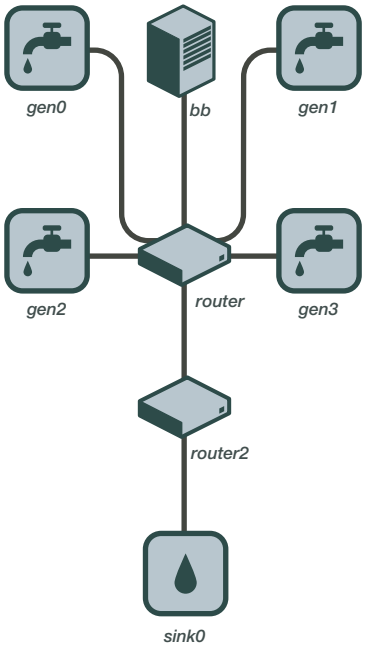

figura 6 – cenário simulado

Cada fonte inicia uma sessão com o envio de uma mensagem *Request* ao BB, mensagem que inclui a taxa de pico. O BB processa o pedido e responde com uma mensagem *Accept* ou *Reject* consoante o teste de admissão seja ou não bem sucedido. As fontes enviam os pacotes gerados por uma ligação ponto a ponto entre dois *routers* com destino a um receptor. Um dos routers mede a utilização da rede e periodicamente envia as medições ao BB. Os algoritmos relevantes abordados neste estudo descrevem-se de seguida.

O algoritmo *Simple Sum* (SS) admite um novo fluxo se a soma da taxa de transmissão requerida com a largura de banda reservada para os fluxos existentes da mesma classe for inferior à utilização alvo:

$$r + v < \alpha$$

onde

> r é a taxa de transmissão do novo fluxo;

> v é a largura de banda reservada para os fluxos anteriormente admitidos;

> α é a utilização alvo da ligação.

O algoritmo *Measured Sum* (MS) admite um novo fluxo se a soma da taxa de transmissão requerida com a estimativa da taxa de transmissão do agregado de fluxos existentes da mesma classe for inferior à utilização alvo:

$$r + \mu < \alpha$$

onde

> r é a taxa de transmissão do novo fluxo;

> μ é a taxa de transmissão do agregado de fluxos estimada;

> α é a utilização alvo da ligação.

Adicionalmente, por cada fluxo admitido o *bandwidth broker* adiciona à estimativa o valor de r.

O algoritmo *Hoeffding Bounds* (HB) combina as características de um algoritmo baseado em parâmetros e de um algoritmo baseado em medições, já que a estimativa de utilização é determinada a partir das medições efectuadas à rede e da largura de banda reservada pelos fluxos. Admite um novo fluxo se a soma da sua taxa de pico com a largura de banda efectiva do agregado de fluxos existentes da mesma classe, determinada a partir dos limites de *Hoeffding*, for inferior à utilização alvo:

$$C(\mu, \{P_i\}_{1 \le i \le n}, \varepsilon) + P < \alpha$$

onde

> P é a taxa de pico do novo fluxo e

> α é a utilização alvo da ligação.

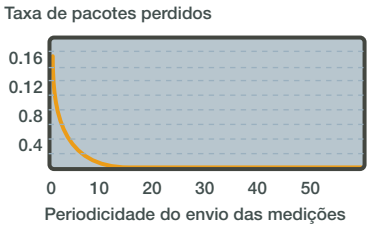

figura 7 – taxa de perdas para o algoritmo MS

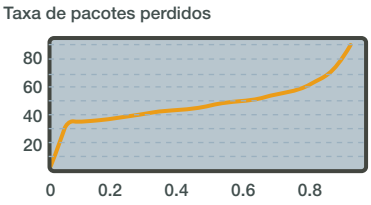

figura 9 – taxa de perdas para o algoritmo HB

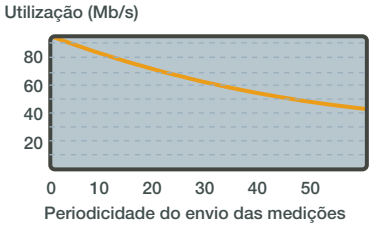

figura 8 – Nível de utilização para o algoritmo MS

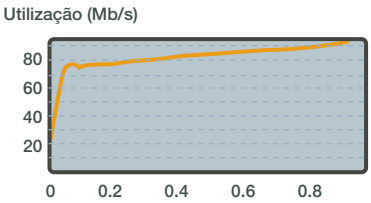

figura 10 – nível de utilização para o algoritmo HB

$C(\mu, \{P_i\}_{1\le i\le n}, \varepsilon)$ é a largura de banda efectiva do agregado de fluxos existente e é dada por

$$C(\mu, \{P_i\}_{1\le i\le n}, \varepsilon) = \mu + \sqrt{\frac{1n(1/\varepsilon)\sum_{i=1}^{n}(p_i)2}{2}}$$

onde

> μ é a medição da taxa de transmissão do agregado de fluxos;

> p_i é a taxa de pico do fluxo i;

> ε representa a probabilidade da taxa de transmissão do agregado de fluxos exceder a largura de banda efectiva $C(\mu,\{P_i\}_{1\le i\le n}, \varepsilon)$.

Os gráficos acima mostram a dificuldade em utilizar um algoritmo baseado em medições numa arquitectura onde o controlo de admissão é centralizado: o BB não consegue obter uma medida instantânea da utilização da rede. Pode verificar-se que a utilização da ligação decresce à medida que se diminui a frequência de envio das medições. Este resultado deve-se ao incremento efectuado à estimativa aquando da admissão de um novo fluxo, tornando o valor desta mais elevado que a utilização real. Como em nenhum momento se decresce a estimativa, esta mantém-se elevada até à recepção da próxima medição, ou seja, o BB, erroneamente, rejeita pedidos de admissão, sendo este algoritmo conservador.

Como foi referido, o algoritmo HB utiliza o parâmetro ε para definir a probabilidade da taxa de transmissão do agregado de fluxos ser superior à largura de banda efectiva determinada. Como tal, este parâmetro define o grau de agressividade do algoritmo, pois valores baixos de ε conduzem a um algoritmo mais conservativo e valores mais elevados a um algoritmo mais agressivo com a consequente maior utilização da ligação e maior taxa de perdas.

Conclusões

Com a observação destes resultados, verifica-se que o aumento da utilização da rede é conseguido a custo de perda de qualidade de serviço, pois, nas simulações em que o algoritmo foi mais agressivo, isto é, onde se verificaram elevadas taxa de admissão, mais elevado foi o valor da taxa de pacotes perdidos.

Algoritmo	Utilização média (Mb/s)
SS*	43,40
MS	85,34
HB	84,01

*O algoritmo SS obteve 0% de pacotes perdidos

tabela 2

Em termos de desempenho, verifica-se que este é idêntico para os dois algoritmos baseados em medições. Isto significa que os algoritmos deste tipo alcançam desempenhos semelhantes independentemente das fórmulas matemáticas que utilizam, sendo apenas necessário escolher os parâmetros apropriados para cada algoritmo. Se se fixar um valor da taxa de pacotes perdidos, por exemplo ~0.5%, seriam obtidos os resultados presentes na tabela 2.

A tabela 2 evidencia a vantagem em usar algoritmos baseados em medições em detrimento dos algoritmos baseados em parâmetros na presença de tráfego proveniente de fontes *ON-OFF*, já que os primeiros permitem uma melhor rentabilização dos recursos. O controlo de admissão efectuado por algoritmos que baseiam as suas decisões nas medições da rede obtiveram, neste cenário, uma utilização média que se aproxima do dobro do conseguido pelo algoritmo que se baseia apenas na informação dada pelas fontes.

Assim, a escolha do tipo de algoritmo a ser utilizado no controlo de admissão terá de se basear no tipo de serviço a que corresponde o novo fluxo. Como se verificou, os algoritmos baseados em medições oferecem uma elevada utilização e, portanto, existe a possibilidade deste tipo de algoritmos falhar os níveis de QoS requeridos. Assim, estes são apropriados para os serviços tolerantes. Por seu lado, a utilização de algoritmos baseados em parâmetros deve ficar restrita aos serviços que exigem garantias estritas ou que garantem que a taxa que é anunciada é a que realmente se vai verificar.

4.2 Algoritmos CAC previstos para um serviço VoD e sua aplicação

A introdução de mecanismos de CAC num serviço VoD, desenvolver-se-á

em duas componentes, com crescente aumento de complexidade e inteligência:

1. por contabilização dos fluxos VoD activos em pontos específicos da rede (DSLAM, elementos agregadores de nível e nível 3 e Servidores de vídeo);

2. adição da observação do estado da rede em todo o percurso entre servidores VoD e acesso do cliente (*router/modem ADSL*).

Na primeira fase, em que a inteligência da entidade de gestão de recursos é reduzida, o BB mantém em memória, associado a cada elemento da rede, um contador que identifica o número de sessões de vídeo activas em cada momento e que cruzam esse elemento. Para esta simplificação, considera-se que todos os filmes requerem o mesmo débito binário na rede para a sua visualização. Aquando da recepção de um pedido para uma nova sessão, o BB verifica, para todos os elementos pertencentes ao caminho entre o cliente e o Servidor de Vídeo, se estes suportam mais uma nova sessão, isto é, se o contador se encontra abaixo de um determinado limiar configurado. Se esta condição for verdadeira então a sessão de vídeo é aceite e é incrementado o contador para todos os elementos no caminho; em caso contrário, a sessão não deve ser activada e o cliente é informado que no momento não é possível a visualização do conteúdo pretendido.

Para a segunda fase, a inteligência, o volume de informação a processar e as tarefas do BB para correlacionar tudo isso, são maiores, uma vez que deixará de apenas contabilizar o número de fluxos activos em determinados pontos da rede, mas deverá conhecer também a utilização/estado da rede no caminho entre o cliente e o servidor de vídeo alvo. Para que tenha um conhecimento efectivo do estado da rede, o BB tem de interagir com uma plataforma de medições de rede, que poderá ser o Altaia [12], de forma a colectar a utilização efectiva de cada elemento da rede. É com base no conhecimento do nível de utilização dos elementos da rede e nas reservas autorizadas anteriormente que deverá decidir sobre a aceitação ou rejeição de novos pedidos.

Tomando P_c como o pico do novo fluxo pertencente à classe de serviço c, e μ_c como a última medição recebida relativa a essa mesma classe para cada elemento do percurso do fluxo, então o *Broker* aceitará uma nova sessão se:

$$P_c + \mu_c < \alpha_c$$

sendo α_c a capacidade máxima para a classe c no elemento de rede x.

De forma a minimizar a dependência do algoritmo da frequência das medidas, propõe-se deduzir o valor de utilização da ligação a partir do rácio entre a utilização medida e a largura de banda reservada.

Cada medição recebida pelo BB corresponde à média da utilização relativa ao último intervalo. Com o cálculo da média da largura de banda reservada para o mesmo intervalo, pode ser estabelecido o rácio referido para derivar a estimativa de utilização cada vez que um novo pedido é recebido. Desta forma, a estimativa de utilização é actualizada ao longo do tempo e não apenas nos instantes em que novas medições são recebidas.

Uma vez que uma medição recebida pelo BB corresponde à média da utilização relativa ao último intervalo, com estes longos, a medição recebida pode estar longe da utilização actual. Assim, as medições recebidas são apenas utilizadas para estabelecer o rácio, sendo que a estimativa de utilização utilizada na expressão é obtida com base neste rácio e na largura de banda reservada no momento da recepção de uma nova medição.

A nova expressão é dada por

$$\mu_{ac} + \mu_{pc} + P_{pc} < \alpha_c$$

em que:

> $\mu_{ac} = R_{ac} \cdot \mu_{mc} / R_{mc}$: estimativa da utilização no momento em que a última medição foi recebida;

> $\mu_{pc} = R_{uc} \cdot \mu_{mc} / R_{mc}$: previsão da largura de banda consumida pelos fluxos que entraram desde o momento em que a última medição foi recebida;

> $P_{pc} = P_c \cdot \mu_{mc} / R_{mc}$: previsão da largura de banda que o novo fluxo irá consumir;

> μ_{mc}: última medição recebida relativa à classe c. Corresponde à média de utilização do último intervalo;

> R_{mc}: largura de banda média reservada relativa ao último intervalo;

> R_{ac}: largura de banda reservada pelos fluxos no momento em que a última medição foi recebida;

> R_{uc}: largura de banda reservada pelos fluxos que entraram desde que a última medição foi recebida;

> Por cada fluxo admitido: $R_{uc} = R_{uc} + P_{pc}$;

> Por cada fluxo terminado: $R_{uc} = R_{uc} - P_{pc}$;

> Ao receber uma nova medição: $R_{uc} = 0$;

> P_c pico do novo fluxo pertencente à classe de serviço c.

> α_c: Utilização alvo para a classe c.

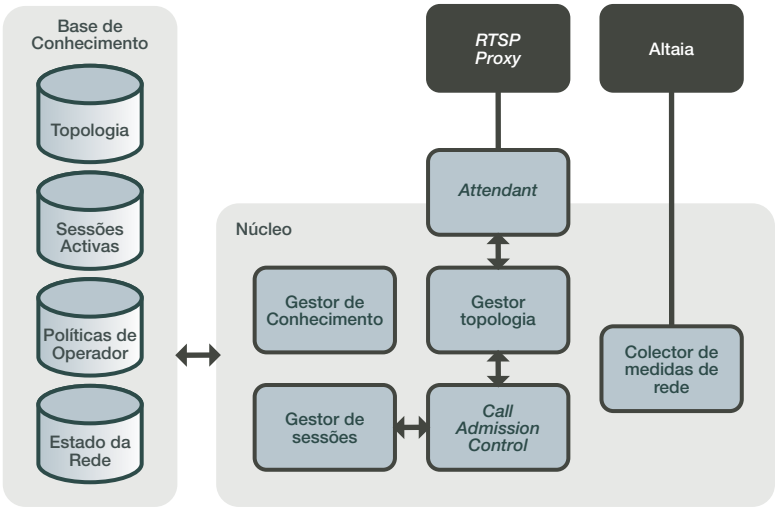

figura 11 – Estrutura interna do BB

A utilização alvo α_c pode constituir um parâmetro ajustável que será utilizado para alterar o grau de agressividade do algoritmo. Para tal, o BB terá de receber estatísticas relacionadas com QoS.

O elemento central para a tomada de decisões (execução dos algoritmos de CAC atrás referidos), tem a estrutura interna ilustrada na figura 11.

É composto por vários blocos funcionais, sendo os mais importantes:

> ***Attendant* –** responsável por processar os pedidos e responder, consoante a decisão, ao *RTSP Proxy*;

> **Gestor de Topologia –** este módulo mantém em memória a topologia actual da rede e executa um algoritmo de forma a encontrar o caminho entre o cliente e o *Video Server*;

> ***Call Admisson Control* –** módulo que executa o algoritmo de CAC. É o centro de todo o processo de decisão;

> **Gestor de Conhecimento –** interface entre o núcleo do *Broker* e as bases de dados que servem como base de conhecimento para o *Broker*;

> **Colector de medidas de rede –** responsável por receber os valores enviados pela plataforma de medidas e armazená-las na base de conhecimentos do *Broker*.

5. Conclusões

A gestão de recursos e o controlo de admissão em redes é um tema que não está livre de polémica sobre a sua real utilidade e aplicabilidade. O aumento da banda disponibilizada pelas várias tecnologias parece indicar que a crescente procura poderá ser sempre respondida com um aumento da capacidade no transporte, a custos baixos. No entanto, no segmento de acesso, em especial quando este se faz com recurso a tecnologias rádio móveis, isso não é verdade. Também recentes estudos [13] vêm provar que mesmo para a rede de transporte existirão limitações no futuro, atendendo à crescente disponibilização de banda ao sector residencial. À data de conclusão deste texto, não existem ainda resultados da exploração prática da solução aqui descrita. No entanto existem em laboratório protótipos dos mesmos, o que nos permite concluir sobre a aplicabilidade e benefícios dos mecanismos e algoritmos aqui descritos, nomeadamente sobre a intercepção e análise de mensagens RTSP. Alguns dos componentes também se encontram integrados em protótipos de projectos internacionais atestando a sua utilidade em redes NGN e B3G do futuro [14] [15].

Como trabalho futuro, pretende-se colocar a solução em operação em ambiente real, o que permitirá parametrizar correctamente os algoritmos definidos e atestar sobre a escalabilidade das soluções encontradas. O âmbito de aplicação poderá ser alargado a outros serviços, permitindo ter um controlo total da globalidade dos serviços e fluxos que cruzam a rede.

Referências
[1] http://www.ipv6.org/
[2] http://www.itu.int/ITU-T/ngn/definition.html
[3] http://www.etsi.org
[4] http://portal.etsi.org/tispan
[5] http://en.wikipedia.org/wiki/*Ethernet*
[6] http://www.mplsrc.com/
[7] http://www.rtsp.org/
[8] http://www.ieee802.org/1/
[9] http://www.ietf.org/html.charters/OLD/diffserv-charter.html
[10] http://www.faqs.org/rfcs/rfc3046.html
[11] http://www.omnetpp.org/
[12]http://www.ptinovacao.pt/produtos/aplicacoes/fol_Altaia.pdf
[13] K. Cho, K. Fukuda, H. Esaki, A. Kato, "The Impact and Implications of the Growth in Residential User-to-User Traffic," ACM Sigcomm 2006, Pisa, Italia, Setembro de 2006
[14] Ricardo Azevedo, André Oliveira, Francisco Fontes, "Distributed, Scalable QoS Control Architecture for Next Generation Networks", 15th IST Mobile & Wireless Communications Summit, Mykonos, Greece, June 2006
[15] http://www.ist-daidalos.org/

Francisco Fontes, licenciado em Engenharia Electrotécnica e de Computadores, Ramo de Electrónica e Telecomunicações, pelo Instituto Superior Técnico (Set/91) e Doutorado pela Universidade Politécnica de Madrid (Nov/00) na área de gestão distribuída de redes de telecomunicações. Em Set/91 iniciou a sua actividade profissional na PT Inovação (CET) tendo-se especializado em tecnologias de rede de banda larga, especialmente nas tecnologias ATM e IP. Actualmente os seus interesses situam-se nas áreas de Qualidade de Serviço em redes multi-serviço e na sua evolução para arquitecturas RPG All-IP.

Ricardo Azevedo, mestrado em *Internet Computing* pelo Queen Mary College - Universidade de Londres em 2006 e Licenciado em Eng. de Computadores e Telemática pela Universidade de Aveiro em 2004. Foi bolseiro de investigação do IT desde 2004 a Fev. de 2006, com trabalho desenvolvido na área de QoS em redes heterogéneas de 4ª geração. Desde Fev. 2006 é contratado a termo pela PT Inovação, tendo por objectivo o desenvolvimento uma entidade gestora de recursos a integrar na plataforma ShipNet. Participação em diversos projectos de I&D no âmbito do IST na área de QoS e *Network Management*. Conhecimento de diversas linguagens de programação e interesse pelas áreas de QoS, *Admission Control Algorithms*, *Network Measurements* e mobilidade em cenários de redes heterogéneas.

André Oliveira, licenciado em Engenharia Electrónica e Telecomunicações pela Universidade de Aveiro em 2004, encontra-se presentemente a realizar dissertação de mestrado. É bolseiro de investigação no Instituto de Telecomunicações desde 2004, desenvolvendo trabalho de I&D na temática das redes de telecomunicações, especificamente nas áreas de Qualidade de Serviço (QoS) em redes heterogéneas de próxima geração e redes de comunicações multimédia.

Gonçalo de Lemos, licenciado em engenharia electrónica e de telecomunicações pela Universidade de Aveiro, encontra-se actualmente a realizar uma dissertação de mestrado, enquanto investigador do Instituto de Telecomunicações de Aveiro, nas áreas da qualidade de serviço (QoS) e comunicações multimédia.

Jorge Loura, licenciado em Engenharia de Computadores e Telemática pela Universidade de Aveiro em 2006. É bolseiro de investigação do IT desde 2006, participando no desenvolvimento de projectos na área de QoS.

João Ferreira, licenciado em Engenharia informática e Computadores no ramo de Sistemas Computacionais pelo Instituto Superior Técnico. Em 1993 começou a trabalhar na PT Comunicações na direcção de Engenharia na área das redes de dados. Esteve envolvido desde então na implementação e desenvolvimento de serviços na rede ADSL, *Backbone* IP, Rede *Ethernet*, MPLS, WIFI e pilotos de Vídeo sobre IP. É actualmente responsável pelo departamento de Desenvolvimento de Serviços e Pilotos da PT Comunicações.

Pedro Bandeira, licenciado em Engª Electrotécnica e Computadores, Ramo de telecomunicações pelo IST. Trabalhos em *Watermarking* e processamento de Vídeo Digital. Colaborou com a Telepac na área das redes de dados, VOIP e videoconferência sobre IP. Colaborou com a PT Prime na áreas das redes de dados empresariais (VPNs, IPSEC, VOIP e multiconferência sobre IP).

Desde 2002 é colaborador da PT Comunicações, trabalhando na área do IPTV, VoD e segurança de conteúdos em redes IP.

