Universidade do Minho
Escola de Engenharia

Helder Marco da Mota Pereira

Aplicação de Políticas em Redes 3GPP

Outubro de 2009

Helder Marco da Mota Pereira **Aplicação de Políticas em Redes 3GPP**

UMinho | 2009

Universidade do Minho
Escola de Engenharia

Helder Marco da Mota Pereira

Aplicação de Políticas em Redes 3GPP

Mestrado de Engenharia Informática

Trabalho efectuado sob a orientação do
Professor Doutor Paulo Manuel Martins de Carvalho
e do
Engenheiro Jorge Miguel Marques Dias e Sousa

Outubro de 2009

Agradecimentos

Em primeiro lugar, as minhas palavras de apreço vão para os meus pais. Em segundo lugar, gostaria de agradecer ao meu orientador, o Professor Doutor Paulo Manuel Martins de Carvalho, pela paciência e dedicação demonstradas para com o meu trabalho. Para ele, o meu muito obrigado. Uma palavra de agradecimento ao Eng. Jorge Sousa, que me acompanhou na PT Inovação durante o decorrer do meu trabalho, pela disponibilidade e prontidão com que me foi apoiando. Um agradecimento particular também ao Eng. Raul Rodrigues da PT Inovação, que teve um papel muito importante para a realização do trabalho conducente a esta dissertação.

Também gostaria de agradecer a todos os elementos que constituem a equipa de Redes de Dados do Departamento de Redes e Protocolos da PT Inovação, que de forma directa ou indirecta contribuíram para o sucesso deste trabalho.

Resumo

O aumento substancial da utilização da banda larga móvel tem representado um desafio cada vez maior para os operadores de telecomunicações. Este crescimento tem feito transparecer a crescente importância de existir um controlo e policiamento de tráfego dos utilizadores eficaz ao nível das redes móveis, bem como a capacidade de fornecer um serviço diferenciado e à medida, consoante os perfis ou necessidades dos utilizadores. A solução para este novo desafio que se apresenta, não pode passar unicamente pela instalação constante de maior largura de banda. Os recursos são finitos e a instalação de cada vez mais largura de banda, sem haver um controlo significativo do crescimento da utilização da infra-estrutura de rede, atingirá eventualmente um limite.

Soluções de controlo e policiamento de tráfego começam a tomar um papel fundamental nas redes dos operadores. Estas soluções permitem aos operadores garantir um bom comportamento da rede e, simultaneamente, garantir aos seus clientes uma determinada qualidade de serviço. Permitem igualmente vislumbrar um modelo de negócio moderno, onde haverá a venda de garantias de qualidade de serviço ao invés do modelo tradicional, onde aquilo que se vende é a largura de banda.

Esta dissertação propõe uma solução de controlo de tráfego baseado em políticas, com fortes bases nas directivas definidas pelo organismo de normalização internacional 3GPP no que concerne ao controlo de tráfego em redes de telecomunicações. Nesse sentido foi desenvolvida uma solução capaz de controlar equipamentos de controlo de políticas de rede que sigam as linhas orientadoras da norma definida pelo organismo 3GPP. Esta solução é capaz de fazer controlo de tráfego baseado no perfil do utilizador, permitindo desta forma a oferta de serviços diferenciados, enquanto paralelamente introduz mecanismos para o controlo equilibrado da infra-estrutura de rede e respectiva qualidade de serviço. Com o intuito de testar a solução desenvolvida foram realizados testes que demonstram claramente que a solução está habilitada a introduzir controlo de Qualidade de Serviço em infra-estruturas de rede que suportem esta funcionalidade. A forma utilizada para atribuir perfis de utilização aos diversos utilizadores garante também que a solução é suficientemente flexível para assegurar diferentes modelos de negócio.

Abstract

The substantial increase in the use of mobile broadband represents a continuous challenge to telecom operators. This growth has revealed the increasing importance of having good control and policing of user's traffic at the mobile network level, as well as providing personalized and differentiated services according to the profiles and needs of the users. The solution to this new challenge cannot rely on the continuous installation of additional network bandwidth. Resources are limited and the provisioning of higher bandwidth capacity without a significative control over the network infrastructure usage will eventually reach an upper limit.

Traffic control and policing solutions are assuming an importante role within operator networks. Those solutions enable operators to ensure good network behaviour, while guaranteeing a certain quality of service to their customers. They also allow operators to envision a modern business model, where the sale of service quality will superseed the traditional model, where what is sold is bandwidth quantity.

This thesis proposes a traffic control solution based on network policies, strongly supported by the directives defined by the international standards organization 3GPP in what relates to traffic control in telecommunication networks. In this way, a solution capable of controlling network equipment responsible for policy configuration has been implemented according to 3GPP guidelines. This solution will be able to control traffic based on the user profile, thus enabling provision of differentiated services while, at the same time, introducing mechanisms for a balanced control of the network infrastructure and its quality of service. The tests carried out to evaluate the developed solution show that the solution is able to provide proper QoS control in network infrastructures that support this funtionality. The method used to assign service profiles to users also guarantees that the solution is flexible enough to allow different business models.

Conteúdo

Lista de Figuras

Lista de Tabelas

Lista de Acrónimos

3GPP	3rd Generation Partnership Project
A-RACF	Access-Resource Admission Control Function
AAA	Authentication, Authorization and Accounting
ADSL	Assymetric Digital Subscriber Line
AF	Application Function
ATM	Asynchronous Transfer Mode
AVP	Attribute-Value Pair
BD	Base de Dados
BGF	Border Gateway Function
CAC	Call Admission Control
CCA	Credit Control Answer
CCR	Credit Control Request
CEO	Chief Executive Officer
COPS	Common Open Policy Service
CRF	Charging Rules Function
DSCF	Data Session Control Function
DSCP	Data Service Control Point
DSGW	Diameter Signaling Gateway
ETSI	European Telecommunications Standards Institute

GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
GRE	Generic Rules Engine
HSDPA	High-Speed Downlink Packet Access
HSPA	High Speed Packet Access
IETF	Internet Engineering Task Force
IMS	IP Multimedia Subsystem
IN	Intelligent Network
IP	Internet Protocol
IP-CAN	IP-Connectivity Access Network
MSISDN	Mobile Subscriber Integrated Services Digital Network Number
NGN	Next Generation Networks
OCS	Online Charging System
OFCS	Offline Charging System
P-CSCF	Proxy Call Session Control Function
PACF	Policy and Admission Control Function
PBNM	Policy Based Network Management
PCC	Policy and Charging Control
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Rules Function
PDA	Personal Digital Assistant
PDF	Policy Decision Function
PDP	Packet Data Protocol
QoE	Quality of Experience
QoS	Quality of Service
RACS	Resource Admission Control Sub-System

RADIUS	Remote Authentication Dial In User Service
RAA	Reauthentication Answer
RAR	Reauthentication Request
RCEF	Resource Control Enforcement Function
RTDAP	Real-Time Data Application Part
SCE	Service Control Engine
SDP	Service Data Point
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SLR	Service Location Register
SLS	Service Level Specification
SM	Subscriber Manager
SNMP	Simple Network Management Protocol
SPDF	Service Policy Decision Function
SPR	Subscriber Profile Repository
TISPAN	Telecommunications and Internet Converged Services and Protocols for Advanced Networking
TLS	Transport Layer Security
TPF	Traffic Plane Function
UMTS	Universal Mobile Telecommunications System
VoD	Video on Demand
VoIP	Voice over IP

Capítulo 1

Introdução

Nos dias de hoje tem-se assistido a uma enorme expansão geográfica no que concerne ao acesso à Internet que naturalmente traz consigo um aumento do número de utilizadores, nomeadamente através do surgimento de tecnologias de acesso de banda larga diversas (ex: ADSL[1], WiFi, WiMAX, 3G, HSPA[2]) acompanhadas igualmente por um crescente número de dispositivos capazes de utilizar essas mesmas tecnologias (computadores de secretária, computadores portáteis, PDAs[3], telefones móveis, leitores MP3, ...) levando a um aumento natural no consumo de largura de banda. Perante este cenário, os operadores de telecomunicações têm sido confrontados com a necessidade de lidar com a crescente utilização do acesso à Internet, crescimento esse que poderá provocar congestionamento nas ligações de acesso à Internet.

Para o aumento exponencial de utilização de largura de banda por parte dos utilizadores têm contribuído diversos factores adicionais[1], de que são exemplo: (i) o uso inadvertido de software suspeito tal como spyware ou spam; (ii) a utilização excessiva de tráfego *peer-to-peer* (p2p) através de aplicações como o BitTorrent ou mesmo Emule; (iii) o crescimento da utilização de streaming áudio (Skype, Google Talk, Last.FM,); ou (iv) os serviços de vídeo (YouTube, Vídeo-On-Demand). Como tal, as redes de comunicações dos respectivos operadores tendem a operar perto de níveis de utilização que se poderão traduzir numa degradação indiferenciada do serviço fornecido para os utilizadores que dele usufruem.

Este nível de acesso à rede é também possível, em parte, devido à implementação de uma tarifa única de acesso. Tarifas de acesso fixas, diárias ou mensais, com ou sem limite de *downloads*

[1] Assymetric Digital Subscriber Line
[2] High Speed Packet Access
[3] Personal Digital Assistant

imposto, contribuíram para a existência desta situação, em que existe uma utilização desequilibrada dos recursos da rede. Isto representa um problema para todos os operadores, independentemente da tecnologia de acesso que utilizam. Por estes motivos, torna-se fundamental que um operador seja capaz de oferecer um serviço diferenciado aos seus utilizadores, baseado nas necessidades e nos perfis de cada um. Os operadores móveis têm vindo a fazer enormes investimentos na sua infra-estrutura no sentido de providenciarem acessos de banda larga em grande escala com tecnologias como o ADSL, UMTS[4] ou HSDPA[5], de forma a providenciarem um melhor serviço aos seus utilizadores. No entanto, o rápido crescimento no número de utilizadores e um fraco controlo da utilização dos recursos, conduzem a um uso desmedido da rede e, consequentemente, a uma fraca qualidade do serviço, contribuindo para um desgaste da imagem do operador, tornando o retorno do investimento difícil. Os operadores tentaram solucionar estes problemas, limitando a largura de banda ou introduzindo políticas universais de utilização aceitáveis. Como resultado destas medidas, todos os clientes recebem o mesmo tratamento, mesmo que não exagerem na utilização do serviço.

Com um mercado extremamente competitivo, os clientes exigem qualidade de serviço e tratamento diferenciado, bem como inovação. Os utilizadores não querem um serviço degradado como resultado de um pequeno número de utilizadores abusivos. A Internet é hoje um veículo fundamental para qualquer negócio, e como qualquer outro meio, o utilizador tem o direito de exigir garantias de qualidade de serviço face ao contracto estabelecido[2].

Neste contexto, o controlo de tráfego baseado no cumprimento de políticas - *Policy Enforcement* - dará um valioso contributo no sentido de oferecer melhores garantias de serviço ao utilizador. Consequentemente, decisões em tempo real de *Policy Enforcement* levam à satisfação do cliente, através da oferta de um serviço à medida, proporcionando garantias de qualidade de serviço. Ao operador ou fornecedor de serviço permite-lhe uma gestão mais racionalizada e sustentada da própria infra-estrutura de rede, abrindo também oportunidades para novos modelos de negócio.

1.1 Motivação e Objectivos

Os primeiros equipamentos com capacidade para realizar funções de *Policy Enforcement*, começaram por utilizar interfaces proprietárias para receberem a informação de quais as políticas a

[4]Universal Mobile Telecommunications System
[5]High-Speed Downlink Packet Access

aplicar a um utilizador. Naturalmente, os produtos que controlavam esses equipamentos estavam limitados também pelo facto de que, caso quisessem comunicar com outros equipamentos de outros fabricantes, teriam que implementar a interface de comunicações com esses equipamentos. A arquitectura *Policy and Charging Control* do organismo 3GPP vem tentar colmatar esta falha. Através da especificação de um interface de comunicação normalizado com o equipamento responsável pela aplicação de políticas na rede, assim como a normalização do seu comportamento, a entidade responsável pelo controlo desse equipamento não mais estará presa a um único fabricante. Desta forma esta entidade poderá controlar equipamentos de diversos fabricantes, equipamentos esses responsáveis pela aplicação de políticas numa rede. A arquitectura 3GPP, através da interface Gx, implementa uma interface normalizada para controlo do equipamento responsável pela aplicação de políticas na rede. Esta dissertação pretende, precisamente, dar resposta a esta necessidade.