04

Cenários de IP Multicast Seguro


Bernardo Cardoso, João Matos

palavras-chave:
Multicast, IPTV, Segurança, Criptografia

A evolução tecnológica no mercado das telecomunicações aponta no sentido de que a voz, os dados e mesmo a difusão de vídeo, sejam transmitidos na rede sob a forma de pacotes IP. Neste momento está a iniciar-se a comercialização de serviços de *triple play* em que se incluem, num mesmo par de cobre, todos estes serviços, estando a ser dada uma especial atenção à área do IPTV e do VOD. Embora seja possível tirar partido de tecnologias de segurança semelhantes às utilizadas nas actuais distribuições de televisão digital via satélite e via cabo, nos cenários de IPTV é possível tirar partido da natureza bidireccional da rede e criar soluções de IP *Multicast* Seguro sem recurso a cartões, que apresentam características de segurança vantajosas em relação a estes e têm potencial para uma substancial redução de custos ao nível do investimento e da operação.

1. Introdução

A revolução tecnológica no mercado das telecomunicações aponta no sentido de que múltiplos serviços, tais como, a voz, os dados e mesmo a difusão de vídeo, sejam transmitidos na rede sob a forma de pacotes IP. Esta tendência de oferta de serviços integrados sobre redes IP pode ser explicada pela possibilidade de partilha dos recursos de transmissão e comutação, pela adopção de técnicas unificadas de gestão e pela versatilidade introduzida na operação destes serviços, o que favorece a emergência de novos mercados.

Actualmente esta revolução caracteriza-se pelo início da comercialização de serviços de *triple play* em que se incluem a voz, os dados e o vídeo. Esta terceira componente pode ainda ser dividida em duas partes: a que diz respeito à distribuição de vídeo em *broadcast*, à qual se associa normalmente o termo IPTV e à distribuição de vídeo *unicast* denominada por VOD.

A componente IPTV tem modelos de negócio semelhantes à televisão paga, actualmente fornecida pelos operadores de televisão por cabo. Partilha também de problemas semelhantes, nomeadamente a questão do acesso ilegal aos canais. Para a solução destes problemas podem ser empregadas metodologias idênticas às usadas pela televisão digital. Estas recorrem a cartões de segurança e à cifra dos conteúdos. No entanto, nos cenários de IPTV é possível tirar partido da natureza bidireccional da rede e criar soluções *cardless* que apresentam características de segurança com vantagens evidentes em relação aos cartões, bem como o potencial de redução de custos associados. Este artigo vai focar-se numa solução deste último grupo, denominada *IP Multicast* Seguro.

2. Cenários IPTV, *Multicast* e Segurança

A tecnologia IPTV permite a distribuição de conteúdo televisivo usando o protocolo IP. A transmissão é assegurada através de técnicas de difusão *IP Multicast*. Neste tipo de cenários associa-se um canal de televisão a um grupo *multicast*. Quando seleccionam um determinado canal na sua STB, os subscritores estão a realizar, em termos IP, uma associação a um grupo *multicast*.

Os modelos de negócio associados ao IPTV compreendem vários cenários de distribuição, que correspondem aos actuais produtos de televisão paga. Os mais normais e que estão hoje disponíveis nos operadores de televisão por cabo são a tarifa plana (*flat rate*) associada a um determinado pacote de canais e os canais *premium*, que são pagos à parte, um a um.

Na televisão por cabo analógica, as soluções técnicas para a implementação destes produtos recorrem a filtros de frequências para diferenciar os vários pacotes de canais e ao uso de descodificadores para os canais *premium*. Em termos de televisão digital a solução normalizada passa pela cifra dos conteúdos e pelo uso de STBs com sistemas de acesso condicional e respectivos cartões. Estes sistemas, embora

bastante seguros e flexíveis, têm o problema de, no caso de ser comprometida a sua segurança, obrigarem à troca física do cartão e eventualmente do próprio receptor.

Nos sistemas de *triple play* sobre tecnologia xDSL, como existe uma ligação ponto a ponto entre a central e o cliente, é possível usar soluções em que o controlo do acesso aos canais é feito no DSLAM, com matrizes de acesso entre portas xDSL e grupos de *multicast*. Este tipo de solução, embora viável, padece de alguns problemas. Primeiro é preciso que o DSLAM suporte esta facilidade, depois existem vários problemas de gestão e manutenção dos equipamentos que se complicam bastante quando se utilizam dispositivos de fabricantes diferentes. Colocam-se também questões associadas ao acesso aos canais *premium*, dado que a sua difusão em aberto e em pacotes IP permite com relativa facilidade a sua redifusão, quer para clientes autorizados (outros receptores dentro da casa), quer para terceiros não autorizados.

Para além destes cenários simples, os produtos mais elaborados, como o pagamento por visualização (*Pay Per View* – PPV) muito associado à transmissão de eventos desportivos e filmes e o vídeo quase a pedido (*Near Vídeo On Demand* – NVOD), apresentam complicações acrescidas à metodologia de controlo directo no DSLAM. Estes produtos obrigariam a um acesso constante à configuração do equipamento e continuariam a não estar protegidos contra a redifusão.

Para resolverem estes problemas, alguns operadores de IPTV estão a optar por soluções semelhantes às usadas na televisão digital, isto é, utilização de sistemas de acesso condicional com recurso a cartões. Aliás, os sistemas são em muitas situações exactamente os mesmos e, portanto, partilham dos mesmos problemas aquando do comprometimento da sua segurança. No entanto é possível, tirando partido da bidireccionalidade da rede, desenvolver sistemas simultaneamente seguros e que não precisam de utilizar cartões (*cardless*), podendo reagir muito rapidamente a uma situação de comprometimento de segurança. Em IPTV estas soluções denominam-se por *Multicast* Seguro.

3. *Multicast* Seguro

A abordagem *multicast* seguro [1], aplicada ao IPTV, adiciona à tecnologia *multicast* tradicional a cifra dos conteúdos, permitindo que apenas os receptores devidamente autenticados e autorizados tenham acesso a determinados canais de vídeo.

A principal vantagem desta solução, em relação aos esquemas de acesso condicional, tipicamente usados nos sistemas de televisão digital, é que num cenário de IPTV, devido à bidireccionalidade disponível na rede, o elemento responsável pela gestão das chaves e dos acessos não se encontra num cartão que está em casa do cliente. Neste cenário, esse elemento encontra-se centralizado nas instalações do operador que, a pedido, vai fornecendo as chaves cliente a cliente, canal a canal.

Ao estarem cifrados os dados e ao ser necessária a autenticação e a manutenção de uma sessão com um servidor, para receber as chaves indispensáveis à decifra, reduzem-se substancialmente os problemas associados ao acesso fraudulento aos conteúdos. Limita-se também a redifusão para terceiros não autorizados. Esta necessidade de comunicar com um servidor garante também a eliminação da duplicação de receptores, a chamada "clonagem", pois é muito fácil ao servidor detectar tentativas de estabelecimento de sessões duplicadas.

Ao residir no servidor e não no cliente, como nos sistemas de cartões, o controlo do acesso de canais permite que não sejam possíveis manipulações dos cartões/ receptores para haver acesso a canais não incluídos no pacote subscrito. Uma vez que grande parte da inteligência associada ao controlo reside no servidor central é possível realizar alterações profundas aos sistemas sem necessidade de trocas de cartões, com a redução dos respectivos custos associados, quer em termos monetários, quer em termos de imagem.

Por fim, esta tecnologia tem a vantagem de, em caso de detecção de tentativa de fraude, poder identificar rapidamente o cliente associado, já que é possível estabelecer uma relação física entre o mesmo e a porta xDSL por onde estão a fluir os dados.

4. Implementação do *IP Multicast* Seguro

4.1 Arquitectura

A solução proposta apresenta os seguintes elementos: servidor de chaves, STBs e *streamers* de IPTV. O servidor de chaves é responsável pelo processo de geração e distribuição de chaves, pela autenticação e pela autorização de clientes. É da competência do servidor de chaves a geração dos contextos criptográficos usados nos pedidos de canal (STEK/STAK) e no canal de distribuição de chaves (KEK/KAK). As STBs e os *streamers* de IPTV actuam essencialmente ao nível da transformação de dados, i.e. na cifra e decifra dos canais de vídeo. Os *streamers* de IPTV são responsáveis pela transformação dos conteúdos áudio e vídeo num fluxo de pacotes IP e no envio cifrado desse mesmo fluxo. Têm ainda como função a geração das chaves de cifra a utilizar no canal de vídeo (VEK) e a sua distribuição. As STBs são responsáveis por obter todos os elementos criptográficos necessários (STEK/STAK, KEK/KAK e VEK) para a decifra do fluxo de vídeo.

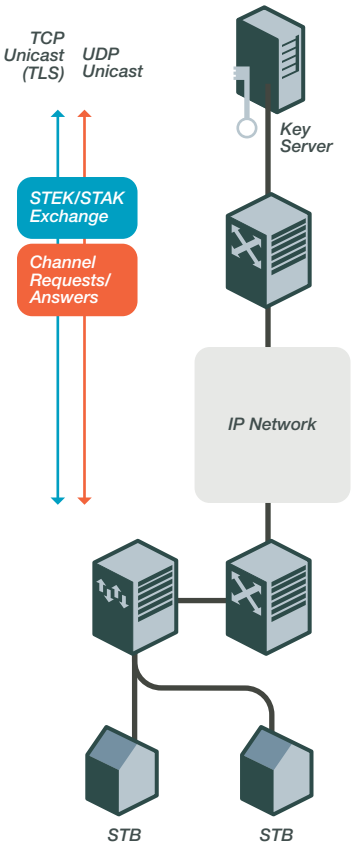

figura 1 – Ligações usadas entre STBs e o *Key Server*

4.2 Funcionamento básico

Na solução proposta são adoptadas três tipos de ligações: ligações TCP *unicast*, UDP *unicast* e ligações UDP *multicast*. A ligação UDP *unicast* é utilizada durante os pedidos de canal efectuados pelas STBs ao servidor de chaves, no refrescamento do contexto criptográfico do canal de distribuição de chaves (KEK/KAK) e na sinalização dos períodos de anúncio e refrescamento das chaves de cifra de vídeo (VEK) entre o servidor de chaves e os *streamers* de IPTV (figura 1). As ligações *multicast* são usadas para a transmissão do fluxo de vídeo cifrado e dos fluxos de distribuição de chaves de cifra dos canais de vídeo.

O processo de difusão de múltiplos canais de vídeo usando *IP multicast* seguro inicia-se quando um cliente liga a respectiva STB. Nesse momento é estabelecida uma ligação TCP segura entre a STB e o servidor de chaves, recorrendo ao protocolo TLS [2]. A STB solicita o contexto criptográfico de sessão (STEK/STAK) ao servidor de chaves através dessa ligação segura, que lho entrega após a verificação da autenticidade da STB. Na posse da STEK/STAK, a STB pode dar início ao envio de pedidos de canal ao servidor de chaves, com o intuito de receber o contexto criptográfico do canal de distribuição de chaves (KEK/KAK) associado ao canal subscrito pelo cliente (figura 2). Neste pedido de canal, a mensagem vai cifrada/autenticada com as chaves STEK/STAK e encapsulada em pacotes UDP *unicast*.

Em simultâneo com o envio de pedido de canal ao servidor de chaves, a STB envia mensagens *Internet Group Management Protocol – version 2* (IGMPv2) [3] ao *Access Router* (AR) para se associar ao grupo *multicast* do canal, dando assim início à recepção de tráfego *multicast*. As mensagens IGMPv2 são:

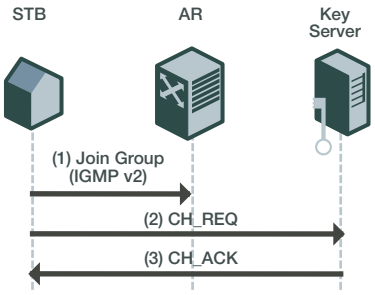

figura 2 – Pedido de canal

> Mensagem “JOIN”, para passar a receber os pacotes *multicast* do canal desejado;

> Mensagem “LEAVE”, para deixar de receber os pacotes do canal anterior.

O canal de vídeo solicitado pelo utilizador é transmitido para o endereço *multicast* específico do canal, sendo transportado sob a forma de pacotes *Secure Real-time Transport Protocol* (SRTP) [4] cifrados com a chave de cifra do canal (VEK). Esta chave é distribuída periodicamente em *multicast* conjuntamente com o próprio canal de vídeo, i.e., enviada para o mesmo endereço *multicast* do canal mas para um porto diferente, cifrada/autenticada com a KEK/KAK respectiva. As STB decifram a VEK do canal usando a KEK obtida no pedido de canal. Após obtenção da VEK as STB decifram os pacotes SRTP associados ao fluxo de vídeo (figura 3).

Todos os contextos criptográficos possuem um tempo de vida máximo que é dependente dos seguintes factores:

> do algoritmo de cifra utilizado;

> da sua resistência a ataques de força bruta;

> do tamanho de eventuais janelas *anti-replay.*

Assim, o refrescamento de contextos criptográficos surge como

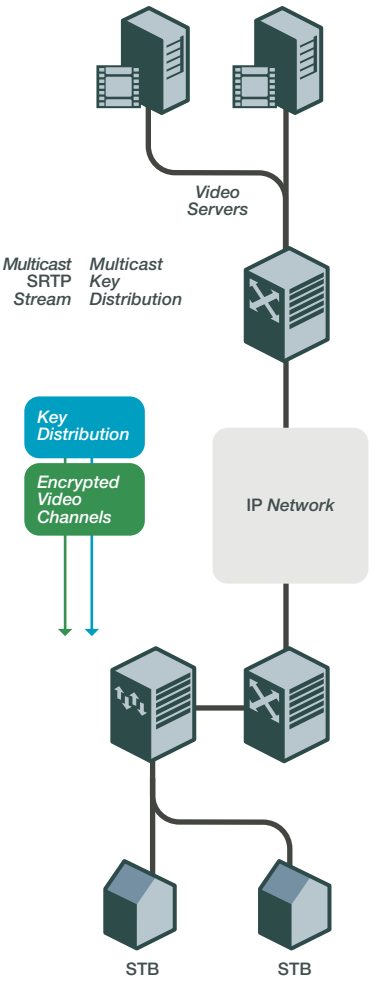

figura 3 – Difusão de vídeo e chaves

obrigatório e deverá acontecer num intervalo de tempo inferior ao seu tempo de vida máximo.

No caso do refrescamento da VEK, e de forma a garantir a continuidade da visualização do canal solicitado, existe um mecanismo de sincronização entre os *streamers* de IPTV e as várias STB envolvidas, relativamente à nova VEK (distribuída no canal *multicast*). Este sincronismo é feito com base no número de sequência dos pacotes SRTP. No que respeita ao contexto criptográfico de sessão (STEK/STAK) e ao contexto criptográfico do canal de distribuição de chaves (KEK/KAK), é da competência das STB a sua renovação após estes terem expirado.

5. Conclusão

A implementação dos serviços de *triple play* está a intensificar a procura de soluções que permitam a disponibilização de serviços com segurança e produtos flexíveis, nomeadamente em termos de televisão paga sobre tecnologia IPTV.

As soluções baseadas em *multicast,* e principalmente as abordagens em volta do *multicast* seguro, apresentam-se como um método bastante eficaz e acessível para a disponibilização de serviços de IPTV.

A implementação de *IP Multicast* Seguro, aqui proposta, apresenta características bastante interessantes em termos da capacidade de disponibilização de serviços sem comprometer a segurança da solução sendo mais acessível e com melhores funcionalidades que um tradicional sistema baseado em cartões.

Referências
[1] Thomas Hardjono, Brian Weis, "The Multicast Group Security Architecture," in draft-ietf-msec-arch-05.txt from MSEC, 2004, Work in Progress
[2] T. Dierks and C. Allen, "The TLS Protocol Version 1.0," in RFC 2246.
[3] W. Fenner, "Internet Group Management Protocol, Version 2," in RFC 2236, 1997.
[4] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)," in RFC 3711, 2004.

Bernardo Cardoso, licenciado em Auditoria Contabilística e Mestre em Gestão da Informação, está integrado no grupo de Tecnologias Multimédia, onde tem estado envolvido em vários projectos relacionados com televisão e vídeo digital, com particular enfoque em sistemas baseados em MPEG e DVB, nomeadamente sistemas de transmissão de televisão digital e interactividade. Esteve fortemente envolvido no projecto de Televisão Digital Interactiva da TV Cabo, ao nível do desenvolvimento de aplicações, conteúdos, testes e análise de desempenho. Mais recentemente dedica-se ao desenho, implementação e avaliação de sistemas de IPTV, especialmente sobre soluções xDSL.

João Carlos Matos, licenciado em Engenharia Informática pela Escola Superior de Tecnologia e Gestão de Felgueiras. No INESC Porto foi responsável pelo desenvolvimento de um Protótipo de IP *Multicast* Seguro para transmissão de vídeo. Actualmente é estagiário na PT Inovação e a sua actividade centra-se essencialmente em cenários de IPTV.

Nota
Este projecto foi desenvolvido em conjunto com o INESC Porto ao abrigo do Plano de Inovação. Da parte do INESC estiveram envolvidos os Professores Manuel Ricardo e António Pinto.

01

Programação Orientada a Aspectos em Telecomunicações

Cristiano Malanga Breuel (INB)

palavras-chave:
Programação Orientada a Aspectos, Engenharia de *Software*, Requisitos Transversais, Modularidade, Linguagens de Programação, Sistemas de Telecomunicações.

Em sistemas de telecomunicações, é comum que existam requisitos transversais a todas as suas funcionalidades, como gestão de fusos horários, segurança, multi-operadoras e multi-tecnologias. A Programação Orientada a Aspectos surgiu para permitir uma melhor modularização desse tipo de requisito. Neste artigo, oferecemos uma breve introdução sobre este novo paradigma de programação, e fazemos uma análise das perspectivas de sua utilização no âmbito de sistemas de telecomunicações.

1. Introdução

Desde o surgimento dos computadores, a ciência da computação busca formas de tornar a sua programação mais fácil e eficiente. Para atingir esses objetivos, é preciso sempre levar em conta as habilidades e defeitos das duas partes envolvidas: humanos e computadores. Os primeiros têm uma capacidade inigualável de reconhecer padrões e criar abstrações. Entretanto, quando se trata de tarefas repetitivas, são muito lentos e propensos a erros. Já os computadores são extremamente rápidos e imunes a erros em tarefas repetitivas, mas não possuem as capacidades cognitivas dos humanos.

Portanto, a otimização do processo de programação tem de seguir sempre o mesmo sentido: permitir aos humanos expressarem-se através de abstrações, símbolos e descrição de padrões, deixando para o computador o máximo possível de tarefas repetitivas.

A programação orientada a aspectos é um grande passo nesse sentido. Antes dela, muitos requisitos transversais não podiam ser expressos de forma modularizada, manifestando-se na forma de repetição de código espalhado pelo programa.

2. Requisitos transversais

As boas práticas de engenharia de *software* recomendam que se criem sistemas com modularidade. Ou seja, funcionalidades relacionadas devem ser agrupadas em módulos, de forma que se possa compreendê-las e mantê-las de forma isolada. Entretanto, nas linguagens de programação tradicionais (orientadas a objetos e procedurais), isso nem sempre é possível. Existe uma classe de requisitos que, por natureza, não podem ser isolados em módulos, porque interagem com muitas outras funcionalidades. Na programação orientada a aspectos, eles são chamados de requisitos transversais ou interesses transversais ("*crosscutting concerns*").

Alguns dos exemplos mais comuns desses requisitos são segurança, *logging*, tratamento de erros, atualizações de estado. Entretanto, em muitas aplicações e domínios, existem interesses transversais específicos. Na área das telecomunicações, esses exemplos são abundantes.

3. Programação orientada a aspectos

A proposta da programação orientada a aspectos é a de modularizar em uma única unidade de *software* (denominada um aspecto) os requisitos transversais, cuja implementação anteriormente ficava espalhada por diversas unidades.

A forma como isso é feito é através da inserção automática de código para tratar esses requisitos nos pontos adequados. Assim, tudo o que o programador de um aspecto precisa fazer é especificar onde essa inserção deve ser feita e o que deve ser feito nesses pontos. As especificações de onde e o que são traduzidas respectivamente em dois elementos dos aspectos: os *pointcuts* e os *advices*.