Cada vez mais é necessário que os operadores controlem eficazmente os recursos da sua rede, garantindo o bom funcionamento da mesma e potenciando uma maior satisfação dos clientes. Como tal, a construção de um componente capaz de interagir com um equipamento responsável pela aplicação de políticas numa rede, seja de que fabricante for e desde que implemente a interface Gx do organismo 3GPP, irá trazer, indubitavelmente, uma mais valia aos operadores, potenciando novos modelos de negócio, através da oferta de serviços diferenciados. Esta solução permitirá agrupar os clientes por perfis de utilizador, garantindo que o tráfego de cada utilizador recebe um tratamento de acordo com o seu perfil. Naturalmente, será possível priorizar o tratamento dado a um determinado utilizador desde que o seu perfil assim o determine. A utilização de uma solução de *Policy Enforcement* ajudará a garantir igualmente o bom funcionamento da rede, ajudando a que os clientes tenham acesso ao serviço contratado e que picos de utilização repentinos possam ser melhor controlados e amortizados.

Nesse sentido, pretende-se construir um componente capaz de controlar os diversos equipamentos de *Policy Enforcement* que, de momento, utilizem a interface Gx do organismo 3GPP. Este componente será capaz de aplicar políticas aos utilizadores consoante o perfil que lhes é atribuído. No futuro, este componente poderá fazer a ligação com a camada de serviço dos operadores, permitindo que as diversas aplicações possam comunicar requisitos de QoS específicos. Existe uma natural aproximação às especificações delineadas pelo organismo 3GPP no que concerne à comunicação com equipamentos de *Policy Enforcement* uma vez que esta parece ser a tendência seguida pelos diversos fabricantes de equipamentos deste género.

1.2 Sumário das principais contribuições

A principal contribuição resultante do trabalho desenvolvido traduz-se na proposta e implementação de uma solução de policiamento de tráfego que segue de perto as linhas definidas pelo organismo 3GPP para controlo de tráfego. É utilizado um componente já existente na PT Inovação, o PACF, para o qual foi desenvolvido um plugin, assim como outras estruturas de apoio, que permitem implementar uma solução para controlo de tráfego e gestão de QoS. Esta solução consegue diferenciar cada utilizador que inicia sessão na rede, agrupando os utilizadores por perfil de utilização. Desta forma, é conseguida uma maior aproximação dos modelos de negócio dos operadores aos próprios mecanismos de configuração da rede, fazendo com que a modificação nas configurações dos equipamentos de rede seja motivada pelas regras de negócio definidas pelos operadores.

1.3 Organização da dissertação

No presente capítulo fez-se uma introdução à necessidade de introduzir medidas de controlo e gestão de recursos nas redes dos operadores de comunicações. Foi também explicada qual a motivação e os objectivos que se pretendem atingir com o presente trabalho.

No capítulo 2 será feita uma abordagem ao Controlo de Tráfego Baseado em Políticas, o PBNM[6], que importância é que o PBNM assume nas redes de nova geração, as denominadas NGN[7], e de que forma a filosofia PBNM foi adoptada por alguns organismos de normalização internacional, nomeadamente o IETF[8][3] através da sua arquitectura de *policy* genérica, o ETSI-TISPAN[9][4] através do RACS[10] e o 3GPP[5] através da arquitectura PCC[11]. Neste capítulo, será feita uma análise mais pormenorizada à arquitectura PCC do 3GPP, dado que esta arquitectura é o principal alvo de estudo desta dissertação.

No capítulo 3 será feita uma apresentação da solução actual de PE[12] na PT Inovação e serão demonstrados os passos seguidos para a implementação de uma nova solução de policy adaptável

[6] Policy-based Network Management
[7] Next Generation Networks
[8] Internet Engineering Task Force
[9] European Telecommunications Standards Institute - Telecommunications and Internet converged Services and Protocols for Advanced Networking
[10] Resource Admission Control Sub-System
[11] Policy and Charging Control
[12] Policy Enforcement

e escalável, que parte de um *policy server* genérico e que o especializará, através da construção de um plugin e de respectivas políticas de suporte, no componente funcional PCRF[13] da arquitectura PCC do 3GPP.

No capítulo 4 serão apresentados alguns cenários de utilização da nova implementação, assim como serão mostrados quais os testes efectuados para validar a implementação da nova solução de *policy*.

Por fim, no capítulo 5 serão apresentadas as conclusões desta dissertação, assim como as respectivas contribuições prestadas e possível trabalho futuro a desenvolver de forma a tornar a solução ainda mais rica em funcionalidades.

[13]Policy and Charging Rules Function

Capítulo 2

Controlo de tráfego baseado em políticas

Neste capítulo será feita uma apresentação geral do Controlo de Tráfego Baseado em Políticas (conhecido como PBNM[1]), a sua evolução, desde o seu início menos conseguido, até aos nossos dias, onde o PBNM é claramente exposto como uma arquitectura normalizada. Irá igualmente ser apresentada a arquitectura geral de *Policy* preconizada pelo IETF[2] e de que forma essa arquitectura contribuiu para o desenvolvimento e especificação de uma arquitectura de *Policy* pelo organismo de normalização internacional ETSI-TISPAN[3] e também pelo 3GPP[4]. Dar-se-á especial atenção à abordagem seguida pelo organismo 3GPP, uma vez que esta dissertação segue de perto essa abordagem.

2.1 Primeiras abordagens ao Controlo de Tráfego Baseado em Políticas

As soluções PBNM[6] surgiram como uma possível solução para a problemática do controlo da QoS[5]. A grande promessa inicial destas soluções seria oferecer um maior controlo sobre a QoS percepcionada pelas aplicações e, consequentemente, pelos respectivos utilizadores.

Nesta fase, era dado uma grande ênfase ao facto de que o controlo da rede poderia ser en-

[1] Policy-based Network Management
[2] Internet Engineering Task Force
[3] European Telecommunications Standards Institute - Telecommunications and Internet converged Services and Protocols for Advanced Networking
[4] 3rd Generation Partnership Project
[5] Quality of Service

tregue aos responsáveis de mais alto nível de uma empresa, como os CEOs[6]. Na prática, isto significa que os gestores da empresa seriam capazes de efectuar a configuração da rede através de software com uma interface simplificada e baseada no conceito "point and click", sem que para isso precisassem de ter conhecimento dos mecanismos inerentes à configuração de uma rede. A ideia é a de que, o gestor, através da selecção de algumas opções, poderia configurar a rede para ter um comportamento bem definido sobre os fluxos de tráfego.

A QoS é um conceito relativamente complexo e representa muito mais do que um simples aumento ou diminuição da largura de banda. A QoS incide sobre um conjunto de métricas que irão determinar qual o tratamento fornecido a um determinado agregado de tráfego. Uma vez que diferentes aplicações têm diferentes requisitos de QoS, a quantidade de parâmetros a configurar poderão sofrer variações, consoante se trate de uma aplicação de conversação em tempo real, que precisa de valores de atraso bem limitados e de perda de pacotes bastante baixos, ou uma aplicação de transferência de dados, que poderá ser mais tolerante, quer a atrasos na transmissão, quer à perda de pacotes. De forma a fornecer garantias para os valores destes parâmetros, o controlo da QoS não será feito apenas num ponto, mas sim ao longo do caminho que a *stream* de dados percorrerá e, consequentemente, a configuração da QoS terá que ser atacada ao longo dos equipamentos desse caminho. Desta forma, os valores de configuração da QoS nos diferentes equipamentos de rede determinarão como é que esses equipamentos irão reagir perante os diversos fluxos agregados de tráfego.

O fornecimento de QoS não é uma tarefa trivial. O grande desafio que se coloca é, como controlar a heterogeneidade de equipamentos presente na rede de um operador na perspectiva da satisfação dos níveis do serviço a ser oferecido. Dado que uma rede poderá possuir equipamentos de origem e propósito diversos, a forma como os equipamentos devem ser configurados poderá ser muito distinta entre eles.

As soluções PBNM, que permitem o fornecimentos de serviços diferenciados baseados num controlo estrito da QoS, proporcionam novas oportunidades de negócio. Poder-se-á passar para um modelo onde se oferece qualidade ao invés de quantidade (embora não sejam mutuamente exclusivos). Como a definição de QoS diferenciada é um processo extremamente complicado, a utilização de políticas permitirá, em boa medida, simplificar ou complementar esse processo[7]. A utilização de políticas para controlar o fornecimento de QoS diferenciada é um dos grandes motores por trás das soluções PBNM.

Dados todos estes factores, o entusiasmo em torno do PBNM era bastante grande. A pro-

[6]Chief Executive Officer

messa de que as soluções PBNM permitiriam a simplificação da configuração e manutenção de uma rede, permitindo um maior controlo sobre a QoS fornecida, era uma ideia extremamente aliciante para as empresas, nomeadamente para aquelas que ofereciam serviços IP[7]. No entanto, os problemas das soluções iniciais de PBNM começavam precisamente neste ponto, a excessiva dependência do IP. Uma vez que existiam muitas empresas que utilizavam outro tipo de tecnologia que não o IP (por exemplo, o popular ATM[8]) as soluções PBNM não seriam facilmente aplicáveis.

Outro grande erro cometido nas soluções iniciais de PBNM prendia-se com a abordagem *single-vendor* seguida pelos fabricantes destas soluções. Esta abordagem significa que a interacção com diferentes dispositivos era muito limitada, tipicamente reconhecendo apenas um determinado tipo de equipamento. Para mais, os fabricantes utilizavam interfaces de comunicação proprietários com os seus equipamentos, o que levantava mais uma dificuldade para a inter-operabilidade. Por este motivos, várias soluções PBNM teriam que ser usadas para que se conseguisse controlar os diversos dispositivos. No entanto, a utilização de várias soluções PBNM levantava outro grande problema: a grande dificuldade (ou mesmo impossibilidade) de integração das diferentes soluções PBNM, uma vez que elas próprias eram incompatíveis entre si. Tal situação era facilmente evidenciada pelo simples facto de que não existia um conjunto de normas onde se definisse qual o mecanismo de comunicação e funcionamento das soluções PBNM.

Por último, as soluções PBNM eram muito focadas num determinado tipo de equipamento e numa função específica, o que levantava alguns problemas nas configurações desses equipamentos. Estas soluções poderiam não ter a capacidade de controlar todas as funcionalidades do equipamento. Este problema, por si só, levava também à necessidade de introdução de mais soluções PBNM e consequentemente à dificuldade de integração e interacção que foi previamente referida.

Dadas todas as suas limitações, as soluções iniciais PBNM falharam completamente o seu propósito[6]. Foram mal compreendidas e além disso foram colocadas no mercado sem um conjunto de normas que definissem o seu modo de funcionamento básico ou como deveriam interagir com outras soluções PBNM. Como resultado, as primeiras soluções PBNM sofriam de problemas de escala, uma vez que adicionar mais equipamentos para executar a função de PBNM poderia ser um processo extremamente complexo e difícil.

[7] Internet Protocol
[8] Asynchronous Transfer Mode

Para contribuir ainda mais para a dificuldade destas soluções iniciais de PBNM singrarem no mercado, surge ainda outro problema. Esse problema reside no facto de que as soluções de monitorização de rede ainda estavam muito atrás das soluções de PBNM da altura[6]. Este facto, por si só, significava que não havia um método eficaz para fazer medições de rede com o intuito de determinar a sua real ocupação e, consequentemente, o controlo de admissão à rede estava severamente limitado.

Ir-se-á demonstrar, seguidamente, como é que as soluções PBNM actuais resolveram os seus problemas iniciais e como é que são aplicadas hoje em dia.