Os *pointcuts* são expressões que selecionam um conjunto de *joinpoints* em um programa. Um *joinpoint* é

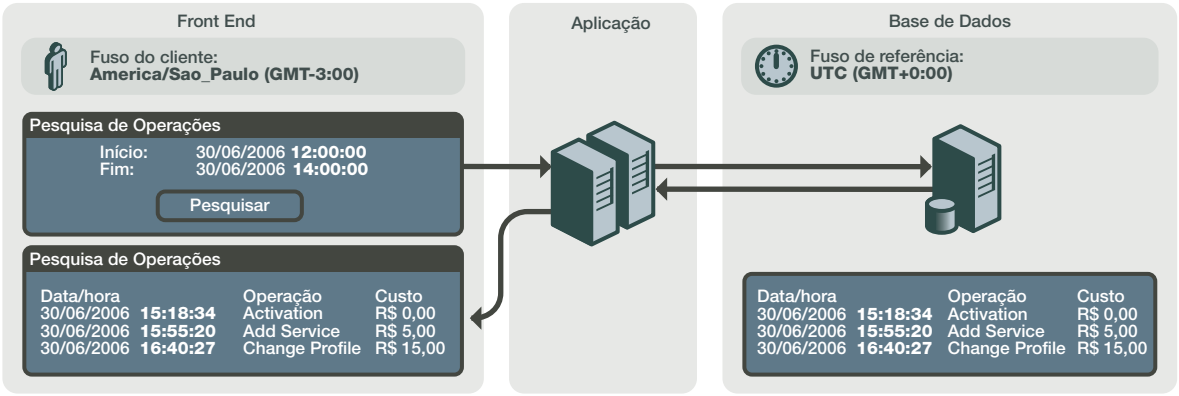

figura 1 – Exemplo de processo com gestão de fusos horários

qualquer ponto de interesse no programa, como a definição de um método ou um ponto de tratamento de exceção. Um *pointcut* é análogo ao que representa uma *query* SQL para um banco de dados relacional, mas em vez de selecionar linhas em tabelas ele seleciona *joinpoints* de um programa.

Um *advice* é um trecho de código que é associado a um *pointcut* para implementar o requisito transversal. Ele será inserido nos *joinpoints* selecionados pelo pointcut associado a ele. Normalmente, o *advice* é implementado na mesma linguagem do resto do programa, como um método comum.

No exemplo que é dado mais à frente, estes conceitos ficarão mais claros.

4. Ferramentas disponíveis

Existem diversas linguagens e *frameworks* orientados a aspectos disponíveis no mercado, quase todos de código livre. Todas as principais linguagens modernas, como Java, C#, C++ e C têm pelo menos uma ferramenta disponível para programação orientada a aspectos.

A linguagem mais adiantada nesse sentido é Java. A primeira linguagem orientada a aspectos de sucesso, AspectJ [1], é baseada em Java, e diversas outras semelhantes surgiram em seguida. Quase todas seguem o mesmo paradigma da pioneira, baseado nos conceitos de *pointcuts* e *advices*.

Existem também alguns *frameworks* que implementam os conceitos de orientação a aspectos sem que seja necessária uma extensão da linguagem ou um compilador específico. Eles se valem de descritores em XML ou anotações no código para definir os elementos que não existem nas linguagens orientadas a objetos.

O artigo [2] faz uma boa comparação entre as principais linguagens e *frameworks* orientados a aspectos do mercado. Nos exemplos deste artigo, utilizaremos a linguagem AspectJ, por ser a mais conhecida e madura e também por ter a sintaxe mais clara e compreensível.

5. Aspectos em telecomunicações

Em sistemas de grande porte em telecomunicações, como os de tarifação e gerência de linhas de telefonia móvel, a implementação e manutenção de requisitos transversais é particularmente custosa. Além disso, este tipo de requisito surge com grande freqüência.

Portanto, a redução do custo e esforço relativos a esses requisitos pode ser um fator de grande melhoria competitiva para um sistema. Essa melhoria não vem apenas através da redução de prazos e custos, mas também pela melhoria de qualidade, propiciada pela eliminação da necessidade de replicar código relativo aos requisitos transversais por todo o sistema, que sempre aumenta a taxa de erro humano.

A seguir, daremos um exemplo da implementação de um desses requisitos, com o uso de orientação a aspectos, para ilustrar o seu potencial.

6. Exemplo: Fusos Horários

Em um sistema de telecomunicações para um país de grande dimensão territorial, como o Brasil, surge a questão de lidar com diversos fusos horários. Este é um requisito transversal do sistema, que pode ter impacto em quase todas as suas funcionalidades.

Para exemplificar as vantagens da programação orientada a aspectos, vamos implementar a funcionalidade de gerência de fusos horários, em uma versão simplificada de um sistema de gestão de clientes de uma operadora móvel. Os requisitos desse sistema são:

> Cada cliente tem um fuso horário associado;

> Todas as informações na base de dados são armazenadas em um fuso horário de referência (UTC);

> As operações e consultas sobre um cliente devem sempre receber e exibir as datas/horas no fuso do cliente;

> Os relatórios que mostram dados de diversos clientes devem exibir as datas/horas no fuso de referência.

A figura 1 ilustra o processo.

Os requisitos acima têm conseqüências transversais para a aplicação de atendimento do sistema. Eles significam que:

> Todas as datas/horas fornecidas pelos usuários nas telas em consultas e operações sobre um

cliente devem ser convertidas para o fuso de referência antes de serem passadas para o banco de dados;

> Todas as datas/horas lidas do banco de dados em consultas e operações sobre um cliente devem ser convertidas para o fuso horário do cliente antes de serem exibidas.

Em um sistema convencional, sem o uso de programação orientada a aspectos, isso exigiria a inserção manual de código para fazer as conversões necessárias nos pontos relevantes de todo o sistema.

O exemplo consiste em uma classe que implementa a busca de operações de um cliente na base de dados, classes de domínio para representar um cliente e uma operação, e um aspecto de tratamento de fuso horário em operações com clientes. Existe também uma classe utilitária para conexão à base de dados e uma classe de negócio que invoca aquela responsável pela pesquisa dos dados. A estrutura do projeto é ilustrada na figura 2.

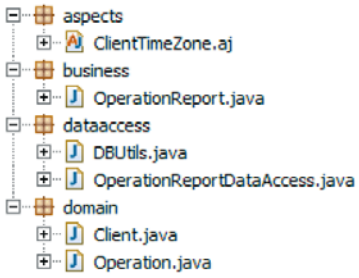

figura 2 – Estrutura de classes e aspectos do exemplo

A classe que implementa a busca das operações do usuário está mostrada na figura 3. O que ela faz é conectar-se à base de dados, executar a *query* que busca os dados e construir um objeto representando a operação, para cada uma das linhas. Como se vê, não, há nela qualquer código para tratamento de fusos horários.

A seguir, temos o aspecto que trata os fusos horários. A sintaxe de definição dos *pointcuts* e *advices* não óbvia a princípio. Para evitar sobrecarregar este artigo com detalhes sobre a linguagem, recomendamos a quem quiser se aprofundar que consulte a documentação do AspectJ, que encontra-se na referência [1], para entender os detalhes do funcionamento.

```
public class OperationReportDataAccess {

    public List<Operation> getClientOperations(Client client,
            Timestamp initialTime, Timestamp finalTime) throws SQLException {

        List<Operation> result = new ArrayList<Operation>();

        Connection conn = DBUtils.getConnection();

        PreparedStatement ps = conn.prepareStatement(
                "SELECT * " +
                "FROM OPERATIONS " +
                "WHERE CLIENT_ID=? " +
                "AND OPERATION_TIME BETWEEN ? AND ?");

        ps.setLong(1, client.getId());
        ps.setTimestamp(2, initialTime);
        ps.setTimestamp(3, finalTime);

        ResultSet rs = ps.executeQuery();

        while (rs.next()) {
            Operation operation = new Operation();
            operation.setTime(rs.getTimestamp("OPERATION_TIME"));
            // ...
            result.add(operation);
        }

        rs.close();
        ps.close();
        conn.close();

        return result;

    }

}
```

figura 3 – Classe de negócio da consulta de operações

```
public aspect ClientTimeZone {

    // Contante que define o fuso horário de referência
    private static final TimeZone DEFAULT_TIMEZONE =
        TimeZone.getTimeZone("GMT");

    // Pointcut que seleciona os métodos de acesso a dados de clientes
    pointcut clientDataAccess(Client client):
        args(client, ..) &&
        execution(public * dataaccess.*.*(Client,..));

    // Pointcut que seleciona as atribuições de variáveis do tipo Timestamp
    // em operações na base de dados, dentro de métodos de acesso a dados de
    // clientes
    pointcut setTimestamp(Client client, Timestamp t):
        cflow(clientDataAccess(client)) &&
        args(.., t) &&
        call(* java.sql.Statement+.setTimestamp(.., Timestamp));

    // Advice que trata conversões de fuso horário nos pontos selecionados pelo
    // pointcut setTimestamp
    void around(Client client, Timestamp t): setTimestamp(client, t) {

        TimeZone tz = client.getTimeZone();
        Timestamp t2 = convertTimeZone(t, tz, DEFAULT_TIMEZONE);

        proceed(client, t2);
    }

    // Pointcut que seleciona as obtenções de campos do tipo Timestamp em
    // consultas à base de dados, dentro de métodos de acesso a dados de
    // clientes
    pointcut getTimestamp(Client client):
        cflow(clientDataAccess(client)) &&
        call(* java.sql.ResultSet+.getTimestamp(..));

    // Advice que trata conversões de fuso horário nos pontos selecionados pelo
    // pointcut getTimestamp
    Timestamp around(Client client): getTimestamp(client) {

        TimeZone tz = client.getTimeZone();
        Timestamp t = proceed(client);
        Timestamp t2 = convertTimeZone(t, DEFAULT_TIMEZONE, tz);

        return t2;

    }

    // Método utilitário para conversão de timestamps entre fusos
    private Timestamp convertTimeZone(Timestamp t, TimeZone fromZone,
            TimeZone toZone) {
        Calendar result = Calendar.getInstance();
        result.setTime(t);
        int fromZoneOffset = fromZone.getOffset(t.getTime());
        int toZoneOffset = toZone.getOffset(t.getTime());
        int gap = toZoneOffset - fromZoneOffset;
        result.add(Calendar.MILLISECOND, gap);
        return new Timestamp(result.getTime().getTime());
    }

}
```

figura 4 – Aspecto de tratamento de fusos horários

Essencialmente, o que o aspecto faz é capturar chamadas a dois métodos da API de acesso a bases de dados do Java (JDBC): setTimestamp (da interface java.sql.Statement) e getTimestamp (da interface java.sql.ResultSet). Os filtros aplicados garantem que a captura ocorre apenas em chamadas feitas dos métodos que tratam operações e consultas de um cliente específico, ou seja, que possuem um objeto *Client* como primeiro parâmetro. Este objeto é capturado no parâmetro "*client*" dos *pointcuts*. Para o *pointcut* do método setTimestamp, é capturado também o parâmetro do tipo Timestamp que este recebe.

O *advice* que é ativado nas chamadas ao setTimestamp obtém o Timestamp passado como parâmetro, converte-o para o fuso horário de referência, e continua a chamada usando o valor convertido como parâmetro em lugar do que foi passado inicialmente.

O *advice* das chamadas a getTimestamp efetua a chamada normalmente, captura seu valor de retorno, converte-o para o fuso horário do

cliente e devolve esse valor como se fosse o retorno do método interceptado.

Ou seja, para o programador da regra de negócio, é como se os métodos setTimestamp e getTimestamp já cuidassem de toda a conversão necessária.

Entretanto, digamos que os requisitos de negócio mudam. Em vez de exibir os dados no fuso horário do cliente, deve existir na tela uma opção em que o usuário da aplicação escolha um fuso qualquer como referência.

Além de criar o controle na tela e o mecanismo para armazenar o fuso escolhido na sessão, basta trocar a primeira da figura 5 pela segunda.

```
TimeZone tz =
client.getTimeZone();
```

```
TimeZone tz =
TimeZone)SessionManager
.getSession().getAttrib
ute("timeZone");
```

figura 5 – Mudança de regra para tratamento do fuso

Na segunda linha, SessionManager.getSession() é um método estático que dá acesso ao objeto que representa sessão em curso na aplicação, talvez através de uma variável local à *thread*. Se esse objeto não estivesse disponível de forma estática, seria possível "puxá-lo" de níveis superiores da pilha de chamadas, da mesma forma que foi feito com o objeto *Client*.

Note-se que não é necessário fazer qualquer mudança nas classes de negócio ou acesso a dados. Além disso, se forem criadas novas operações que atuam sobre clientes e lidam com datas, elas terão os recursos de fuso horário aplicados automaticamente.

7. Conclusões

A programação orientada a aspectos pode ser uma ferramenta valiosa na criação de sistemas complexos como os de telecomunicações. Com ela, não apenas podemos melhorar os prazos e custos da implementação desses sistemas, como também melhoramos a sua qualidade e podemos oferecer recursos complexos como vantagem competitiva.

O uso da programação orientada a aspectos está crescendo, e ela está em um estado de maturidade em que já pode ser aplicada a qualquer sistema que utilize uma das linguagens com suporte a esta tecnologia.

Referências

[1] Aspectj project. http://eclipse.org/aspectj/.
[2] KIRSTEN, Mik. AOP tools comparison, Part 1: Language mechanisms. IBM DeveloperWorks, Fevereiro/2005 (http://www-106.ibm.com/developerworks/java/library/j-aopwork1/).

Curriculum Vitae

Cristiano Breuel formou-se em Engenharia de Computação pela Universidade de São Paulo e cursa mestrado em Ciência da Computação, com especialização em Programação Orientada a Aspectos, na mesma instituição. Trabalha na PT Inovação desde 2002, tendo participado do desenvolvimento, implantação e suporte de diversas versões da solução NGIN ProCare na operadora móvel Vivo. Atualmente, é responsável pelo desenvolvimento de diversos módulos da solução ProCare unificada para esta operadora.

02

O Impacto da Iniciativa JAIN™ no Desenvolvimento de Serviços de Rede Inteligente

Flávio Matiello (INB), João Batistella (INB)

palavras-chave:
JAIN, Java, JCC, IMS, NGIN, SCP, IN, SS7

O mercado de telecomunicações hoje enfrenta diversas dificuldades: resultados operacionais das operadoras não conseguem pagar os altos investimentos realizados, a competitividade no mercado é cada vez maior e a necessidade de lançar novos produtos e diferenciar-se dos concorrentes exige ainda mais investimentos. As operadoras têm que trabalhar com prazos cada vez mais curtos e orçamentos apertados, de forma a atender a demanda imposta por novos produtos.

Estes novos produtos, portanto, precisam ser desenvolvidos de forma simples, prática e eficiente, mas esbarram na complexidade de integração com diversas tecnologias, protocolos e fornecedores terceiros. Nesta perspectiva, surge a tecnologia JAIN, apoiada e especificada por um conjunto de empresas do ramo das telecomunicações para facilitar, agilizar a padronizar o desenvolvimento de novos serviços de forma simples, escalável, portável e eficaz, trazendo benefícios para fornecedores e operadoras.

A tecnologia Java já provou sua força no mundo da tecnologia da informação (TI), promovendo a portabilidade e padronização das aplicações, e facilitando a comunicação entre plataformas heterogêneas de *software*. Agora Java quer conquistar o segmento das telecomunicações, atualmente dominado por interfaces proprietárias.

Este artigo tem por objetivo apresentar a iniciativa JAIN, um esforço para a padronização das aplicações de telecomunicações, que promete oferecer a este segmento benefícios já alcançados em TI.

1. Introdução

Com o aumento da competição entre as operadoras de telefonia móvel, fica evidente a importância de se oferecer novos e exclusivos serviços, de modo a conquistar novos clientes e manter o *market-share* já alcançado. As oportunidades de mercado são únicas e o lançamento de novos produtos e serviços têm que ocorrer no tempo certo, sob a pena de perda de receita. O modelo de competitividade utilizado pelas operadoras deve-se basear cada vez mais na capacidade de diferenciação e inovação de seus produtos, evitando a briga por preços que as tem lesado nos últimos anos.

A inteligência existente nas redes móveis, como recurso capaz de viabilizar novos produtos ou serviços, fica sendo, portanto, a principal arma das operadoras para se diferenciar das demais e ampliar a rentabilidade de seus negócios.

Como ferramenta estratégica, as redes inteligentes deverão permitir disponibilizar novos produtos e serviços de forma rápida, escalável e eficaz.

Já os fornecedores destes novos produtos e serviços enfrentam uma série de dificuldades como prazos apertados, ausência de mão de obra especializada e integração com elementos e interfaces de diversos fornecedores, cada um com as suas peculiaridades.

Estes novos serviços requerem facilidade de desenvolvimento, agilidade e eficácia na sua integração, escalabilidade, eficiência operacional, carregamento e execução dinâmica de *software*. A tecnologia Java é a candidata ideal para prover tais competências.

Antes de pensar em mudar todos os dispositivos existentes em um operador para suportar Java, é preciso considerar o grande investimento já realizado para a montagem de infra-estrutura, o que torna inviável qualquer tentativa de se substituir todo o *software* existente por Java. Desta forma, é preciso escolher um caminho evolutivo e não revolucionário para se atingir o objetivo. A tecnologia Java API *for Integrated Networks* (JAIN) provê uma elegante solução para o problema. JAIN é um conjunto de API de redes integradas para a plataforma Java, que contém um *framework* para a criação e integração de soluções (ou serviços) em redes de pacotes (IP, ATM), *wireless* e PSTN (*Public Switched Telephone Network*). O objetivo da API JAIN é oferecer portabilidade de serviços, convergência de redes e segurança no acesso (serviços ficam fora da rede).

A API JAIN é definida e especificada por um grande número de empresas do segmento das telecomunicações (a comunidade JAIN, composta por empresas como, por exemplo, Nokia, Motorola, Sun, Vodafone, etc.), de acordo com um processo muito bem documentado, conhecido como *Java Community Process* (JCP).

A iniciativa JAIN está divida em duas categorias de API:

1. *Interfaces Container* – especifica

as API para ambientes de execução de serviços;

2. Interfaces de Aplicação – especifica API para o desenvolvimento de serviços sobre redes diversas (IP, PSTN, *Wireless*, etc).

A tecnologia JAIN faz uso da portabilidade fornecida pela plataforma Java, padronizando o fluxo de informações desde a lógica de sinalização, através da definição de API, e provê um *framework* para criação, teste e *deployment* de serviços que fazem uso desta API. Os pontos fortes da tecnologia se concentram em portabilidade, convergência de redes e segurança no acesso:

> **Portabilidade de Serviços:** o desenvolvimento de aplicações é, atualmente, limitado por interfaces proprietárias. A portabilidade de aplicações é praticamente inexistente. Isto aumenta o custo de desenvolvimento de aplicações, eleva o tempo para entrega e prejudica a manutenção do *software*, pois este requer mão de obra extremamente especializada e, muitas vezes, de alto custo. A abordagem do JAIN é substituir as interfaces proprietárias por interfaces Java uniformes, possibilitando o desenvolvimento de aplicações portáveis;

> **Convergência de Redes:** Uma chamada ou sessão nas aplicações atuais tipicamente ocorre apenas sobre um tipo de rede: PSTN, dados ou *wireless*, embora já existam *gateways* que permitem a integração entre estas redes. O modelo de chamadas de alto nível da tecnologia JAIN inclui facilidades para o tratamento de chamadas, onde uma chamada é entendida como uma sessão multimídia, entre múltiplas partes, utilizando vários protocolos, sobre uma rede integrada;

> **Segurança no acesso:** Aplicações ou serviços de comunicação podem correr dentro da rede do operador ou completamente fora desta. A interface JAIN *Service Provider* permite que serviços que estão fora da rede acessem recursos para executar funções da rede de forma segura.

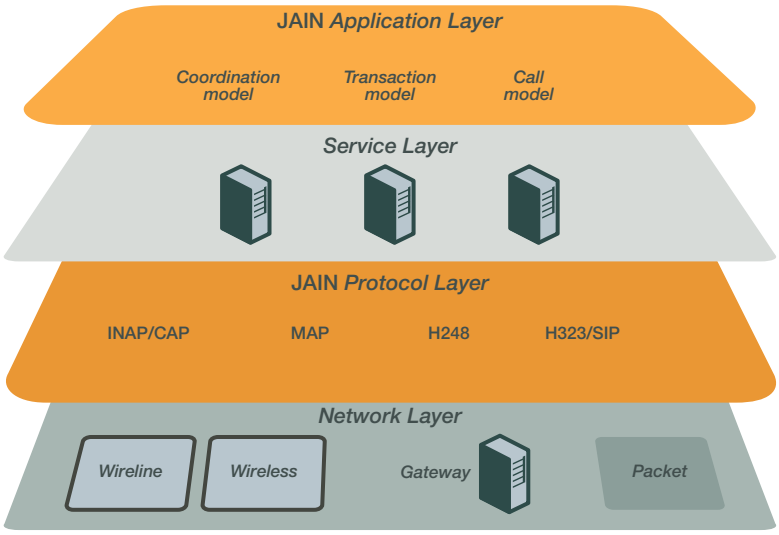

figura 1 – Camadas da arquitetura JAIN

A tecnologia JAIN cria um novo ambiente para desenvolvedores e usuários de serviços de comunicação, através de sistemas construídos sobre padrões de mercado que tem a garantia de funcionar sobre redes que os suportem. As oportunidades de negócio para os novos serviços serão imensas, pois, ao abrir a rede para aplicações Java, está criada uma oportunidade para se utilizar milhares de serviços portáveis e integrados, ao invés dos poucos atualmente disponíveis. A utilização de aplicações Java permite também o reuso de uma vasta quantidade de componentes já existentes, e o apoio de ferramentas de desenvolvimento, depuração e testes, além de possuir a disposição uma grande quantidade de desenvolvedores pertencentes à comunidade Java, trazendo, a médio prazo, uma redução no custo do desenvolvimento e nos prazos de entrega.