2.2 O Controlo de Tráfego Baseado em Políticas actual

A Internet que utilizamos no dia a dia, assenta nativamente num modelo *best-effort*[8]. O que esta constatação significa é que, um utilizador, por omissão, será servido consoante a infra-estrutura de rede tenha capacidade para tal. Esta abordagem não permite a oferta de serviços diferenciados, uma vez que não contempla a inspecção de tráfego e respectivo tratamento diferenciado. Dado que nesta abordagem não existe possibilidade de fornecer tratamento diferenciado ao tráfego, também não existe uma forma eficaz de planear a infra-estrutura da rede para que esta possua um bom comportamento global. Portanto, dadas estas limitações, a abordagem que se seguiu para o planeamento das infra-estruturas de rede foi sobre-dimensionar a infra-estrutura relativamente à largura de banda expectável que será exigida da rede.

O sobre-dimensionamento da rede é feito na esperança de que, o excesso de largura de banda, seja capaz de conseguir dar resposta a todos os fluxos de tráfego que entram na infra-estrutura sem haver atraso ou degradação perceptível (a percepção subjectiva da QoS denomina-se por QoE[9][9]). Através desta abordagem não existe uma solução concreta para o problema sustentado de uma infra-estrutura de rede, havendo, sim, um eterno adiar da solução. Se a rede se está a aproximar do seu limite de funcionamento, então aquilo que se faz não é racionalizar, mas sim, adicionar mais recursos à rede que se traduzam num aumento da largura de banda disponível. O sobre-dimensionamento da largura de banda numa infra-estrutura de rede não pode ser a resposta para a gestão eficaz do tráfego da mesma[6]. Apesar de ser uma abordagem relativamente simples, esta abordagem tem as suas consequências negativas. Este tipo de abordagem é estático e não prevê a necessidade de um auto-ajuste às condições variáveis da rede, havendo portanto uma baixa resiliência da infra-estrutura. Esta abordagem possui também o inconveniente de que,

[9]Quality of Experience

durante o maior período de tempo em que há tráfego na rede, existe um claro desperdício de recursos, uma vez que, durante este período, os níveis de ocupação da rede são tipicamente baixos. No entanto, esta abordagem apresenta também um problema diametralmente oposto. Uma rede sobre-dimensionada, embora possua quantidades de largura de banda muito generosas, não está preparada para dar uma resposta eficaz a picos de utilização repentinos. Se por alguma razão o volume de tráfego ultrapassa a capacidade da rede então teremos uma situação de baixo desempenho da rede perante esta situação, não estando a mesma preparada para absorver estes picos de utilização repentinos.

Existe também o problema de que aplicações cada vez mais sofisticadas requerem serviços igualmente cada vez mais sofisticados[6]. Esses serviços exigem uma resposta bem definida e espectável da rede. A aplicação exclusiva do sobre-dimensionamento da rede como solução para o problema da gestão da largura de banda da rede não garante, de forma alguma, que se possam implementar serviços na rede que precisem de um tratamento diferenciado[10]. Nesse sentido, o sobre-dimensionamento da rede não poderá ser a resposta definitiva para o planeamento e gestão dos recursos de uma infra-estrutura de rede.

O PBMN pretende dar resposta a esta problemática através da utilização de modelos de informação e de um estreitar de relações com os modelos de negócio permitindo a instanciação dos mesmos ao nível da configuração dos elementos de rede. O PBNM posiciona-se como uma filosofia de excelência para potenciar cada vez mais o funcionamento da infra-estrutura de rede e tornar a sua configuração e gestão em processos bastante simplificados[7]. O PBNM permite, efectivamente, fazer uma gestão inteligente dos recursos de rede, tornando esses mesmos recursos numa oportunidade rentável de negócio[11]. Permite igualmente, de forma concisa, abstrair os detalhes de comunicação com os equipamentos de rede e, em simultâneo, providenciar um método de alto nível para criação de políticas que irão determinar o funcionamento da infra-estrutura de rede[12]. O sistema PBNM terá a capacidade de traduzir as políticas criadas através de métodos mais alto nível em parâmetros de configuração mais baixo nível, inteligíveis pelos equipamentos de rede. Um sistema deste género terá, igualmente, de ser capaz de controlar um grande número de equipamentos de rede, possivelmente com interfaces de comunicação bastante distintas, mas cuja especificação e implementação é bem conhecida.

O *Policy Based Network Management* apresenta 6 grandes objectivos:

- Fornecer um serviço melhor do que o típico *best-effort*, e que este possa ser proposto como uma oferta comercial;

- Simplificar o controlo e configuração dos equipamentos e serviços de rede;
- Diminuir a quantidade de activos humanos necessários para configurar a rede;
- Definir o comportamento de uma rede ou de um sistema distribuído;
- Controlar a complexidade cada vez maior associada à configuração dos equipamentos de rede;
- Utilizar as necessidades e procedimentos de um modelo de negócio para configurar os equipamentos de rede de forma a darem resposta a essas mesmas necessidades.

Como foi anteriormente referido, a Internet foi desenvolvida como um serviço *best-effort*. O que isto significa é que, na Internet, não existem garantias concretas de entrega de tráfego. Será feito o melhor esforço para que os procedimentos corram bem, mas não há uma garantia implícita que tal irá acontecer[13]. Essa falta de garantias é preocupante, e começaram a surgir cada vez mais utilizadores ou serviços com necessidades específicas. Ou seja, começou a surgir a necessidade de garantir níveis mínimos de largura de banda, níveis máximos de atraso ou de perda de pacotes, etc. Portanto, começa a haver uma corrente de utilizadores que espera que o seu tráfego Internet possa ter um comportamento bem definido e esperado, e para os quais o serviço *best-effort* não é suficiente. Para dar resposta a esta necessidade, foram criados modelos de serviços orientados a fornecer um comportamento esperado para o tráfego Internet. Modelos como o Diffserv[14] ou Intserv[15] permitem dar resposta a esta problemática. A QoS, será então um meio para fornecer aos utilizadores um serviço melhor do que o *best-effort*, o qual terá um comportamento previamente definido e como tal futuramente expectável.

A oferta de QoS à medida não é trivial e, naturalmente, existem alguns desafios para a sua implementação. Uma QoS consistente apresenta, por si só, uma grande complexidade de implementação. Além disso, existe uma enorme variedade de serviços que poderão usar e tirar partido da QoS, serviços esses que poderão ter requisitos bastante distintos. Sendo assim, a primeira grande motivação para a utilização do PBNM é a promessa de utilização de um conjunto de normas para controlar os diferentes mecanismos de QoS que serão necessários para implementar um determinado serviço de rede.