A tecnologia JAIN também abre caminho para a convergência entre Internet e PSTN de forma mais simples, permitindo que os serviços sejam desenvolvidos de forma abstrata, independentemente da rede em que serão disponibilizados.

2. Arquitetura JAIN

A arquitetura JAIN é composta por um conjunto de API Java de redes integradas e de um ambiente para executar componentes ou aplicações JAIN em redes PSTN, de pacotes (IP, ATM) e redes sem fio. A abordagem JAIN separa a lógica de serviço de questões específicas da rede em uso. Isso significa que os serviços desenvolvidos podem ser disponibilizados em qualquer tipo de rede com suporte a JAIN.

A figura 1 ilustra a arquitetura JAIN subdividida em camadas.

> **JAIN *Protocol Layer*** – padronização Java para protocolos de rede, que especifica interfaces para IP, redes sem fio e protocolos de sinalização. Dentre estes protocolos estão SIP, TCAP, ISUP, INAP, MAP, SIP, MGCP, Megaco (H.248) e H.323;

> **JAIN *Application Layer*** – padronização de interfaces para o desenvolvimento de aplicações independentes do protocolo de rede utilizado. Dentre as principais especificações estão: JCC (Java *Call Control*) e SLEE (*Service Logic Execution Environment*).

A camada de protocolo na arquitetura JAIN é baseada em uma padronização Java para protocolos específicos (SIP, H.248, H.323, TCAP, ISUP, INAP/CAP, MAP, etc.). Provendo interfaces padronizadas

para o acesso a diferentes protocolos em um modelo Java orientado a objetos, aplicações e pilhas de protocolo podem ser dinamicamente substituídas e, ao mesmo tempo, temos um elevado grau de portabilidade para o *software* na camada de aplicação, que pode utilizar pilhas de variados fornecedores.

A camada de aplicação na arquitetura JAIN proporciona um modelo genérico de chamada (ou sessão) sobre redes heterogêneas. A idéia fundamental é prover uma única interface entre as camadas de serviço e rede, permitindo a criação de serviços independentes da rede.

Na figura 2 temos uma representação em camadas completa da arquitetura JAIN.

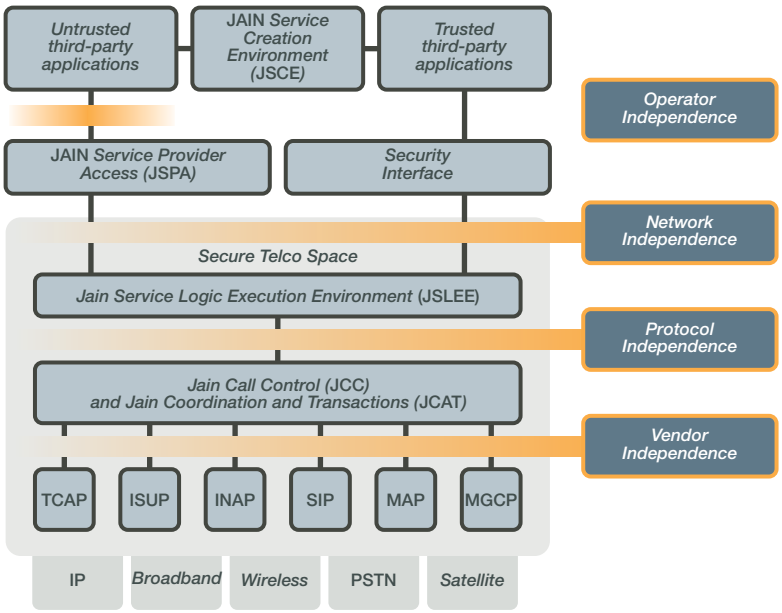

figura 2 – Detalhe das camadas JAIN

Um dos componentes centrais nesta arquitetura é o JSLEE (Java *Service Logic Execution Environment*). JSLEE está para as aplicações de comunicação como J2EE (agora conhecido como Java EE) está para os sistemas de informação. Trata-se de um *container* que provê um conjunto completo de serviços necessários a um ambiente de suporte a aplicações de comunicação, tais como controle transacional, persistência, balanceamento de carga, segurança, *pool* de objetos ou conexões, etc. Isto é similar aos *Enterprise Java Beans* na especificação Java EE.

A figura 3 ilustra uma topologia de aplicações que faz uso do JAIN de modo a integrar diversas tecnologias.

Neste exemplo temos diferentes tecnologias interagindo para o tra-

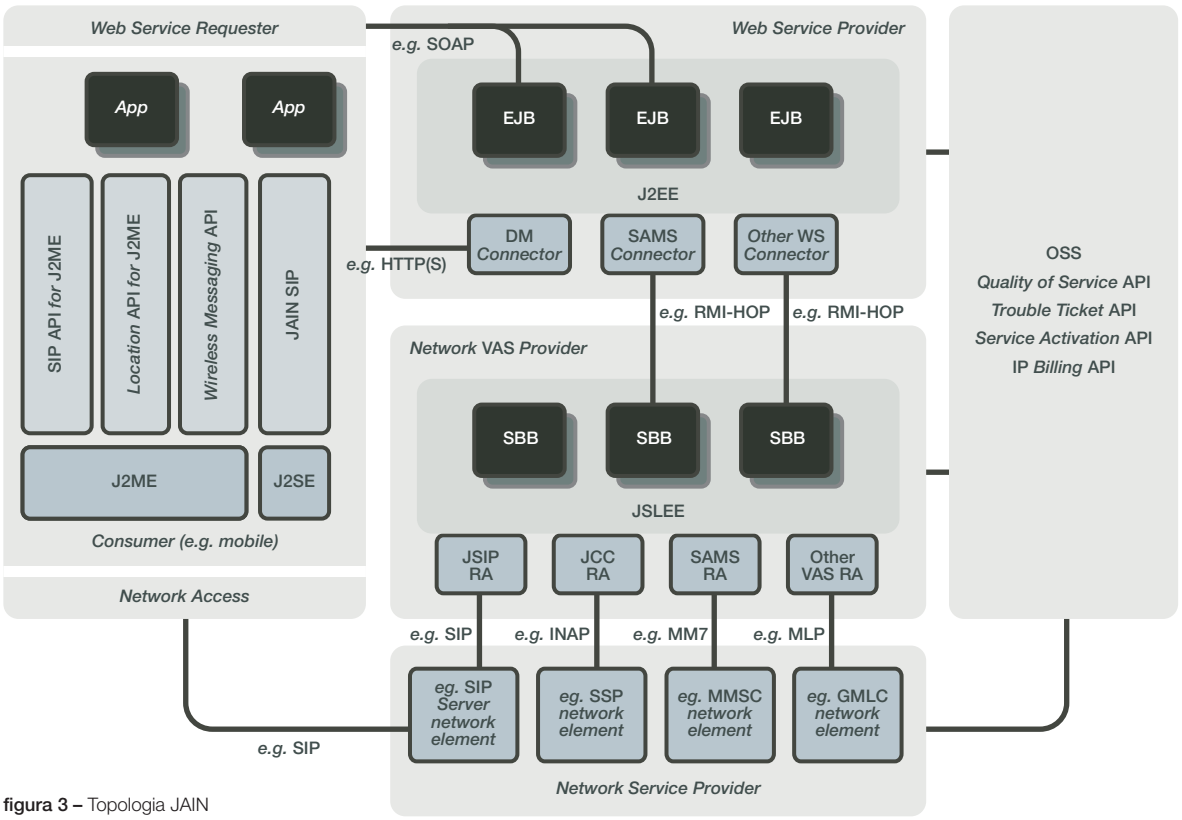

figura 3 – Topologia JAIN

tamento de chamadas (sessões). Por exemplo, uma chamada INAP pode chegar através do JCC RA (*Resource Adapter*), o que disparaaria a execução de um SBB (*Service Building Block*), o qual, com a ajuda de um EJB, invoca a lógica que verifica se um determinado cliente possui saldo para realizar uma chamada.

3. SLEE

O SLEE, conforme já referido, é um container de serviços definido pela JSR-22 [8], que provê um conjunto de funcionalidades e infra-estrutura para facilitar o desenvolvimento e *deployment* de serviços, fornecendo recursos como:

> Suporte a *queue* de eventos;

> Integração com JCC para tratamento de eventos;

> Definição de conceitos básicos para homogenização de serviços como SBB, *Event, Activity, Service* e *Profile*, os quais falaremos mais a frente;

> Tolerância a falhas, replicação de estado de SBBs e fluxo de execução;

> Suporte a *tracing* de chamadas, alarmística, gerenciamento e *timers*.

Vamos abaixo explicar um pouco mais sobre cada um dos componentes definidos pelo SLEE:

> **SBB: *Service Building Block***, é um componente reutilizável que define uma unidade funcional. Pode ter seu estado replicado entre outros servidores SLEE de forma a garantir que o serviço continua a atender os clientes na eventual queda de um servidor. Este componente se registra no *container* SLEE para ser avisado na ocorrência de determinados eventos;

> ***Profile*:** São os dados relativos a um determinado domínio de informação relativo a chamada, ou seja, em termos práticos, são os dados dos clientes que estão realizando a chamada que o serviço registrado no servidor SLEE está tratando;

> **JCC: *Java Call Control***, é uma especificação funcional que um *Resource Adaptor* (RA) deve implementar, de forma a tornar transparente ao serviço que é executado dentro de um conjunto de SBBs qual a tecnologia de rede ou protocolo de comunicação utilizado.

O SLEE é, portanto, um ambiente padronizado para execução de aplicações de comunicação, de forma análoga a um *container* Java EE no que diz respeito a sistemas de informação. Serviços desenvolvidos para o padrão JAIN SLEE são estruturados como componentes reutilizáveis orientados a objetos, e que podem ser executados em qualquer implementação que atenda à especificação JAIN SLEE.

4. Java *Call Control*

Desenvolvida pelo JCC Expert Group [7], a API JCC provê interface para um modelo genérico de chamadas, que oferece ao programador uma poderosa e conveniente abstração para manipulação de chamadas (ou sessão), e para o gerenciamento da interação entre aplicações e chamadas. Além disto, a API JCC é extensível, desta forma, se novas funcionalidades são requeridas, elas podem ser incorporadas de forma modular. Um requisito específico da JCC é ser independente de protocolo, isto é, aplicações criadas sobre o JCC não precisam se preocupar com questões específicas do protocolo de rede sobre o qual vão ser executadas.

A API JCC é uma interface Java para a criação, monitoramento, controle e manipulação de sessões de comunicação independente da tecnologia de rede utilizada.

JCC permite que as aplicações desenvolvidas segundo a especificação possam ser executadas em qualquer plataforma que suporte a API, aumentando o mercado para estas aplicações. A API JCC também permite que os provedores de serviços de rede possam oferecer, de forma rápida e efetiva, novos serviços aos usuários finais, através do desenvolvimento interno ou de terceirização, o que é facilitado pelo fato de se lidar com uma interface aberta.

A figura 4 ilustra como seria o uso da especificação JCC em uma implementação real. Por esta visão fica clara a independência das aplicações desenvolvidas em relação à tecnologia de rede utilizada.

A API foi especificada de modo a modelar os aspectos físicos e lógicos envolvidos em uma sessão (chamada) sob a forma de objetos. Aplicações interagem com este modelo segundo o padrão de projeto *Listener*, onde objetos se registram para ser notificados de eventos que lhe possam ser úteis no contexto de alguma lógica.

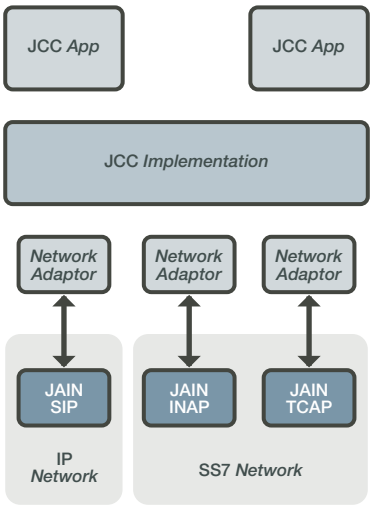

figura 4 – Visão de implementação JCC

4.1 Cenários de utilização de JCC

Em termos práticos, a implementação de uma lógica utilizando JCC necessita da codificação de um listener que é acionado na ocorrência de determinados eventos.

A seguir temos dois exemplos de cenários básicos de chamadas e a descrição sucinta das API invocadas no *listener* implementado. Vamos considerar no nosso exemplo apenas os eventos para uma das partes envolvidas na chamada, de forma a facilitar o entendimento.

Os cenários apresentam apenas os eventos relevantes ao cliente A:

> Chamada Originada – Cliente A origina chamada com sucesso para cliente B (figura 5);

> Chamada Terminada – Cliente B origina chamada com sucesso para cliente A (figura 6).

5. Arquitetura JAIN e NGIN

Um dos grandes desafios para nossas soluções de rede inteligente é identificar a melhor forma como a arquitetura de serviços baseados no *framework* JAIN pode-se integrar aos serviços que atualmente são fornecidos na solução NGIN.

Os serviços atualmente baseados na arquitetura NGIN tratam chamadas sinalizadas via protocolo SS7 (MAP, CAP, INAP) e sessões de dados e eventos via protocolo GTP [9] e OSA *Charging* [10]. Estes serviços são desenvolvidos com ferramentas proprietárias e correm em processos estáveis e escaláveis.

O cenário que pode propiciar uma boa integração e convergência entre as duas arquiteturas de soluções destacadas aqui (NGIN e JAIN) é a criação de uma camada capaz de tratar eventos oriundos dos dois cenários, rodando um processo Java *standalone* ou dentro de um *container* de aplicações Java EE (conforme demonstrado

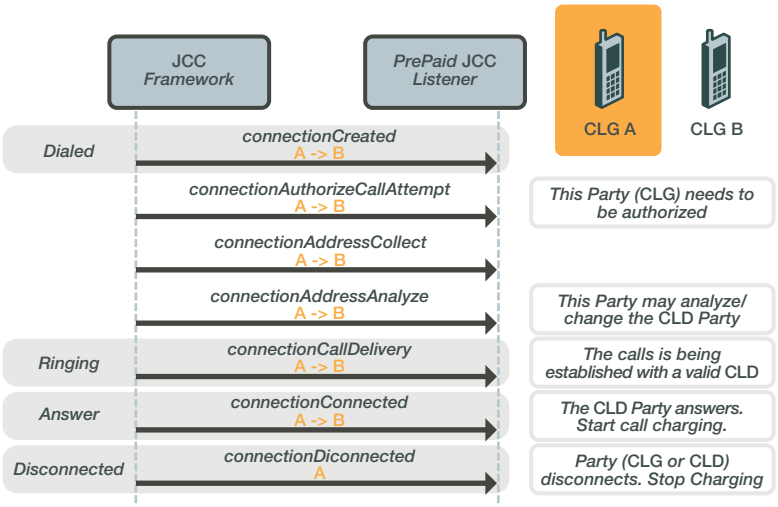

figura 5 – Fluxo JCC para Chamada Originada

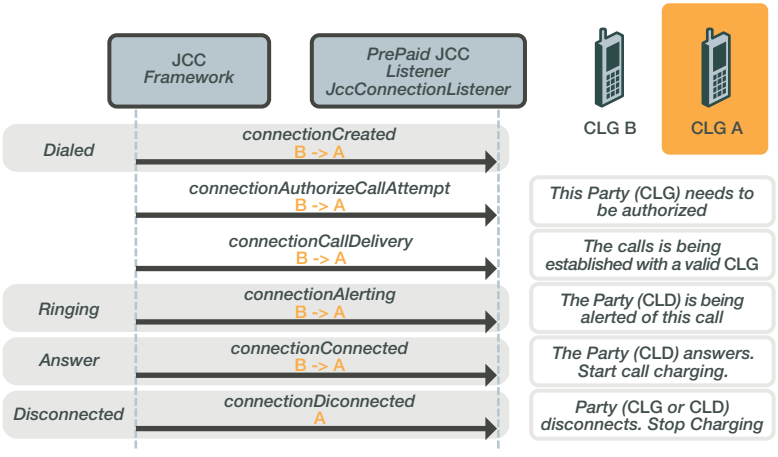

figura 6 – Fluxo JCC para Chamada Terminada

na figura 7), permitindo os seguintes ganhos:

> Separação dos dados da lógica de negócio;

> Instanciação de lógicas de serviços em qualquer uma das duas arquiteturas;

> Utiliza Java e todo seu conjunto de ferramentas, componentes e benefícios para a construção das lógicas de negócio.

A arquitetura de serviços proposta na figura 8 mostra uma abordagem física para união das arquiteturas tradicionais e *All-IP*.

A solução proposta pelo JAIN, então, vem-se integrar numa arquitetura de rede inteligente contendo

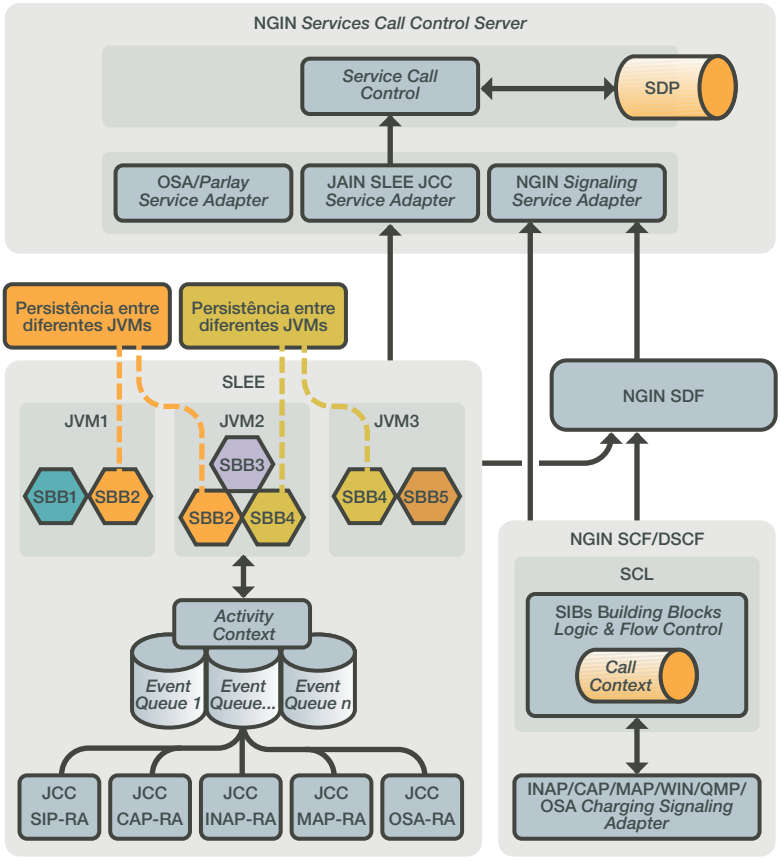

figura 7 – Integração de serviços JAIN e NGIN

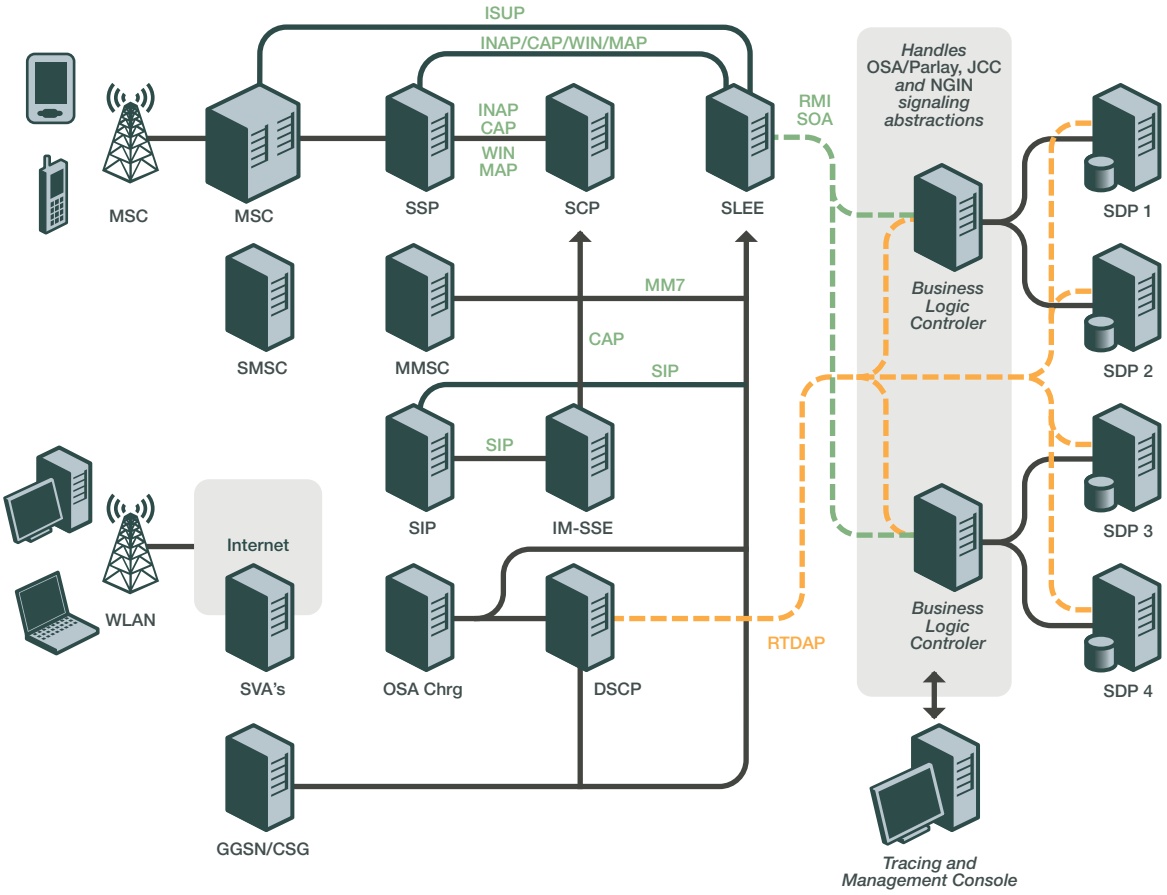

figura 8 – Arquitetura física de serviços JAIN e NGIN

o NGIN e diversos outros elementos de rede, com suporte a uma grande diversidade de protocolos de sinalização, permitindo que a mesma lógica de controle trate eventos de chamadas e sessões de dados.

6. Serviços Genéricos

Um grande avanço que pode ser realizado através da mudança de paradigma para o desenvolvimento em Java é a criação de serviços genéricos que agilizem, através de uma lógica robusta e estável, a criação de serviços de rede inteligente tradicionais como o de cartões pré-pagos.

Este modelo de serviço genérico deve fornecer uma arquitetura que seja suportada por chamadas encaminhadas no contexto JAIN ou no contexto NGIN.

Conforme a figura 9, o serviço genérico deve atuar em dois níveis:

> **Lógica de Negócio** (BCL): prover uma lógica funcional com controle de fluxo que permita de forma simples e eficaz tratar uma chamada, sessão de dados ou evento, fornecendo já pronta a integração com as bibliotecas NGIN para controle de serviços, saldos, notificação e *logs* já utilizadas nos serviços mais modernos em arquiteturas tradicionais;

> **Lógica de Sinalização:** prover a adaptação de questões como parâmetros de protocolo e localização, específicos das lógicas de sinalização NGIN, trazendo uma abstração total para a lógica de negócio (tal abstração já existe no caso do JAIN).

Esta arquitetura de serviços genéricos proposta pode ser implantada de diversas formas:

> **Anexo em um SDP:** funcionaria como uma nova entidade do SDP, separando a lógica de negócio e a base de dados, recebendo eventos via RMI, RTDAP ou *web-services*.

> **Em um ServiceLogicServer:** funcionaria como uma nova entidade NGIN, podendo fazer a agregação de vários SDP do ponto de vista de lógica de negócio (ou até mesmo da forma inversa), permitindo assim escalar a solução no ponto exato de maior sobrecarga identificado (rede, lógica de negócio e base de dados) , recebendo eventos via RMI, RTDAP ou *web-services*;

> **Em um servidor SLEE:** funcionaria já dentro de uma arquitetura JAIN, sendo necessária a implementação de um *Resource Adapter* com JCC para o NGIN.

É importante salientar que esta estrutura proposta é independente da utlização de JAIN, pois permite trazer para o mundo Java as lógicas de negócio com os seguintes ganhos:

> Melhores IDE e ferramentas de *debugging;*

> Um universo de componentes disponíveis;

> Suporte a *multi-threading* e AOP (*Aspect Oriented Programming*);

> Utilização de componentes de mercado que são robustos, flexíveis e estáveis;

> Utilização das boas práticas e padrões de projeto existentes para o desenvolvimento orientado a objetos.

7. Java *Specification Requests* para comunicações

Para que o leitor possa compreender a abrangência e ambição da iniciativa JAIN, apresentamos a seguir uma lista com todas as 21 especificações Java em andamento, no que diz respeito às comunicações.

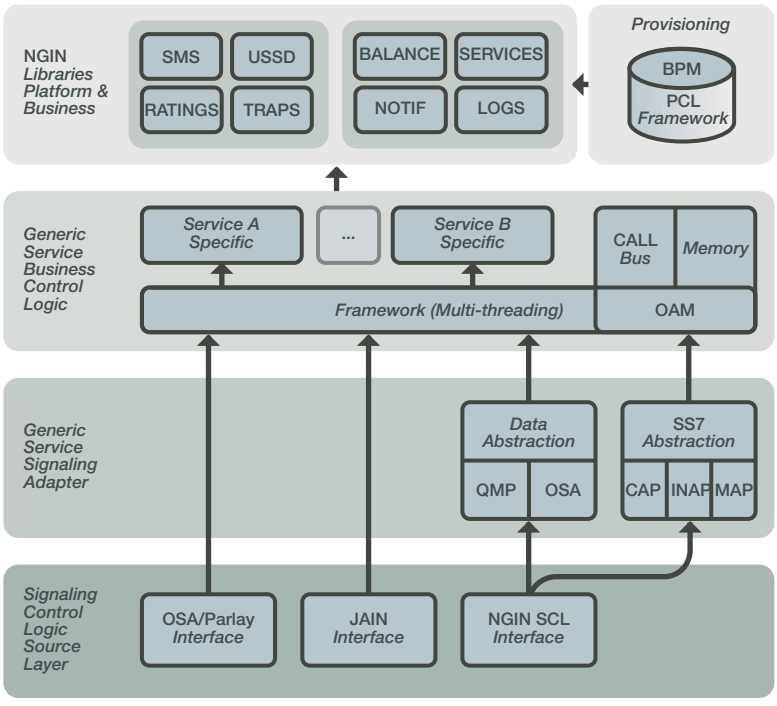

figura 9 – Arquitetura Serviços Genéricos

JSR	Descrição
JAIN SIP 1.1	Interface portável para compartilhamento de informação entre clientes e servidores SIP, provendo elementos de controle de chamadas, viabilizando redes integradas convergentes.
SIP API *for* J2ME 1.0	Define uma API SIP para clientes J2ME.
JAIN MGCP 1.0	Suporte ao protocolo Media *Gateway Control Protocol* (RFC 2705)
JAIN MEGACO	Suporte ao protocolo MEGACO/H.248
JAIN SDP	Descreve mensagens para sessões multimídia
JAIN ENUM	Define uma API para consultas e aprovisionamento de número de telefone no padrão E.164
JAIN TCAP 1.1	Especificação para a camada TCAP do protocolo SS7
JAIN INAP 1.0	Especificação para camada INAP do protocolo SS7
JAIN JCC 1.1	Especifica interfaces para o controle de chamadas (ou sessões).
JAIN JCAT	*Java Coordination and Transaction* (JCAT)
Java Payment API (JPay)	Especifica uma API aberta para pagamentos
JAIN *Presence*	Interface para gerência e manipulação de informações de presença
JAIN *Instant Messaging*	API multi-protocolo para serviços de mensagens instantâneas
JAIN *SIMPLE Instant Messaging*	Interface para troca de mensagens entre clientes SIMPLE
JAIN *SIMPLE Presence*	Interface para manipulação de informação de presença entre cliente e servidor SIMPLE
JAIN SIP *Lite*	Define uma API de alto nível para pilha SIP, de modo a permitir o ágil desenvolvimento de aplicações.
JAIN *Service Creation Environment* (SCE) - SCML	Visa simplificar a criação de serviços portáveis de telecomunicações para execução em containers JSLEE.
JAIN *Service Creation Environment* (SCE) - Java	API Java para suportar a criação de serviços portáveis de telecomunicações para execução em *containers* JSLEE.
Server API *for Mobile Services* (SAMS): *Messaging*	Define uma API multi-protocolo para envio e recebimento de mensagens curta de texto (SMS) e mensagens multimídia (MMS).
SIP *Servlets* 1.0	Define uma API de alto nível para servidores SIP. Permite que aplicações SIP sejam instaladas e gerenciadas tal como as baseadas na tecnologia *Servlet*.
JAIN *Service Logic Execution Environment* (JSLEE) 1.0	Define as interfaces, classes, exceções e requisitos para o desenvolvimento de serviços portáveis e *frameworks* de aplicação de modo que, uma vez que o serviço foi desenvolvido, funcionará em qualquer ambiente SLEE.

tabela 1 – Especificações Java no JCP para comunicações

8. Conclusão

A tecnologia JAIN existe, é uma realidade e está sendo estudada e evoluída por muitas empresas do setor, entretanto, ainda não está em larga exploração nas operadoras. Sua documentação ainda é um pouco escassa, mas cresce a cada dia o número de implementações de seus elementos, e novas empresas anunciam produtos compatíveis com esta tecnologia.

Acreditamos que nossos sistemas deverão ter suporte a este tipo de tecnologia, que hoje não tem um tempo de resposta e estabilidade suficientes para utilização em um sistema de larga escala como um pré-pago, mas acreditamos que a arquitetura de serviços poderá evoluir com esta linha paralela, para uma futura integração total.

O uso de JAIN para serviços de redes *All-IP* é uma ótima opção, sendo um reforço de grande peso neste tipo de arquitetura. Para redes tradicionais com grande volume, tempos de resposta extremamente curtos e necessidade de estabilidade altíssima, JAIN ainda é uma promessa, sendo a arquitetura de rede e sinalização atual uma melhor opção.

9. Próximos Passos

Colocamos abaixo os próximos passos para as pessoas interessadas em estudar o assunto:

> Entendendo o funcionamento do SLEE: ler artigos [5] e [6] sobre o JAIN e SLEE;

> Experimentando o SLEE: Baixar o servidor SLEE *open-source* da Mobicents, exemplos e o *plugin* [1];

> Documentação das JSRs do JCP: Analisar as JSRs e as especificações fechadas e suas respectivas implementações em [2].

Referências

[1] Implementação open-source SLEE hospedado no portal java.dev.net. Ver: https://mobicents.dev.java.net/
[2] Lista das JSRs da iniciativa JAIN no JCP. Ver: http://jcp.org/en/jsr/tech?listBy=2&listByType=tech
[3] Artigo sobre SLEE da Sun. Ver http://java.sun.com/products/jain/article_slee_principles.html
[4] Artigo sobre SLEE da jainslee.org. Ver http://jainslee.org/slee/factors.html
[5] Introducing JAIN SLEE Tutorial. Ver http://jainslee.org/downloads/jainslee-tutorial-04.pdf
[6] Sun JAIN SLEE Tutorial. Ver http://java.sun.com/products/jain/JAIN-SLEE-Tutorial.pdf
[7] Sun Microsystems, "JAIN Java call control (JCC) API Java Specification Request (JSR) 21", 2001. Ver http://jcp.org/jsr/detail/21.jsp
[8] Sun Microsystems, "JAIN SLEE API Specification (JSR) 22". Ver http://jcp.org/en/jsr/detail?id=22
[9] GPRS Tunnelling Protocol. Ver http://en.wikipedia.org/wiki/GPRS_Tunnelling_Protocol.
[10] Parlay/OSA Specifications. Ver http://www.parlay.org/en/specifications/.

Flávio Matiello, graduado em Ciência da Computação pela UFSCar e pós-graduando em Administração de Empresas pela FGV. Analista de Sistemas da PT Inovação Brasil desde 2002, esteve envolvido no suporte, implantação, requisitos e desenvolvimento da solução NGIN no Brasil com os serviços de voz e dados, trabalha atualmente com a equipe de desenvolvimento de serviços Corporativos e de Dados para a Vivo no Brasil.

João Batistella, graduado em Engenharia de Computação pela Unicamp e pós-graduando em Gestão de Projetos pela FGV. Engenheiro da PT Inovação Brasil desde 2003, esteve envolvido no suporte, implantação e desenvolvimento da solução NGIN no Brasil, envolvido com diversos módulos da solução NGIN ProCare, trabalha atualmente com a equipe de desenvolvimento de módulos do ProCare para a Vivo no Brasil.

03

SCA – Sistema de Controlo de Acessos

Clara Guerra,
João Filipe (Maisis),
José Antunes (Maisis)

palavras-chave:
SCA, Utilizador, Perfil, Sistema de Gestão, Subsistema, Ponto Intermédio, Tipo de Acesso, Centro de Gestão, Domínio Gerido, Família de Domínios Geridos, Matriz de Possibilidades, Língua, Língua Preferencial, Multilingue, *Single Sign On*, NGIN

Criar uma visão integrada para o cliente das aplicações/sistemas que a PT Inovação disponibiliza hoje, muitas vezes dispostas em ambiente de portal aplicacional, pode ser um factor crítico de sucesso. A disponibilização de um único ponto de autenticação e um controlo centralizado dos privilégios de cada utilizador dos vários sistemas, contribui para dar essa visão mais integrada das soluções. É adicionalmente um factor de obtenção de sinergias internas e de garantia de segurança para os clientes.

O conceito de *single sign on* e a gestão centralizada de utilizadores e perfis de acesso são objectivos primordiais do Sistema de Controlo de Acessos (SCA). Este é um sistema independente, que se apresenta como uma solução de gestão de segurança e que é facilmente integrado com qualquer aplicação ou conjunto de aplicações duma mesma plataforma empresarial.

Este artigo apresenta o SCA nas suas características técnicas e funcionais, mas também na perspectiva das vantagens apercebidas por algumas empresas clientes que já o integraram nas suas aplicações e portais e que começam a descobrir nele uma solução de controlo de acessos e segurança com um potencial de transversalidade e unicidade aplicável a todos os seus sistemas de gestão.

1. Enquadramento

Um sistema de Controlo de Acessos dá suporte a outros sistemas, colmatando nestes a necessidade da gestão e controlo dos acessos via interface de utilizador. O SCA da PT Inovação apresenta uma solução modular para gerir utilizadores e privilégios de acesso de um ou mais sistemas em simultâneo. As suas duas componentes básicas, GUI (*Graphical User Interface*) e API (*Application Programming Interface*), permitem uma integração simples com qualquer sistema cliente. Através da sua componente GUI, o administrador de acessos configura o sistema alvo, o universo gerido, os utilizadores, perfis, etc. Através da componente API, o SCA fornece os serviços de autenticação, validação de perfis de acesso e todo o tipo de consultas que o sistema cliente necessite fazer, em tempo real, sobre a informação registada no SCA.

Sendo o controlo de acessos uma necessidade básica comum a todos os sistemas que interagem com utilizadores humanos, é prática corrente que cada sistema tenha o seu próprio módulo de controlo de acessos. A arquitectura NGIN, preocupada com a importância de que todas as aplicações que coexistissem numa mesma plataforma tivessem uma política de controlo de acessos comum, veio alterar a prática corrente. A ideia era usar o mesmo módulo de gestão de acessos para alimentar um repositório centralizado dos utilizadores e privilégios de todas as aplicações presentes numa dada instanciação ou plataforma, possibilitando que um utilizador se autenticasse uma única vez para ter acesso às várias aplicações em que era autorizado. Assim nasceu o NGIN SCA, herdando parte de um SCA desenvolvido anos atrás para os projectos LONGA e GEREX. Nos últimos tempos perdeu a sigla NGIN e voltou à designação inicial, dado que já não é só para as soluções NGIN que o SCA se confirma como a ferramenta de controlo de acessos de referência. Vários sistemas de gestão de redes e serviços desenvolvidos na PT Inovação e instalados na PT Comunicações, TMN e outras empresas PT têm vindo a adoptar o SCA.

Este artigo pretende dar a conhecer a solução SCA da PT Inovação, realçando as suas características técnicas mas também a mais valia que apresenta para os clientes na sua utilização. Começa por enquadrar o SCA na breve história do seu percurso, desde a sua concepção inicial à fase actual. Um olhar mais técnico apresenta-nos em seguida a arquitectura, tecnologias e componentes funcionais. Segue-se uma explanação das principais entidades domínio e de como se manipulam na interface de administração do SCA. Realce dado a dois aspectos fundamentais: a capacidade de comutação automática da língua da interface gráfica, de acordo com a preferência de cada utilizador (atributo configurável mediante os dicionários disponíveis) e a contribuição do SCA para a implementação de um único ponto de autenticação (*single sign on*) quando o acesso às diversas aplicações se faz através de um portal. A finalizar remete-se a palavra para os clientes de

soluções PT Inovação, dando-lhes a oportunidade de exporem as suas opiniões sobre este tema. Algumas conclusões encerram o artigo, resumindo contextos passados, presentes e futuros nesta temática do controlo de acessos nas aplicações.

2. Componentes Funcionais

O SCA apresenta-se em três componentes principais distintos:

> A interface gráfica para administração de utilizadores, perfis, aplicações, universos geridos e outra informação. É suportada em tecnologia *web* e disponibilizada via *Internet*/*Intranet* a partir de um servidor aplicacional (e.g. *Tomcat*, *Weblogic*, etc.);

> A API que permite às aplicações interagirem com o SCA para autenticação e validação das permissões dos seus utilizadores. Esta API é disponibilizada em versão Java e é linkada na origem com as aplicações que pretendem interrogar a base de dados do SCA para efeitos de controlo de acessos;

> A base de dados Oracle, onde está o repositório da informação gerida, bem como um conjunto de procedimentos PL/SQL que constituem uma API intermédia de acesso aos dados.

A figura 1 ilustra as componentes de GUI, API e BD do SCA.

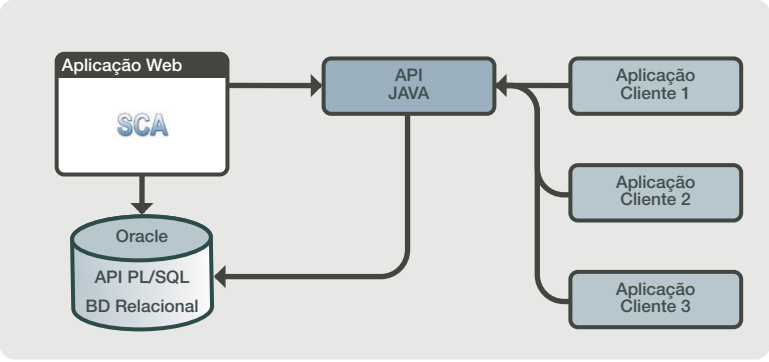

figura 1 – Componentes Funcionais do SCA

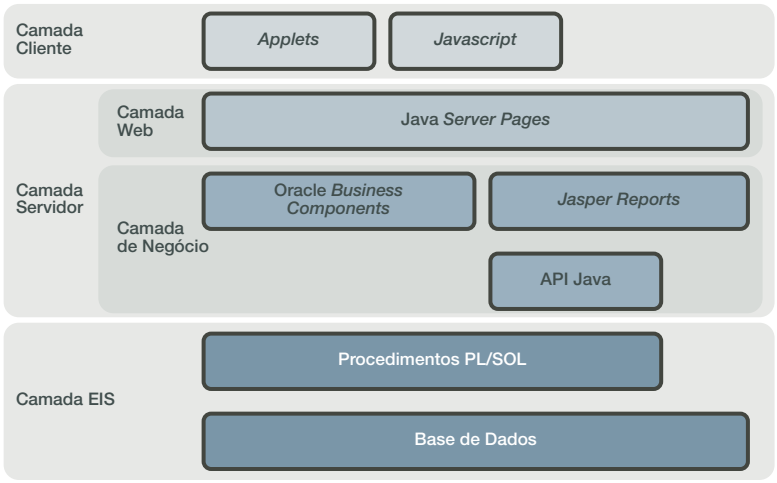

figura 2 – Decomposição em camadas

3. Arquitectura

A aplicação de administração do SCA baseia a sua arquitectura no modelo da plataforma J2EE (Java 2 *Enterprise Edition*), que apresenta uma arquitectura distribuída e por camadas. Assim, a lógica da aplicação é dividida em componentes, de acordo com as suas funcionalidades, e corre em diferentes máquinas conforme a camada e ambiente a que pertence (figura 2).

Os componentes da aplicação SCA dividem-se nas Camadas de Cliente, Servidor e EIS (*Enterprise Information System*) ou de Base de Dados. A Camada de Servidor pode ainda, para efeitos de melhor percepção das suas funcionalidades, ser dividida em Camada *Web* e Camada de Negócio.

O *software* utilizado em todas estas camadas é livre, excepção feita para a camada de Base de Dados que assenta em *software* Oracle.

Nas linhas seguintes descrevem-se os principais componentes de *software*:

> **Aplicação *Web* –** Conjunto de páginas JSP e classes Java que implementam a interface de administração do SCA, visualizável num *browser* cliente;

> **Jasper *Reports* –** Código *open source* que suporta a elaboração de relatórios *web*;

> **Oracle *Business Components* –** Componentes Java que realizam a interacção da aplicação *Web* com a base de dados;

> **API Java –** Código Java que implementa a biblioteca de serviços de controlo de acessos usado pelas aplicações clientes, incluindo a própria aplicação de administração SCA;

> **Base de Dados –** Contém o repositório dos dados do SCA bem como *packages* de PL/SQL usadas quer pela API Java, quer pela aplicação *Web*.