2.2.1 Simplificação da gestão e configuração dos equipamentos de rede

Um dos grandes pilares do PBNM é a simplificação da gestão e configuração dos equipamentos e serviços de uma rede. Consiste numa série de mecanismos que permitirão fazer intervenções precisas nos diferentes serviços de rede. Nesse sentido, o PBNM irá implementar mecanismos que poderão condicionar o fluxo de tráfego ao longo da infra-estrutura de rede assim como definir um conjunto de mecanismos complexos que permitirão a implementação de um determinado serviço de rede. A capacidade de um sistema PBNM poder controlar e condicionar os fluxos de tráfego trazem consigo um efeito colateral extremamente benéfico: uma camada de segurança adicional[6]. Como existe uma fase de admissão de fluxos à rede, esses mesmo fluxos podem ser recusados caso não estejam em conformidade com a política definida pelo operador. Este controlo garante também que o acesso aos recursos de rede ou mesmo à informação disponibilizada pela rede será feita apenas por utilizadores autorizados para tal, reforçando ainda mais a ideia de oferta de serviços diferenciados. A utilização de métodos de análise de tráfego mais avançados, como por exemplo, o DPI[10], potenciarão o cumprimento estrito das políticas definidas.

A capacidade de controlar vários equipamentos distintos é um grande trunfo do PBNM. Esta capacidade é conseguida através da abstracção. Para atingir o necessário nível de abstracção, as políticas que irão condicionar o funcionamento da rede não são escritas na linguagem directamente implementada pelos equipamentos de rede, mas sim através de uma linguagem intermédia que será posteriormente, através de uma camada de abstracção adicional, traduzida para a linguagem específica dos equipamentos de rede. Através deste método é conseguida uma simplificação da configuração da rede através da abstracção. Outro ponto muito importante ao qual o PBNM consegue dar resposta é a recuperação da configuração dos equipamentos de rede para um estado funcional. Muitas vezes, para dar resposta a problemas ocasionais que possam surgir numa rede, poderá haver necessidade de intervenção directa na configuração dos equipamentos de rede. O acumular destas pequenas intervenções poderá trazer problemas sérios de baixo desempenho da rede, colocar o equipamento num estado inconsistente, ou mais grave ainda, impossibilitar a implementação dos diversos serviços de rede[6]. O PBNM, através da manutenção do estado dos diversos equipamentos da rede, poderá repor uma configuração funcional no respectivo equipamento, caso essa necessidade surja. Dessa forma, poder-se-á restaurar as capacidades básicas da rede e voltar a deixá-la num estado consistente, capaz de, num curto espaço de tempo, receber os diversos serviços que deve suportar diminuindo o impacto causado por eventuais *downtimes*.

Um maior nível de abstracção das operações de manutenção e configuração de rede trás con-

[10]Deep Packet Inspection

sigo uma outra vantagem: a diminuição da dependência de mão de obra especializada. Dada a maior automatização de processos, o número de pessoas efectivamente necessárias para a manutenção e configuração da rede é inerentemente menor. Esta característica permite minimizar um problema associado à típica resistência à entrada de novas tecnologias: o custo. Esse custo associado à introdução e utilização de tecnologias emergentes é tipicamente derivado da necessidade de aprendizagem das novas tecnologias, e toda a complexidade que elas acarretam. Uma vez que a relação custo/benefício da introdução de novas tecnologias acabava por ser desfavorável à empresa, a adopção dessas mesmas tecnologias era tipicamente lenta ou mesmo inexistente. As tecnologias para controlo da largura de banda sofreram, precisamente, com esta situação. Em grande medida, acabava por ser mais fácil planear uma infra-estrutura de rede através do sobre-dimensionamento do que através de uma planificação cuidada, que levaria em linha de conta tecnologias de gestão eficaz da QoS de uma rede. O sobre-dimensionamento acabava por ser mais barato do que a implementação de tecnologias de gestão de QoS. Havia também o problema da interoperabilidade. A utilização de diversas tecnologias, tipicamente distintas entre si, fazia com que, para que elas fossem capazes de funcionar em conjunto, teria que haver uma especificação extremamente pormenorizada das suas funcionalidades. Ao existir um modelo de informação que identifica as capacidades dos equipamentos de rede e respectiva forma como devem interagir, essa necessidade deixa de existir. Portanto, a ideia é a de que, a menor dependência de mão de obra humana especializada, fará com que a rede seja mais resiliente e tenha uma capacidade maior em se adaptar às alterações dinâmicas que ocorrem. A definição do funcionamento do produto PBNM terá que ser feita por pessoas extremamente especializadas, claramente conscientes tas suas funções.

Existe um cada vez maior número de pessoas e aplicações a utilizar a rede. Nesse sentido, a rede tem que ser suficientemente flexível para aceitar os procedimentos que irão permitir que esta responda bem quando existe uma necessidade de a configurar ou actualizar para que possa receber um grande número de utilizadores. Nessa medida, uma solução PBNM terá que ter a capacidade de gerir um grande número de utilizadores/equipamentos para garantir a escalabilidade da solução. Uma solução PBNM terá igualmente que ter a capacidade de fazer agendamento de alterações à rede, possibilitando intervenções na rede minuciosamente agendadas, de forma a evitar intervenções em períodos de maior utilização.

A complexidade de uma rede reside também no número e variedade de equipamentos que podem constituir a mesma, como já havia sido anteriormente referido. Os diferentes equipamentos poderão possuir diferentes protocolos de comunicação para a realização de alterações às suas configurações. Devido a esta situação existirá uma necessidade de fazer um mapeamento

entre os diferentes comandos dos diferentes equipamentos que permitam atingir um objectivo comum. Um modelo de informação que permite descrever as regras de *policing* e que, simultaneamente, as abstrai da implementação propriamente dita, tal como preconizado no PBNM, permite, efectivamente, atingir esse objectivo. Ao não haver um mapeamento directo das operações de configuração dos equipamentos de rede na linguagem por eles implementada, mas sim recorrendo a uma linguagem intermédia, que é independente da implementação do equipamento de rede, a solução de PBNM terá maior facilidade em lidar com os diversos equipamentos.