4. Conceitos SCA

Sendo o principal objectivo do SCA autenticar e autorizar os utilizadores que pretendem aceder a um dado conjunto de aplicações, existe à partida um certo número de conceitos que são essenciais ao seu domínio de gestão. Descrevem-se a seguir os mais importantes.

4.1 Universo Aplicacional

O universo aplicacional é constituído pelo conjunto das aplicações

alvo do controlo de acessos. Cada aplicação é um **sistema de gestão** que se decompõe em **subsistemas** e **pontos intermédios** de controlo. Cada ponto intermédio é associado a um **tipo de acesso** que corresponde ao privilégio mínimo requerido para ser executado. Esta caracterização do universo aplicacional irá permitir ao SCA definir que funcionalidades cada utilizador pode executar (figura 3).

4.2 Universo Gerido

O universo gerido compreende a informação de negócio de cada aplicação. Algumas características da informação de negócio definem critérios de particionamento desse universo que o SCA utiliza para definir **famílias de domínios geridos**. Este conceito serve para atribuir aos utilizadores privilégios sobre partes da informação gerida (figura 4).

4.3 Estrutura Organizacional

A estrutura organizacional compreende as empresas responsáveis pelas aplicações e seus utilizadores. Cada utilizador pertence a um ramo hierárquico da estrutura, chamado **centro de gestão**, onde lhe é autorizado ou vedado o acesso a determinados domínios de informação (figura 5).

4.4 Perfil

O **perfil** é concedido a um ou vários utilizadores e agrega um conjunto de funcionalidades (subsistemas) sob um determinado tipo de acesso. A definição dos perfis num determinado sistema de gestão é independente dos seus utilizadores e permite a reutilização e combinação de perfis, assim como uma maior simplicidade na sua configuração.

4.5 Utilizador

O **utilizador** é uma pessoa credenciada para aceder a um ou mais sistemas de gestão. O seu registo é constituído pela seguinte informação obrigatória:

> nome (*username*), palavra-passe (*password*), centro de gestão e língua preferencial;

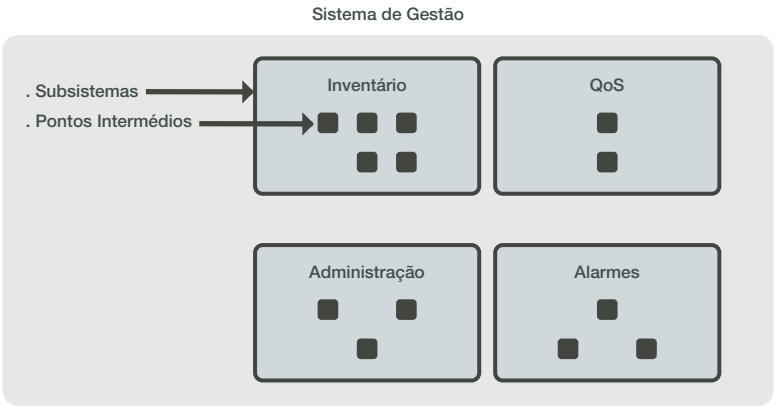

figura 3 – Decomposição do sistema de gestão em subsistemas e pontos intermédios

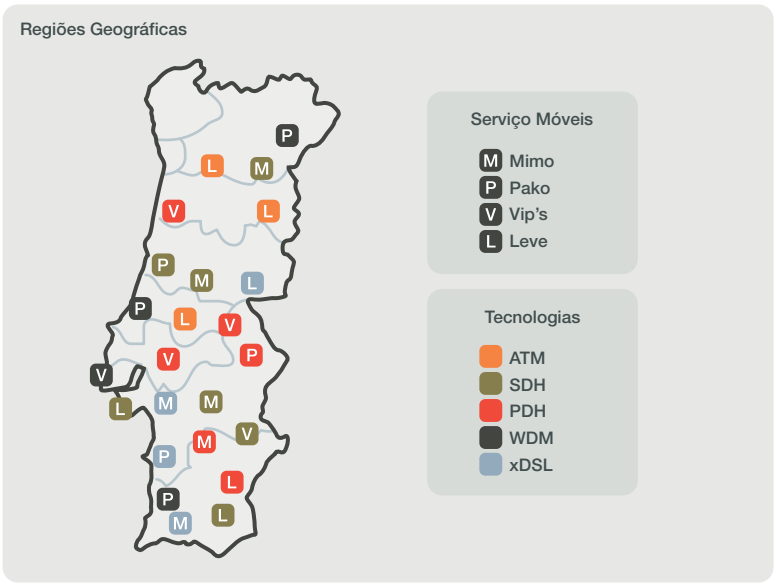

figura 4 – Exemplos de famílias de domínios geridos

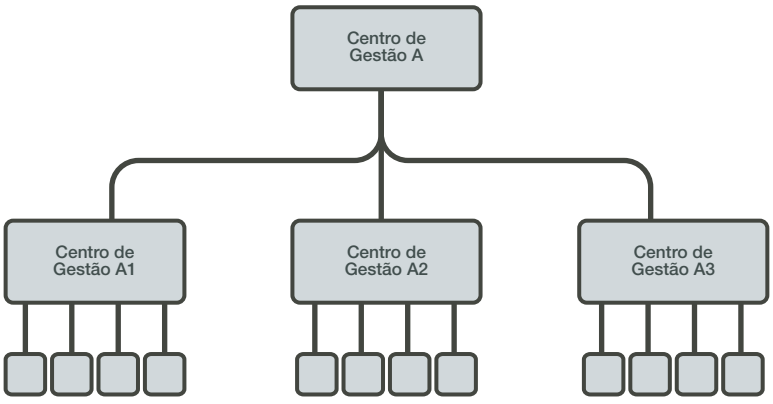

figura 5 – Estrutura organizacional da empresa operadora

e várias características opcionais:

> nome, morada, *e-mail*, máximo de tentativas de *login* e de sessões simultâneas, tempo máximo de sessão e de inactividade, etc.

Completa a informação de utilizador, o conjunto de sistemas autorizados e a lista de perfis determinam o conjunto de funcionalidades que ele pode executar em cada sistema de gestão.

4.6 Tipo de Acesso

O **tipo de acesso** indica o privilégio conferido a um utilizador para um determinado subsistema (mediante a atribuição de um perfil) ou, quando referente a um ponto intermédio, o privilégio mínimo que um utilizador deve possuir para ter acesso a este ponto. Por exemplo Consulta, Operação, Configuração, Administração são tipos de acesso típicos num sistema.

5. Interface Gráfica

A interface gráfica, implementada em páginas *web*, seguiu as recomendações coligidas num trabalho levado a efeito na PT Inovação, com o objectivo de servir de guia ao desenvolvimento de todas as interfaces gráficas das aplicações NGIN [2]. A sua orientação por menus é simples e conduz facilmente ao grupo de entidades que se pretende manipular (figura 6).

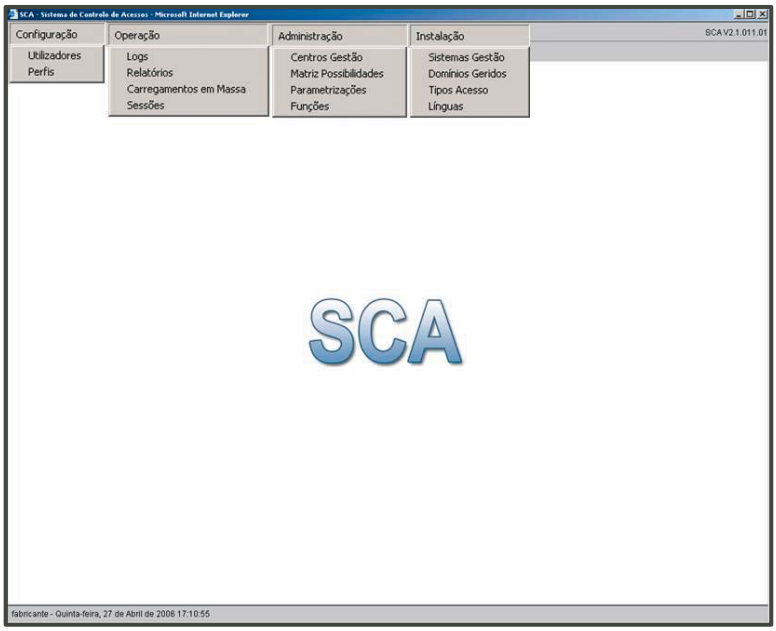

figura 6 – Estrutura de menus do SCA

Os menus do SCA apresentam-se em quatro grupos:

> Configuração;

> Operação;

> Administração;

> Instalação.

O menu **Configuração** é o mais vocacionado a acções do dia a dia, como seja configurar novos utilizadores e perfis genéricos, ou outras acções elementares associadas (alterar, apagar, associar, etc.).

O menu **Operação** permite visualizar os *logs* das acções efectuadas no SCA, bem como consultar diversos tipos de relatórios, efectuar acções de carregamento de utilizadores em massa a partir de ficheiros externos e ainda consultar todas as sessões abertas em aplicações geridas pelo SCA.

O menu **Administração**, como o próprio nome indica, possibilita acções tipicamente de administração, como sejam a definição dos centros de gestão, ou seja, a estrutura organizacional da empresa onde é instalado o SCA, a gestão das atribuições possíveis entre os centros de gestão e os domínios geridos – a matriz de possibilidades –, as parametrizações aplicáveis à organização ou a um determinado centro de gestão e as funções ou cargos aplicáveis aos utilizadores. São acções menos frequentes e que afectarão um grupo ou a totalidade dos utilizadores, sem contudo deixarem de ser normais em exploração.

O menu **Instalação** inclui a definição dos sistemas de gestão, dos domínios geridos, tipos de acesso e línguas. Poderíamos dizer que estas são configurações 'de fábrica', ou seja, aquelas que de uma forma ou outra dependem do *software* instalado. Praticamente todas as alterações a este tipo de entidades obriga à substituição dos executáveis de pelo menos alguma aplicação em campo. Em ambiente de exploração, as opções deste grupo apenas poderão ser consultadas, sendo qualquer outro tipo de acção executada por um *superuser* de nome “fabricante”, de uso muito restrito.

A título de exemplo apresentam-se duas páginas *web* do SCA, a primeira mostrando a lista geral de utilizadores e a segunda mostrando o histórico de utilização compactado referente a todas as sessões de um utilizador num dado período de tempo (figura 7).

6. Como integrar o SCA numa aplicação cliente

Para que as aplicações cliente consigam interagir com o SCA, este disponibiliza uma API em linguagem Java para esse efeito. Essa API tem evoluído à medida das necessidades e solicitações dos clientes, permitindo que estes efectuem "pedidos" de informação ao SCA.

Sempre que a API Java ou a interface *web* de administração do SCA necessitam de informações da base de dados, é invocada uma outra API elaborada em linguagem PL/SQL e que faz a ligação ao modelo de dados.

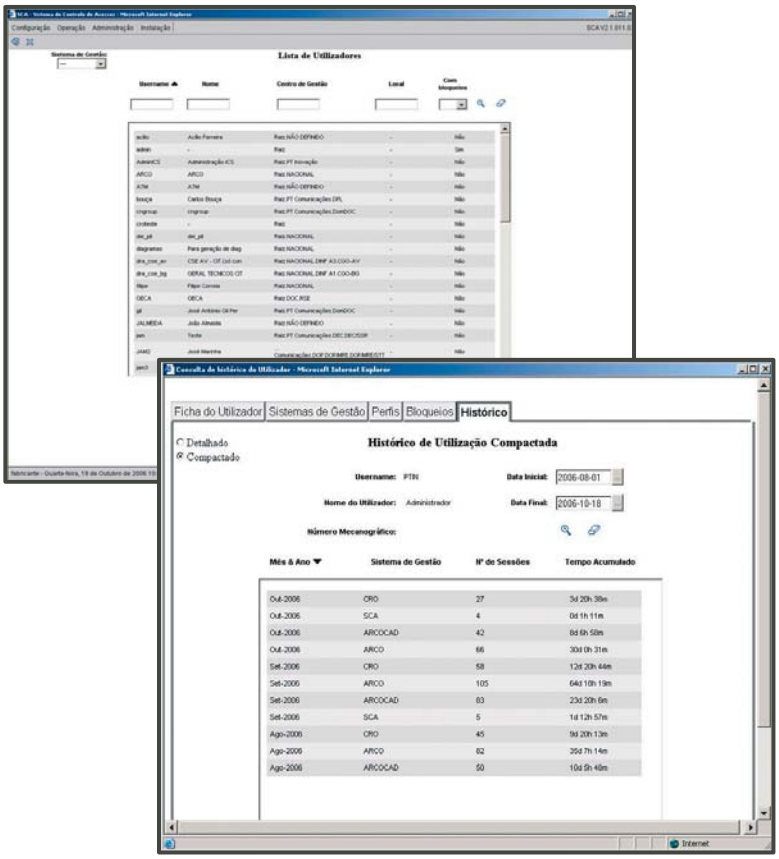

figura 7 – Páginas *web* da aplicação SCA

As API descritas facultam às aplicações cliente operações para autenticar utilizadores, alterar *password*, obter listas de acessos (sistemas de gestão, subsistemas, pontos intermédios, domínios geridos), etc. Através deste tipo de informação as aplicações cliente conseguem moldar-se de forma personalizada para cada utilizador.

Cada aplicação, em sede de desenvolvimento, deverá identificar os pontos de função onde haverá algum tipo de controlo de acessos e incluir nesses locais as invocações aos procedimentos ou métodos da API.

7. Single Sign On

A utilização do conceito *single sign on* permite a um utilizador autenticar-se apenas uma vez e ter acesso a múltiplos sistemas através dessa autenticação. Esta funcionalidade pode ser bastante útil em sistemas que apesar de serem independentes se encontram interligados e coexistem num mesmo ambiente.

Imaginemos o caso de um portal *web* onde a autenticação de um determinado utilizador nesse portal poderá dar acesso, dependendo das permissões desse utilizador, a um conjunto de aplicações (sistemas de gestão). Depois de efectuada a autenticação com sucesso no portal, o utilizador não necessitará de voltar a efectuar a autenticação nos vários sistemas de gestão que coabitam no portal. Para que esta funcionalidade seja implementada, basta que as aplicações tenham a capacidade de enviar ao SCA uma chave que é devolvida por este no momento em que foi efectuada a autenticação no portal. Quando o utilizador pretender entrar num determinado Sistema de Gestão, a chave identifica-o univocamente naquela sessão do portal, em vez do seu *username* e *password*.

O SCA está também preparado para ser uma das aplicações a coexistirem num ambiente portal, sendo possível entrar no SCA através de uma chave. Para isso, apenas será necessário adicionar ao URL de entrada a variável *SCASessionID*.

A partir do momento em que é encerrada a sessão no portal, a chave que identificava o utilizador deixa de ser válida, não sendo possível efectuar a entrada em mais nenhum sistema de gestão usando aquela chave.

8. Internacionalização

Internacionalização traduz-se na disponibilização de uma interface de utilizador multilingue, em ambientes onde se justifica a existência de várias línguas ou idiomas. Para tal, o SCA disponibiliza mecanismos para comutação automática e/ou manual da língua apresentada na interface do utilizador.

Para que as aplicações cliente tenham a possibilidade de disponibilizar também vários idiomas, o SCA tem a capacidade de disponibilizar a língua preferencial do utilizador logo após uma autenticação com sucesso. Deste modo, as aplicações cliente podem também comutar de língua para se adaptarem às preferências dos seus utilizadores (figura 8).

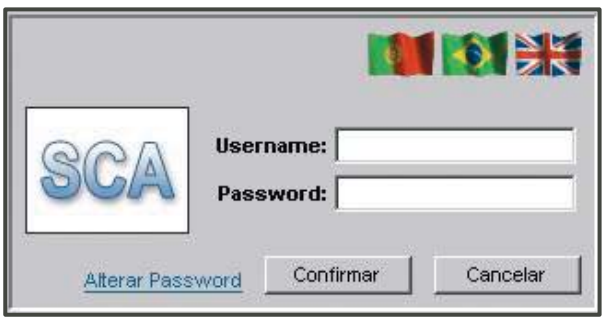

figura 8 – Janela de *Login* do SCA

9. A palavra é dos Clientes

> Ana Graça Gonçalves, da TMN/ DSO, responsável pelo departamento de Gestão de Redes e Serviços, fala sobre o valor acrescentado que encontra no SCA:

"A implementação do SCA nos sistemas de gestão da rede e serviços da tmn permite a optimização do processo de gestão dos utilizadores, através da gestão centralizada do acesso aos vários sistemas. É um grande valor acrescentado para a segurança dos sistemas e da informação."

> João Luís Veríssimo, da PTC/ DEC, responsável pelo departamento de Sistemas de Suporte à Gestão e Operação de Redes, avança com algumas expectativas:

"Uma das grandes utilizações do SCA será o de facilitar a implementação de Portal de Suporte às Operações de Rede (Manutenção e Supervisão) permitindo que as equipas de Supervisão e Manutenção da Rede possam aceder de forma centralizada e rápida aos vários serviços/funcionalidades de suporte ao diagnóstico e recuperação de falhas numa visão de serviço (E2E). Só com o recurso a este tipo de módulo se consegue integrar de uma forma simples e rápida estes vários serviços/funcionalidades, porque normalmente são suportadas por várias aplicações/ sistemas – Sistemas de Gestão de Elemento de Rede e de Rede."

> António Gumerzindo, da PT Inovação Brasil, Gestor de Cliente da VIVO afirma:

"O SCA tem-se mostrado bem adequado às necessidades de uma grande empresa como é a Vivo, respondendo de modo eficaz aos apertados requisitos de segurança que nos foram exigidos. Ao permitir uma gestão unificada dos perfis e usuários do NGIN Care, aliando elevada flexibilidade de customização, tem sido um dos pontos fortes da solução NGIN na Vivo. Com a sua extensão aos outros componentes da solução NGIN na Vivo, o Manager *e o BIT, vai permitir o* single sign-on *na solução NGIN."*

10. Objectivos futuros

Devido ao constante crescimento de utilizadores do SCA, o uso da tecnologia *Oracle Business Components* começa a apresentar algumas desvantagens em sistemas com uma elevada utilização. Assim sendo, está prevista a substituição desta camada por uma tecnologia com melhor comportamento e que responda às solicitações dos clientes de uma forma mais eficiente.

Outro dos objectivos do SCA é conseguir trabalhar sobre vários Sistemas de Gestão de Base de Dados, conseguindo assim uma independência em relação à Oracle que não existe neste momento. Com este ponto ultrapassado, várias aplicações que identificam o *software* Oracle como um constrangimento para o uso do SCA, poderão optar mais livremente pelo sistema de controlo de acessos.

11. Conclusões

O SCA implementa actualmente a função da gestão e controlo de acessos em vários sistemas desenvolvidos na PT Inovação. A chave do sucesso reside seguramente no facto de uma só equipa ter assumido o desenvolvimento centralizado de uma solução única, integrando os requisitos colocados pelos vários projectos e clientes que já o usam e não divergindo para múltiplas versões. Hoje, o sistema SCA implantado nas soluções NGIN na TMN, VIVO (Brasil) e UNITEL (Angola) é o mesmo que integra as várias soluções de OSS na PT Comunicações e na TMN como sejam os sistemas CADREDE, GEREX, SIGO/TTK e ALTAIA.

A par de um esforço constante de evolução e adaptação às necessidades emergentes dos novos sistemas na área da segurança e gestão de acessos, o maior objectivo continuará a ser a sua integração em cada vez mais soluções desenvolvidas na PT Inovação.

Referências

[1] Workshop 'Arquitectura NGIN' - acta de reunião TARGET_20021024.doc, PT Inovação 2002
[2] Pedro Santos, "Interfaces Gráficos NGIN - Normalização GUI", PT Inovação 2003.

Clara Guerra, licenciou-se em Matemática Aplicada, no ramo de Estatística e Computação, pela Faculdade de Ciências de Lisboa, em 1982. Iniciou a sua actividade profissional como Analista de Sistemas na *Time-Sharing* (hoje PT Contact) e mais tarde na COBA. Após uma passagem pela Universidade de Aveiro onde leccionou como assistente no Departamento de Matemática, ingressou em 1987 no Centro de Estudos de Telecomunicações, actual PT Inovação. Aí, participou no desenvolvimento de *software* para centrais digitais nos projectos ETD e ELD. Exerceu funções de investigadora em inúmeros projectos no âmbito dos programas europeus EURESCOM, RACE e ACTS, particularmente nas áreas de TMN, IN e Aplicações. Em 1994 foi nomeada responsável pela Unidade Funcional Engenharia de Sistemas de Informação, cargo que desempenhou durante um ano, tendo continuado a sua actividade na modelização e análise de sistemas de suporte e gestão de redes. Em meados de 2000 incorpora a equipa do projecto de certificação Global PTIN, no papel de Dinamizadora da Qualidade da área de negócio Sistemas de Gestão de Redes (SGR) e mais recentemente da área Sistemas de Suporte e Gestão de Redes (SSG). No âmbito deste mesmo projecto recebeu formação em auditorias de qualidade, tendo a partir daí integrado equipas de auditorias internas na PT Inovação. Ao longo de vários anos e até ao presente é o contacto técnico nomeado para as relações de parceria com a empresa Oracle, tendo durante esse tempo desempenhado funções de administração e *tuning* de bases de dados Oracle. Em Junho de 2005 foi nomeada responsável pelo projecto SCA que disponibiliza uma solução global de gestão e controlo de acessos para os sistemas desenvolvidos na PT Inovação.

João Filipe, bacharel em Engenharia Informática e de Sistemas pelo Instituto Superior de Engenharia de Coimbra. *Degree* em *Computer Science* com especialização em *Software Engineering* na Universidade de Hertforshire. Licenciatura em Engenharia Informática com especialização em Sistemas de Informação no Instituto Superior de Engenharia de Coimbra. Analista na empresa Maisis, envolvido em projectos de Supervisão de Redes Fixas de Telecomunicações. Definição do *workflow* para o *Help-Desk* do Centro de Supervisão da TMN. Concepção e desenvolvimento de procedimentos na TMN, no sentido de melhorar a comunicação e eficácia dos seus fornecedores, tendo como objectivo aumentar a eficácia da rede móvel e assim aumentar a qualidade relativamente aos clientes empresariais e pessoais. Gestor de relações profissionais entre a empresa MAISIS e a PT Inovação. Decisor no processo de negociação de contratos profissionais entre MAISIS e PT Inovação. Gestor de recursos humanos e director de inovação da Maisis. Presidente da direcção da Associação Inova-Ria - Associação de Empresas para uma Rede de Inovação em Aveiro.

José Antunes, licenciado em Engenharia Informática pela Faculdade de Ciências e Tecnologia da Universidade de Coimbra. Actualmente é colaborador da Maisis - Sistemas de Informação, onde integra a equipa de desenvolvimento do Sistema de Controlo de Acessos da PT Inovação.

04

Modelos de Informação para Gestão de Redes de Transporte – Aplicação na Gestão de Redes Netb@nd

Jorge Gonçalves, Ricardo Cadime, Alexandre Laranjeira, Manuel Aguiar, Vitor Mirones

palavras-chave:
Rede de Transporte, Modelo de Informação, Serviço de Transporte, OSS, EMS, MTNM, MTOSI, SID, XML, Neb@nd, AGORA-NG.

Actualmente as Redes de Transporte são baseadas numa diversidade de tecnologias (SDH, PDH, WDM, ATM, *Ethernet*, IP/MPLS, ...) que cobrem as três primeiras camadas do modelo OSI. Esta diversidade e a introdução de novas tecnologias veio complicar as tarefas de planeamento e operação da rede. Para gerirem de forma eficiente estas redes multi-tecnologia, os operadores de telecomunicações evoluiram os seus sistemas de suporte à operação (sistemas OSS) para que estes implementassem os processos operacionais de forma independente das tecnologias de rede. Esta arquitectura de sistemas OSS requer que exista uma camada de abstracção sobre a rede, construída com base em modelos genéricos de rede. Alguns fornecedores de equipamentos e soluções de redes de transporte também adaptaram os seus sistemas de gestão EMS para gerirem a partir da mesma plataforma as diferentes tecnologias de rede que disponibilizam. O interesse mútuo de operadores e fornecedores de soluções de rede levou a que organismos internacionais (ITU-T, TMForum,...) desenvolvessem trabalhos na modelação de Redes de Transporte. Aqui enquadram-se os modelos funcionais (G.805 e G.809) e os modelos de informação (M.3100, MTNM (TMF608) e SID), de referência nesta área. A PT Inovação, como fornecedora de soluções de rede de transporte multi-tecnologia (soluções Netb@nd), construiu um modelo de informação de rede genérico com base nas normas internacionais referidas e aplicou-o na plataforma de gestão AGORA-NG. Esta decisão tem permitido que todas as soluções Netb@nd sejam geridas por esta plataforma. Este artigo visa introduzir os modelos de informação de referência na gestão de Redes de Transporte multi-tecnologia e dar a conhecer como, a partir destes, a PT Inovação construiu o modelo de informação genérico utilizado no sistema AGORA-NG e nos agentes de gestão dos equipamentos. Na parte final do artigo é apresentado um caso de aplicação que consiste na modelação do serviço de transporte de *Ethernet* sobre SDH, utilizando o modelo de informação do AGORA-NG.

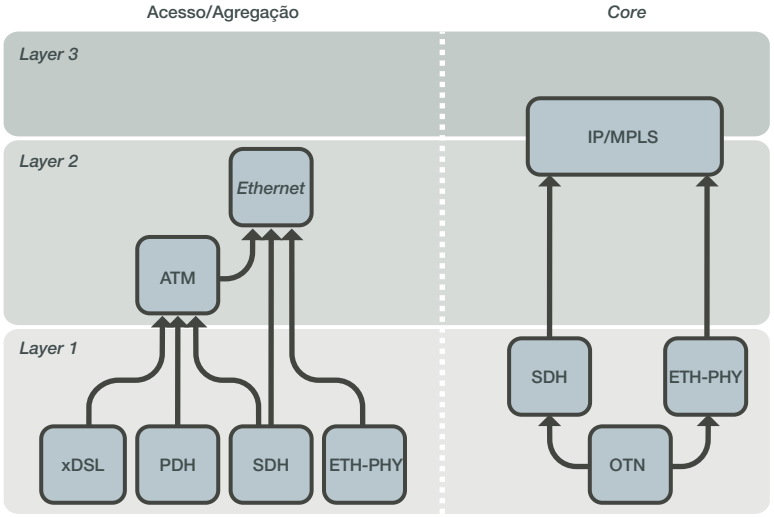

figura 1 - Relações entre Redes de Transporte

1. Redes de Transporte – Caracterização e Evolução

Genericamente, podemos caracterizar Rede de Transporte como sendo uma rede que permite transferir informação de utilizador e de controlo (sinalização e gestão) entre dois ou mais pontos de acesso, utilizando dois modos de ligação possíveis: orientado à ligação (*connection oriented*), tipicamente usado em redes de circuitos; não orientado à ligação (*connectionless*), usado em redes de pacotes.

Até meados da década de 90, o termo Redes de Transporte, era usualmente aplicado a redes de transmissão baseadas em tecnologias de nível 1 (SDH, PDH, OTN, ...). De facto, a verticalização das redes por serviço (voz, dados, televisão) e a tecnologia disponível, levavam a que a integração de serviços na mesma infraestrutura de transporte fosse realizada apenas na camada de transmissão.

No final da década de 90, a introdução da tecnologia ATM na rede dos operadores veio alargar o âmbito das Redes de Transporte ao nível 2. A técnica de multiplexagem assíncrona de pacotes de tamanho fixo, denominadas células, e os mecanismos de QoS da tecnologia, permitiram flexibilizar a utilização de largura de banda e criar um *overlay* sobre as redes de transmissão, com capacidade de transporte multi-serviço. No entanto, a tecnologia ATM mantém a principal característica das redes de circuitos, de operação em modo orientado à ligação.

Nos últimos anos, tem-se verificado uma mudança de paradigma nas Redes de Transporte, com a introdução das tecnologias baseadas em pacotes, IP/MPLS, no *core* da rede, e *Ethernet*, no acesso/agregação. De facto, a consolidação do IP como tecnologia de transporte multi-serviço (dados, voz, vídeo) *end-to-end* é um dado assumido, aliás reflectido na arquitectura das redes de nova geração, NGN[1] e IMS[2], que convergem na escolha desta tecnologia para a camada de transporte. Esta evolução estendeu o âmbito das Redes de Transporte ao nível 3 e a tecnologias que operam em modo não orientado à ligação.

Como ilustra a figura 1, temos actualmente um conjunto diferenciado de Redes de Transporte a operar nos primeiros três níveis do modelo OSI, existindo entre elas uma relação cliente/servidor, em que as redes dos níveis inferiores fornecem serviços de transporte às dos níveis superiores. Vemos que no acesso/agregação há uma predominância das redes baseadas em tecnologias de nível 2 (ATM e *Ethernet*), suportadas numa panóplia de tecnologias de transmissão (xDSL, *wireless*, SDH, PON, ...), enquanto no *core* encontramos a rede IP/MPLS, baseada em tecnologia de nível 3/2,5, suportada tipicamente em redes de transmissão SDH ou WDM.

2. Redes de Transporte – Modelação

Dada a complexidade e diversidade das Redes de Transporte é essencial existir um modelo genérico de rede, independente da tecnologia, que identifique as entidades funcionais e as relações entre elas, de modo a facilitar as tarefas de desenho

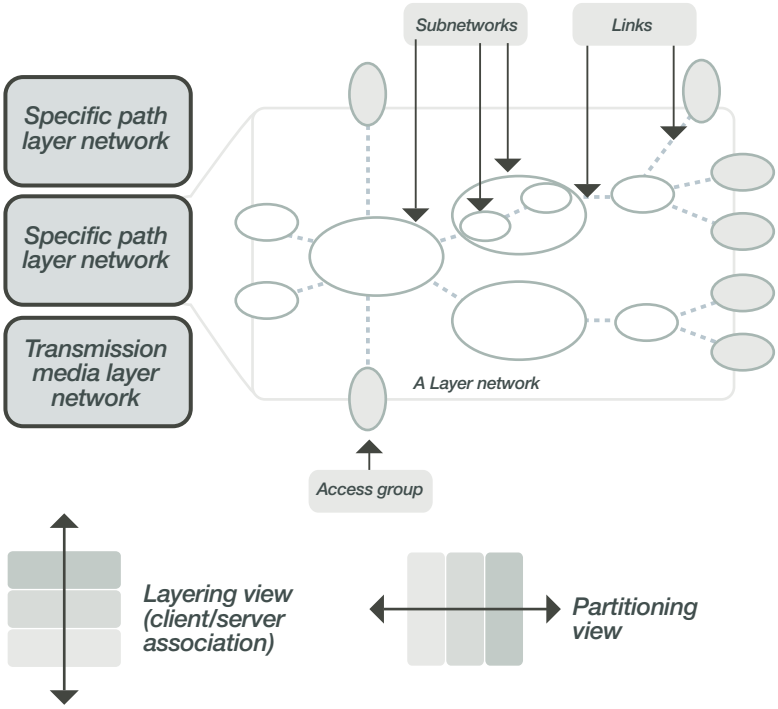

figura 2 - G.805[3] – *Layering* e *Partioning*

e gestão da rede.

Esta modelação é fundamental para os operadores e para os fornecedores de equipamentos/soluções de redes de transporte. Aos primeiros facilita-lhes a construção de uma camada de mediação sobre a rede, de forma a que a arquitectura de sistemas de operação (OSS) seja independente das tecnologias de rede e fornecedores equipamentos. Aos segundos, o modelo genérico permite-lhes criar sistemas de gestão (EMSs) multi-tecnologia, com capacidade de gerir todas as soluções de rede de transporte do seu *portfolio* a partir da mesma plataforma de gestão.

Modelo de Referência

A referência na modelação de Redes de Transporte, é a recomendação G.805[3], do ITU-T, que apresenta um modelo de rede genérico, assente nos conceitos de *layering* e *partitioning*, como ilustra a figura 2. Qualquer Rede de Transporte pode ser estruturada em camadas (*LayerNetworks*), que fornecem serviços de transporte entre si numa lógica cliente/servidor. A topologia de cada camada (*LayerNetwork*) é formada por: subredes (*SubNetworks*), que podem conter outras subredes; pontos de acesso (*AccessGroups*); e ligações (*Links*), que interligam subredes entre si e pontos de acesso a subredes. Os *Links* abstraem recursos, largura de banda e canais, e são suportados na *layer* servidora em *Trails* (ver figura 3).

Numa dada *LayerNetwork*, o serviço de transporte entre pontos de acesso, é modelizado pela entidade *Trail*.

O *Trail* transporta a informação adaptada da *layer* cliente de forma transparente, entre os *TrailTerminationPoints* (TTPs). Esta informação é monitorada *end-to-end* na função de terminação do *Trail*.

A G.805[3] define ainda a entidade de transporte *Connection*. Esta entidade modeliza o transporte da informação caractística da *Layer Network* entre *Connection Termination Points* (CTPs), de forma transparente e não monitorada. Dependendo do âmbito, a *Connection* pode ser especializada em *NetworkConnection, SubNetworkConnection* e *LinkConnection* (ver figura 3). Em resumo, as terminações do *Trail* geram a informação característica da *LayerNetwork*, que é transportada transparentemente através desta pela entidade *NetworkConnection*, que é resultado da concatenação de *Sub-NetworkConnections* e *LinkConnections*.

O modelo G.805[3] aplica-se bem a Redes de Transporte baseadas em tecnologias que operam em modo orientado à ligação. De facto,

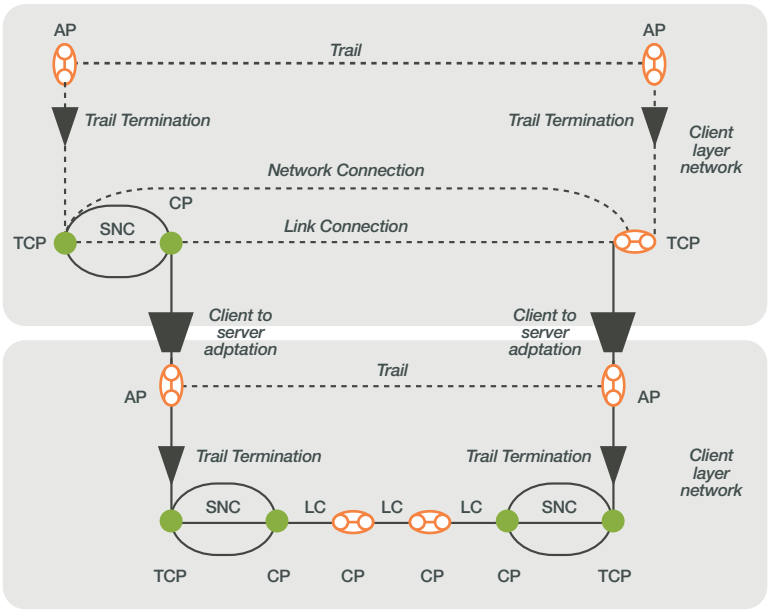

figura 3 - Modelo de Rede de Transporte G.805[3]

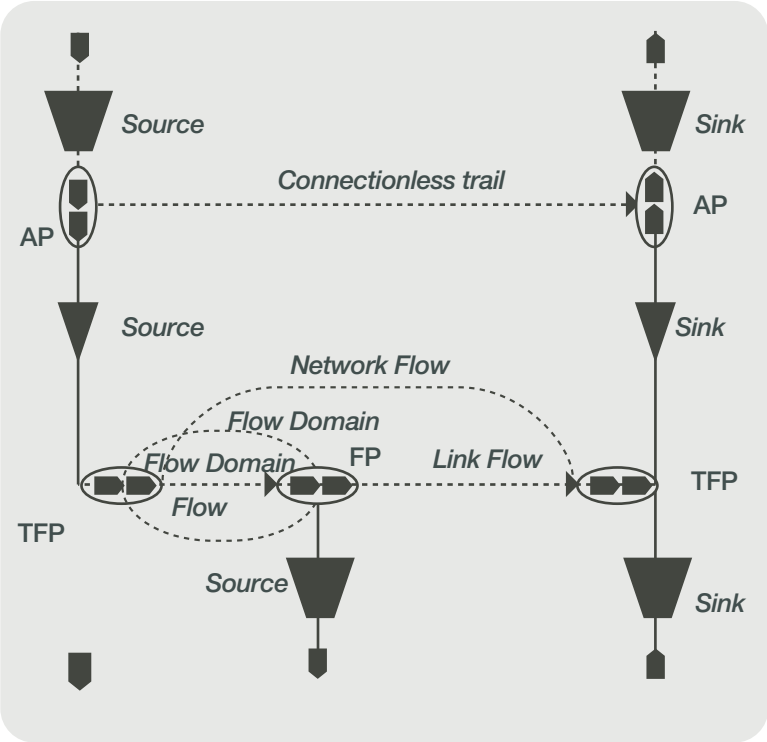

figura 4 - Modelo de Rede de Transporte *Connectionless* G.809[4]

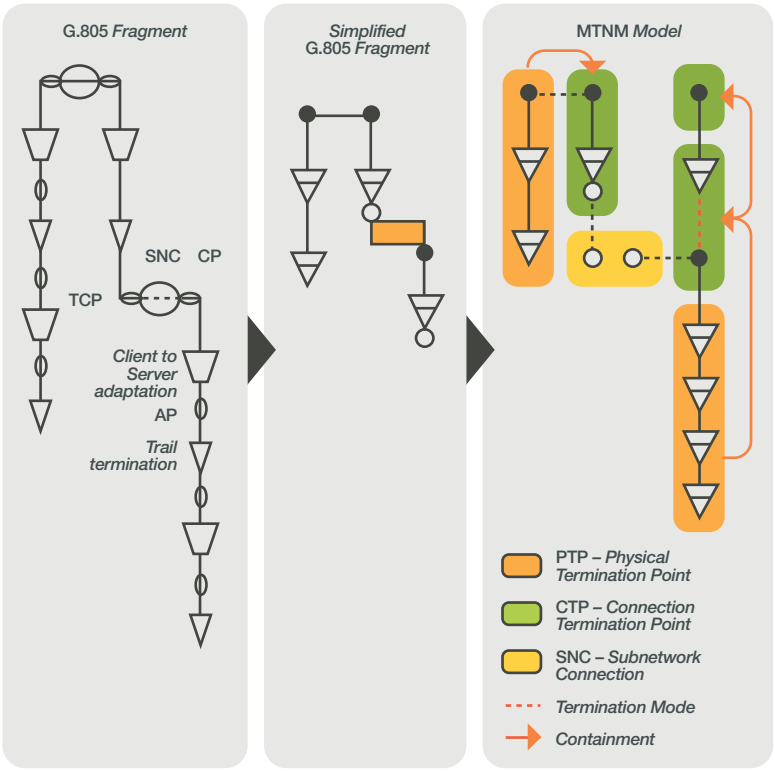

figura 5 - Modelo MTNM TMF608[7]

apesar de genérico, ele foi definido numa altura em que nas redes de transporte predominavam as tecnologias SDH e ATM e em que não era expectável que as tecnologias baseadas em pacotes viessem a entrar neste domínio. Assim, alguns conceitos, como por exemplo *Connection*, não têm muito sentido em redes que operam em modo não orientado à ligação. Deste modo, o ITU-T, na recomendação G.809[4] ajustou o modelo para contemplar este tipo de redes. Basicamente, o modelo G.809[4], como mostra a figura 4, substitui os conceitos de *Trail* por *ConnectionlessTrail*, de *Connection* por *Flow* e de *SubNetwork* por *Flow Domain*, que reflectem melhor o modo de funcionamento destas redes.

Modelos de Informação

O modelo de informação para gestão de redes de transporte, derivado directamente da G.805[3], foi definido na recomendação M.3100[5], do ITU-T. Este é um bom modelo de análise mas tem alguma ineficiência em termos de implementação, dado o elevado número de entidades a que ele conduz. A título de exemplo, se considerarmos a modelação de várias camadas do *layer* físico (ex: Regeneração e Multiplexagem num troço SDH), o número de entidades (*Trail, NetworkConnection*, TTP, CTP) total é multiplicado pelo número de camadas existentes, sem valor acrescentado substancial.

Levando este aspecto em consideração, o TMForum na definição da interface MTNM[6] (*Multi-Technology Network Managent*) entre a gestão de elemento (EML) e a gestão de rede (NML), criou um modelo simplificado da G.805[3] (TMF608[7]), utilizando conceitos de encapsulamento e *containment* (ver figura 5).

Basicamente, o modelo redefine as entidades do tipo *TerminationPoint*, utilizando três conceitos: *PhysicalTerminationPoint* (PTP), *Connec-*

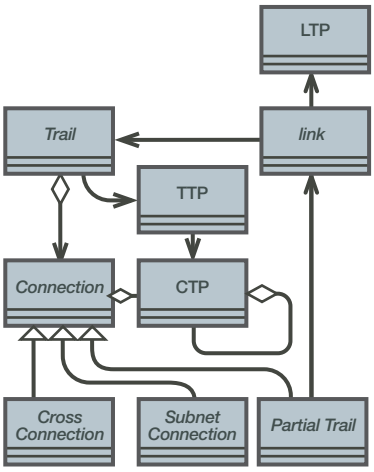

figura 6 - Modelo de rede AGORA-NG[11]

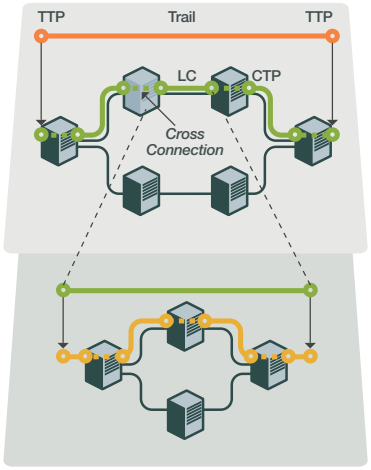

figura 7- Modelo de Rede AGORA-NG[11]

tionTerminationPoint (CTP) e *FloatingTerminationPoint* (FTP).