2.2.2 Regras de negócio ditam o funcionamento da rede

A grande visão do PBNM é possibilitar que as regras de negócio de uma empresa determinem a configuração da rede: *Using business rules to drive network configuration*[6]. Esta visão é extremamente importante para poder tornar o núcleo da rede num centro de lucro efectivo. Como foi anteriormente referido, o problema da QoS foi "resolvido" através do sobre-dimensionamento dos recursos de rede. A verdade é que esta abordagem não resolve o problema, mas sim evita-o no curto ou médio prazo. Além disso, o planeamento da infra-estrutura de uma rede partindo do princípio de que a infra-estrutura deve estar sobre-dimensionada traz, invariavelmente, uma menor potenciação da recolha de dividendos resultantes da exploração da própria rede. Com o sobre-dimensionamento da rede, o problema da QoS continua a existir, e como foi visto anteriormente, esta abordagem faz com que a rede não tenha capacidade de resposta dinâmica a variações na sua utilização.

Nesta fase, a rede ainda não é vista como um possível centro de lucro e como uma forma para potenciar novas oportunidades de negócio. O PBNM permite alterar por completo a visão de que a rede é um centro de prejuízo económico, uma vez que o PBNM abre novas oportunidades e modelos de negócio. A visão deste é a de que, as regras de negócio (que são uma especificação de alto nível conhecidas como SLS[11] e SLA[12]) serão traduzidas para as configurações concretas dos equipamentos, que permitirão implementar um determinado serviço ao longo de toda a infra-estrutura de rede. Portanto, aquilo que é proposto ao cliente final é um conjunto de especificações de alto nível que identificarão qual o serviço que está a ser oferecido (por exemplo, um serviço VoIP[13] terá necessidades de atraso bem controladas e largura de banda mínima garantida). O que o cliente final precisa de saber é que o serviço que lhe está a ser proposto terá o comportamento

[11] Service Level Specification
[12] Service Level Agreement
[13] Voice over IP

contratualizado. A forma como ele será implementado ao longo da infra-estrutura de rede é uma preocupação que não diz respeito ao cliente. Tudo o que interessa é a Qualidade de Serviço percepcionada pelo mesmo (QoE).

Como haverá uma alteração à configuração dos equipamentos de rede para permitir a implementação de um determinado serviço terão que existir algumas preocupações adicionais. A rede não irá fornecer um serviço de cada vez, mas sim uma grande quantidade de serviços em simultâneo. Nesse sentido, tem que haver garantias que as novas alterações não provocarão degradação da QoS previamente oferecida pela rede. Tem que haver igualmente uma garantia prévia de que a admissão de um novo serviço na rede não irá provocar disrupção na qualidade dos serviços actuais[16]. A estreita relação entre o administrador e o operador de rede terá que estar sempre presente. Não pode haver uma oferta adicional de serviços se por ventura a rede não tiver capacidade para os receber sem influenciar negativamente os serviços actualmente em funcionamento na rede. Pode-se então concluir que *Business and Network personell* têm que trabalhar em conjunto de forma a garantirem que os serviços de rede sejam geridos de acordo com as regras de negócio contratualizadas com os clientes e de que a rede está preparada para receber eventuais novos serviços contratados.

2.2.3 Sumário dos benefícios do PBNM

Em jeito de conclusão, pode-se constatar que o PBNM traz inúmero benefícios para as instituições que optem por utilizar esta filosofia. Se nas primeiras abordagens ao PBNM nos deparávamos com abordagens *single vendor* e focadas em tecnologias particulares, sem suporte sem suporte normativo definido, hoje em dia a situação é substancialmente melhor.

O PBNM auto promove-se recorrendo à importância que este deposita nas regras e processos de negócio. Dados todos os avanços na área, o PBNM é uma tecnologia bastante simples de utilizar e perceber[7]. Trouxe inúmeras vantagens, nomeadamente a capacidade de fornecer serviços diferenciados a utilizadores com necessidades distintas. Permite a implementação de serviços cada vez mais complexos através de um diversificado conjunto de equipamentos com um número de pessoas reduzido que são necessárias para os operar. Permite igualmente a simplificação da configuração de dispositivos da rede, da manutenção da mesma e da fácil implementação e controlo dos serviços por esta oferecidos. A utilização de um modelo de informação que garante a interoperabilidade entre os diversos dispositivos da rede foi também uma aposta ganha pelo PBNM. Toda a complexidade de fazer diferentes dispositivos interagir entre si de forma coope-

rativa é grandemente simplificada através da utilização do PBNM. A capacidade de inspecção de tráfego trouxe também vantagens, ao dotar a infra-estrutura de uma maior inteligência para a gestão de tráfego, permitindo que a rede opte por caminhos mais vantajosos de acordo com o perfil de um determinado fluxo de tráfego e que se adapte dinamicamente às condições de tráfego do momento. Esta abordagem permite também que se poupem recursos na construção de uma rede.

Mas de facto, o grande avanço do PBNM verifica-se a partir do momento em que o PBNM foi claramente mostrado como um meio onde as regras de negócio ditavam a configuração e comportamento da rede. Este é um ponto de viragem, a partir do qual o núcleo da rede se torna num ponto de lucro efectivo. Este ponto de viragem tem influência em toda a instituição, cujo objectivo primário é obter lucro. Uma vez que a rede pode agora ser planeada de uma forma inteligente, não há necessidade de se recorrer, única e recorrentemente, ao sobre-dimensionamento para garantir a existência de largura de banda suficiente para todos os utilizadores (garantia que com o sobre-aprovisionamento não é conseguida).

Para além de todas estas vantagens, existe também uma clara melhoria em termos de segurança, através dos mecanismos utilizados para fazer inspecção do tráfego e admissão do mesmo. Fluxos de tráfego não autorizados são imediatamente detectados e descartados. Por fim, a gestão da complexidade da rede é bastante simplificada através da utilização de soluções PBNM.

2.3 O PBNM no contexto das NGN

Como explicado anteriormente, o PBNM apresenta-se como uma forma para definir necessidades de negócio e assegurar que a rede é capaz de fornecer os serviços que os seus clientes precisam. É a capacidade de uma rede dar resposta automática à alteração das suas condições, resposta essa baseada numa série de comportamentos pré-programados, denominadas políticas.