A entidade PTP representa um porto físico e encapsula a terminação de todas as camadas do *layer* 1, nesse porto. Utiliza atributos (*TransmissionParameters*) para descrever as características de cada camada. Deste modo reduz-se significativamente o número de entidades a instanciar.

A entidade FTP é uma entidade interna ao elemento de rede equivalente a um PTP, mas é virtual, isto é, não representa uma interface física. É usado, por exemplo, para modelizar conceitos de *Inverse Multiplexing* (IMA no ATM, VCG no EoSDH, etc.).

A entidade CTP, como vimos, representa a terminação de uma entidade lógica *Connection*. Se o CTP for a terminação de uma *NetworkConnection* está associado a um TTP. Neste caso, o modelo simplifica e permite que o CTP possa conter CTPs da camada cliente. Por exemplo, um CTP VC4 da camada SDH-HOP, que termina um VC4, pode conter CTPs VC12, da camada SDH-LOP.

De notar que o modelo MTNM[7] descreve as entidades de redes de transporte multi-tecnologia geridas por EMS, incluindo, para além das entidades topológicas e de transporte do nível de rede, as entidades físicas e lógicas que caracterizam o elemento de rede.

De acordo com as tendências mais recentes na arquitectura de sistemas OSS, definidas no programa NGOSS[8] do TMForum, o próprio TMForum evoluiu o conceito ponto a ponto (NML-EML) da interface MTNM[6], baseada em CORBA, para o conceito de *bus* de comunicação (OSS - OSS) da interface MTOSI[9], baseada em *WebServices/ XML*. O objectivo é que os OSS de empresa (*Inventory, Alarm, Activation, Performance*) possam comunicar com os recursos de rede através desta interface de mediação. De salientar que o modelo de informação TMF608[7] é seguido em ambas as interfaces, MTNM[6] e MTOSI[9]. Avaliando pelos trabalhos em curso no TMForum de compatibilização do SID[10] (modelo comum de informação da arquitectura NGOSS) com o MTNM[6], é expectável que a especialização do SID[10] para as redes de transporte venha a ser baseada no modelo MTNM (TMF 608[7]).

3. Modelo de Informação para Gestão de Redes Netb@nd

Tendo por base os modelos de referência acima identificados, este capítulo descreve o modelo de informação usado pelo sistema de gestão AGORA-NG[11], na gestão de soluções de rede Netb@nd, da PT Inovação. Estas são essencialmente redes de transporte, que cobrem uma gama variada de tecnologias de transporte, como sejam: PDH, n*64k, xDSL, SDH, ATM e *Ethernet*.

A utilização dos modelos de referência G.805[3], G.809[4] e MTNM(TMF608[7]), na definição do modelo de informação do AGORA-NG, tem como objectivo garantir que o modelo final é suficiente-mente genérico para que o AGORA-NG tenha capacidade para gerir todas as soluções de rede de transporte do *portfolio* PT Inovação.

O modelo de informação é composto pelo modelo genérico de rede e pelo modelo genérico de elemento de rede. O primeiro existe apenas no contexto do sistema de gestão, enquanto o segundo, como veremos no ponto 3.2, é também utilizado nos agentes de gestão dos equipamentos.

3.1 Gestão de Redes Netb@nd – Modelo Genérico de Rede

O modelo genérico de rede utilizado no AGORA-NG[11] está ilustrado na figura 6.

O modelo permite construir topologias de rede muti-tecnologia e descrever serviços de transporte *end-to-end* sobre essas redes.

A topologia é construída com base nas entidades *SubNetwork*, *Link* e CTP (quando o CTP é terminação de rede). Estas entidades estão associadas a uma dada *LayerNetwork*, que identifica uma camada numa dada rede de transporte (ex: HOP *layer* na rede SDH). A informação topológica é utilizada para efeitos de encaminhamento automático de *Trails*.

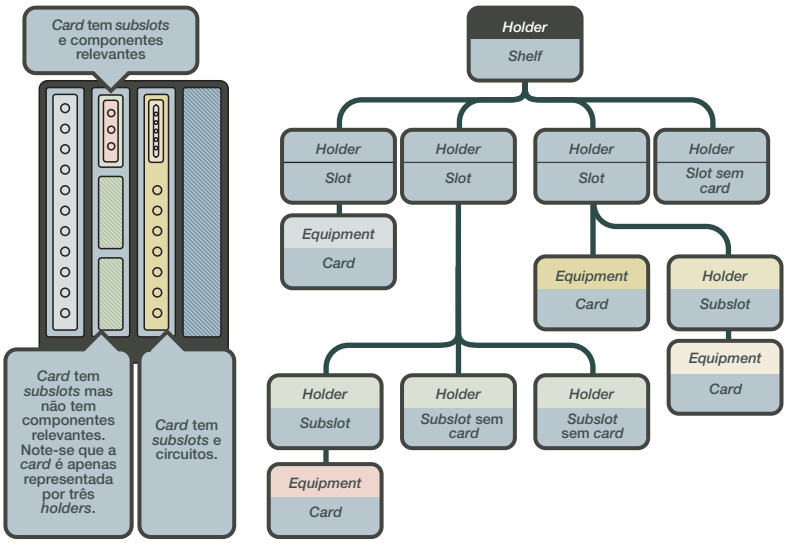

figura 8 - Modelo Físico

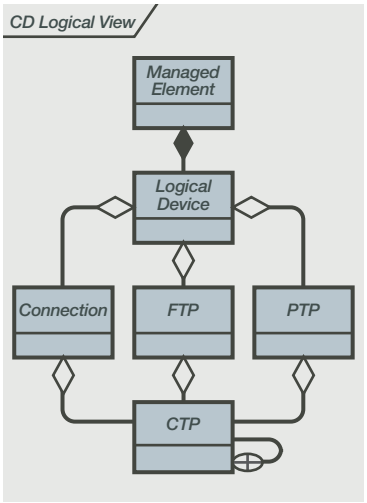

figura 9 - Modelo Lógico

O *Trail* é a entidade que representa o serviço de transporte *end-to-end* (entre TTPs). A terminação do *Trail* (TTP) interliga-o ao CTP ponto de terminação de rede. O percurso do *Trail* é composto por *Connections* (*Crossconnections*, *SubNetwork-Connections* e *LinkConnections*) que lhe garantem a continuidade do fluxo de tráfego ao longo da rede.

A figura 7 ilustra os conceitos implementados no modelo de rede AGORA-NG[11].

3.2 Gestão de Redes Netb@nd – Modelo Genérico Elemento de Rede

O modelo de elemento de rede caracteriza as entidades físicas e lógicas de um equipamento. A adopção de modelos genéricos permite uniformizar a visão do elemento de rede e abstrair as especificidades inerentes aos domínios tecnológicos geridos.

O modelo utilizado no AGORA-NG[11] e nos agentes de equipamentos Netb@nd baseia-se nos modelos MTNM[7] e SID[10].

De salientar a importância da utilização do mesmo modelo no sistema de gestão e nos agentes de gestão. Este ponto facilita a integração de novos equipamentos e tecnologias e reduzir o tempo de desenvolvimento.

O modelo genérico de elemento de rede consiste no modelo físico (figura 8) e no modelo lógico (figura 9).

Modelo Físico

O modelo físico contempla as entidades físicas palpáveis - *rack*, *shelf*, *chassis*, *board*, *physicalport* – destacando-se as relações de *containment* entre as entidades, como ilustra a figura 8.

Os portos físicos, podendo estar contidos em placas ou directamente em *chassis*, são os pontos de acesso ao equipamento. As relações de *containement* facilitam a gestão de inventário das unidades atómicas existentes. Estas relações proporcionam ainda uma correlação entre elementos (um porto de uma placa não pode existir sem a placa e perde o significado quando esta é removida).

Modelo lógico

O Modelo lógico está relacionado com os aspectos e características lógicas associadas à "inteligência" do elemento de rede. A figura 9 apresenta as entidades lógicas que existem no contexto de um elemento de rede.

Da figura destacam-se dois tipos de entidades: as terminações (PTP, FTP e CTP), já descritas anteriormente; e as *Connections* que definem fluxos de tráfego entre terminações. As primeiras sustentam o *layering* enquanto que as segundas sustentam o *partitioning* do modelo. As entidades PTP fazem a ponte entre a parte lógica e física do modelo do equipamento.

4. Caso de Aplicação – Transporte de EoSDH

O objectivo do caso de aplicação é demonstrar a aplicabilidade do modelo genérico de rede do AGORA-NG[11]/Equipamento a um caso concreto de um serviço de transporte. Dada a actualidade da entrada da tecnologia *Ethernet* na rede dos operadores foi escolhido o serviço de transporte de EoSDH.

A solução da PT Inovação nesta área é baseada nos sistemas EMILO-NG[12], que disponibilizam interfaces tributárias *ethernet* FE/GBE e interfaces agregadas STMx. Para o caso de aplicação foi seleccionado o equipamento EMILO-S14E[12]. Este inclui um *switch ethernet* (placa UNICOM-SE) e uma unidade de encapsulamento e concatenação, assegurado pela entidade VCG na placa T8ETH4VC4.

Actualmente a solução EMILO-NG[12] permite disponibilizar os serviços: *Ethernet Private Line* [13](EPL) e *Ethernet Virtual Private Line* [14](EVPL). O EPL[13] consiste no transporte de tráfego *ethernet* entre duas interfaces UNI, de forma dedi-

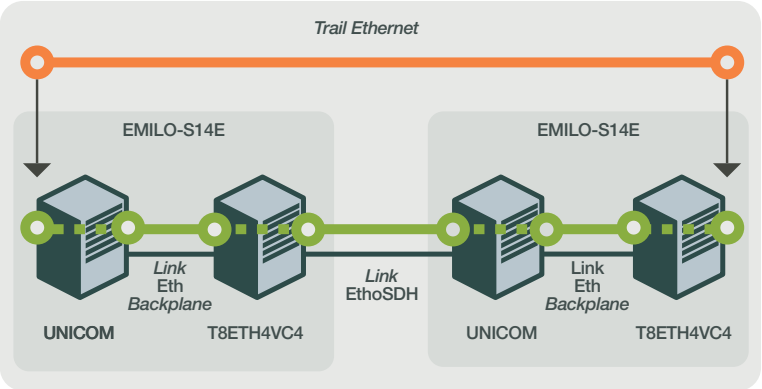

figura 10 - *Trail* EVPL na *LayerNetwork Ethernet*

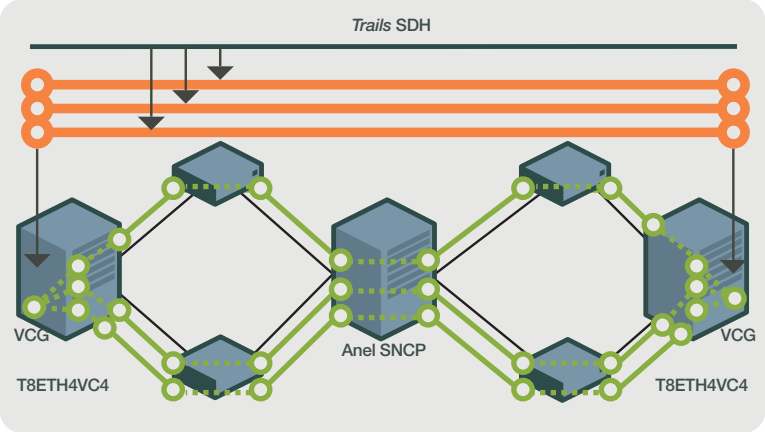

figura 11- Link *Ethernet* suportado em *Trails* VC12

cada. O EVPL[14] consiste no transporte de VLANs entre duas interfaces UNI, de forma dedicada ou partilhada.

Para o caso de aplicação foi escolhido o cenário mais complexo, que consiste no serviço EVPL[14]. Pretende-se ilustrar as entidades envolvidas na modelação do serviço de transporte de VLAN *end-to-end*, através da rede SDH.

No AGORA-NG[11] este serviço é modelizado por um *Trail* EVPL na *LayerNetwork* EoSDH. Esta *layer* é constituída por subredes EMILO-S14E[12]interligadas por *Links Ethernet*. Os pontos de acesso na *LayerNetwork* são os CTPs VLAN associados às interfaces físicas *Ethernet* FE/GbE dos EMILO-S14E[12]. A subrede EMILO-S14E[12] é ainda composta internamente por duas subredes atómicas, a UNICOM-SE e a T8ETH4VC4 interligadas por um *Link Ethernet* (ligação física no *backplane*).

O *Trail* EVPL é então formado pelos TTPs que ocupam os CTPs VLAN situados na extremidade da *LayerNetwork*, pelas *CrossConnections EVC* nas subredes atómicas e ainda pelas *LinkConnections* nos *Links Ethernet*.

O *Link Ethernet*, que interliga as subredes EMILO-S14E[11], é suportado em *Trails* (*Trails* VC12, no exemplo) na *layer* servidora LOP-SDH, como ilustra a figura 11. A adaptação entre camadas é efectuada pela entidade VCG (FTP que implementa a função de *inverse multiplexing*).

Os *Trails* VC12 são encaminhados na *LayerNetwork* SDH-LOP ocupando recursos nas subredes e nos *Links* VC4.

Nos elementos de rede EMILO-S14E[12] os recursos ocupados pelo *Trail* EVPL estão sinalizados na figura 12.

Estão identificadas as *CrossConnections* EVC efectuadas na UNICOM-SE e na T8ETH4VC4. Na primeira existe flexibilidade na comutação, dado que a UNICOM-SE é um *switch ethernet*, enquanto que na segunda o mapeamento é fixo.

Estão igualmente identificadas as *CrossConnections* VC12 efectuadas na T8ETH4VC4 e na matriz AGSTMx, relativas à componente SDH. Na T8ETH4VC4 o mapeamento é fixo.

Por fim, está sinalizada a entidade FTP VCG de adpatação entre a camada *Ethernet* e a camada SDH.

5. Conclusões

A gestão de Redes de Transporte multi-tecnologia requer que os modelos de informação utilizados sejam suficientemente genéricos para

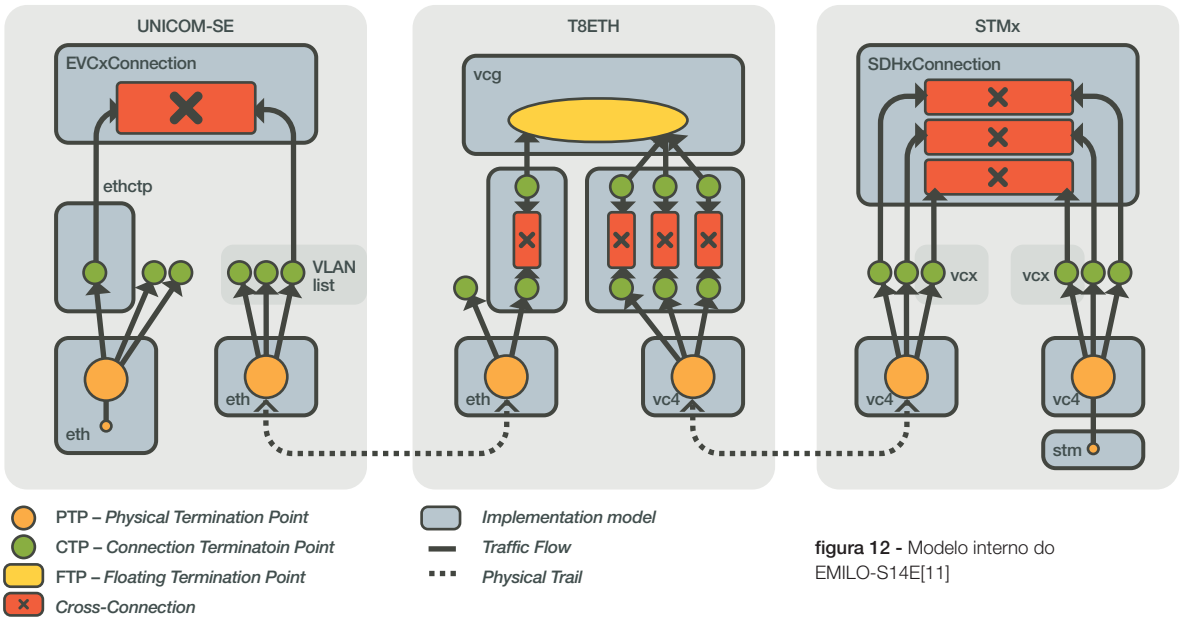

figura 12 - Modelo interno do EMILO-S14E[11]

abstrair as especificidades de âmbito tecnológico da rede.

O AGORA-NG[11], sistema de gestão das soluções de rede Netb@nd, utiliza modelos de rede e de elemento de rede genéricos, baseados nas normas internacionais de referência na modelação de Redes de Transporte (G.805[3], G.809[4], M.3100[5], MTNM (TMF608[7]), SID[10]).

O modelo de informação do AGORA-NG[10] deu à PT Inovação a vantagem competitiva de poder gerir todas as tecnologias de rede de transporte Netb@nd a partir da mesma plataforma de gestão, minimizando custos de investimento e de operação.

Referência

[1] NGN - Next Generation Networks http://www.etsi.org/tispan/
[2] IMS - IP Multimedia Subsystem http://www.3gpp.org
[3] ITU-T G.805 - Generic Functional Architectures of Transport Networks
[4] ITU-T G.809 - Functional Architecture of Connectionless Layer Networks
[5] ITU-T M.3100 - Generic Network Information Model
[6] MTNM - Multi-Technology Network Management Interface; http://www.tmforum.org
[7] Modelo de Informação TMF608 http://www.tmforum.org
[8] NGOSS - Next Generation Operation Support Systems; http://www.tmforum.org
[9] MTOSI Multi-Technology Operation System Interface http://www.tmforum.org
[10] SID - Shared Information Data http://www.tmforum.org
[11] AGORA-NG http://www.ptinovacao.pt
[12] EMILO-NG http://www.ptinovacao.pt
[13] ITU-T G.8011.1- Ethernet Private Line
[14] ITU-T G.8011.2 - Ethernet Virtual Private Line

Curriculum Vitae

Jorge Gonçalves, licenciou-se em Engenharia Electrónica e Telecomunicações, pela Universidade de Aveiro, em 1988, e obteve o grau de mestre pela mesma Universidade, em 1997. Em 1988, começou a trabalhar no Centro de Estudos de Telecomunicações, hoje PT Inovação, na área de comutação digital. Desde 1991 tem participado em projectos europeus (RACE, ACTS, IST e Eurescom), em áreas relacionadas com redes ATM, redes de acesso IP e gestão de redes. Em 1999, assumiu a responsabilidade da Unidade Gestão de Redes de Acesso, sendo presentemente o responsável pela Unidade Gestão de Soluções Netb@nd .

Ricardo Cadime, licenciou-se em Engenharia Electrónica e Telecomunicações, pela Universidade de Aveiro, em 1997, e obteve o grau de mestre pela mesma Universidade, em 2001. Em 2000, começou a trabalhar na PT Inovação, na área de gestão de redes. Participou em projectos europeus (IST e Eurescom), em áreas relacionadas com redes IP (QoS) e gestão de redes. Faz parte actualmente do grupo de Gestão de Redes no SIR/PTIN, sendo responsável pelo projecto AGORA-NG *Assure Pack*.

Alexandre Laranjeira, licenciou-se em Engenharia Electrónica e Telecomunicações, pela Universidade de Aveiro, em 2000. Em 2000, começou a trabalhar na PT Inovação, num projecto do IST sobre gestão de redes IP/WDM. Desde 2001 trabalha no grupo de gestão de soluções Netb@nd, sendo actualmente responsável pelo projecto AGORA-NG *Provision Pack*.

Manuel Aguiar, licenciou-se em Enganharia Electrónica e Telecomunicações pela Universidade de Aveiro em 1994. Ingressou no C.E.T. em 1997 aonde desenvolveu trabalhos relacionados com Sistemas de Gestão para a Rede de Acesso. Integrou a equipa de de Gestão de Soluções Netb@nd tendo sido responsável pelas aplicações *Provision Pack* da plataforma AGORA-NG. Actualmente é o responsável pela unidade de Gestão Local e Qualidade de Serviço da Área de "Sistemas e Infraestruturas de Rede SIR", da PT Inovação.

Vitor Mirones, licenciado em Engenharia Electrotécnica e Computadores, ramo de Telecomunicações, pela faculdade de Engenharida da Universidade do Porto (1994). No CET desde 1995 onde colaborou na área das Tecnologias Mutimédia e Desenvolvimento de Serviços e Aplicações. Actualmente na área de "Sistemas e Infraestruturas de Rede - SIR" da PT Inovação participa no desenvolvimento das componentes de gestão de equipamentos Netb@nd .

Colaboradores do Projecto

Para a implementação deste trabalho foi imprescindível a colaboração das seguintes pessoas, para além dos autores do artigo: Annette Tigeleiro, Sérgio Soares, Manuel Florêncio, Alberto Rocha, Fernando Grangeia, Nuno Guimarães (Withus).