Num contexto de auto-configuração da rede baseada em políticas, o ponto fulcral passa por gerir eficazmente os recursos dessa rede. Na perspectiva das redes IP, pretende-se gerir os recursos da rede de tal forma que, uma grande diversidade de serviços com requisitos de QoS específicos e distintos entre si possuam fiabilidade e garantias de entrega, mesmo funcionando sobre um grande e diversificado número de tecnologias de transporte. O controlo dos recursos da rede passa por controlar o QoS fornecido, dar permissões para a passagem de fluxos de tráfego (controlo de *gating*), segurança, etc. A capacidade de resposta em tempo real do sistema PBNM passa por ser capaz de fazer controlo de admissão, reserva de recursos, aplicação de restrições

de tráfego, etc.

Um sistema PBNM será idealmente conduzido por regras de negócio. O tratamento do tráfego dos utilizadores de uma rede será feito de acordo com os SLA estabelecidos entre estes e o respectivo operador. As características operacionais da rede serão definidas pelo SLS, que é uma parte do SLA. Os parâmetros do SLS serão posteriormente mapeados para parâmetros de configuração da rede. É aqui que reside o poder do PBNM. Esta camada de tradução dos parâmetros de especificação do serviço para parâmetros de configuração da rede é feita através de um *middleware*. Desta forma, alterações às tecnologias de rede de core (por exemplo, adição de um *router* com um protocolo de configuração novo), não será problemático uma vez que a única alteração necessária será adaptar o mecanismo de tradução dos SLS para a nova tecnologia de rede.

Esta camada de tradução entre os SLS e os parâmetros de rede propriamente ditos denomina-se, no contexto das NGN[14], por PDP[15]. O PDP é uma entidade que lida com eventos assíncronos e que fornece respostas com base na SLA definida para o cliente que deu origem ao evento. Essa resposta será interpretada e aplicada por um PEP[16]. Este elemento será o responsável pela admissão de novas sessões e de garantir que o tratamento do tráfego está de acordo com o SLA do cliente.

Uma política consiste numa série de regras que determinam o funcionamento de um sistema[17]. As regras associadas a uma política podem ser provenientes de vários pontos, havendo tipicamente uma distinção entre origem de pontos estáticos (por exemplo, a informação de perfil do subscritor guardada numa base de dados de subscritores) ou então informação dinamicamente fornecida pela rede (qual a QoS disponível, etc.). Uma política pode ser vista como um tripleto Evento/Condição/Acção. Uma abstracção deste modelo pode ser vista na figura 2.1. A ocorrência de um evento irá motivar a verificação de uma condição. A verificação dessa condição como verdadeira irá provocar a execução de uma acção. Por exemplo, a condição poderá ser verificar se o endereço IP se encontra dentro de uma gama pré-definida e, caso se verifique essa condição, a acção a executar poderá ser, por exemplo, autorizar a passagem dos fluxos de tráfego. Portanto, a ocorrência de um evento provoca uma tomada de decisão, que por sua vez irá eventualmente originar a execução de uma determinada acção.

A acção a executar associada ao par Evento/Condição irá determinar a forma como o PEP deverá reagir perante um dado fluxo de tráfego, isto é, quais são os passos necessários para

[14]Next Generation Networks
[15]Policy Decision Point
[16]Policy Enforcement Point

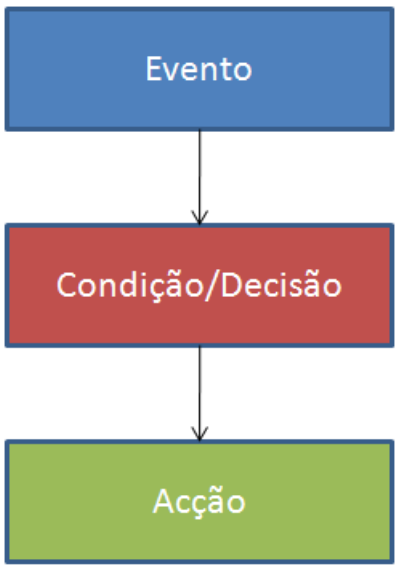

Figura 2.1: Abstracção do modelo Evento/Condição/Acção

que o PEP efectivamente aplique a política definida. A aplicação da política poderá passar por várias acções, tais como, permitir a passagem dos fluxos de tráfego, garantir largura de banda mínima, limitar a largura de banda máxima para um determinado fluxo de dados, fazer marcação de pacotes, priorização de fluxos de tráfego, etc.

O controlo de QoS é um dos ponto mais importantes dos sistemas PBNM. A admissão de novos serviços à rede tem que ser feita de uma forma sustentada, garantindo que a admissão de um novo serviço não irá provocar degradação do desempenho da rede e também garantir que esse serviço irá conviver de forma harmoniosa com os outros serviços previamente definidos. Nesse sentido, se uma reserva de recursos é solicitada, terá que haver uma fase de Controlo de Admissão com o intuito de apurar se a rede tem ou não capacidade para dar resposta ao pedido, que o utilizador tem efectivamente permissões para essa reserva e que as políticas globais de rede aceitam as características da reserva. Portanto, a admissão de novos serviços à rede, além de ter de ser feita de forma a que não haja impacto negativo nos serviços previamente configurados, passa também por verificar se o pedido de reserva é aceite de acordo com as políticas globais da rede (limitações de *codecs*, de tempo, etc.) e se o perfil do utilizador lhe garante a possibilidade de fazer tal pedido (o utilizador poderá estar a tentar fazer uma reserva de recursos à qual não tem direito). Este mecanismo de controlo de admissão servirá para garantir o funcionamento sustentado da rede. Portanto, a admissão de novos serviços à rede passa efectivamente por três passos:

- Autorização - onde é feita a verificação das políticas globais da rede e onde é verificado o perfil do utilizador no sentido de apurar se a reserva poderá ser feita sem violar as políticas da rede;

- Reserva - A rede é sondada para apurar se existe capacidade para aplicar o serviço e respectivos recursos a ele associados. Caso a rede possa dar resposta ao novo serviço, então os recursos são reservados;

- Comprometimento - Os recursos reservados são aplicados e o tráfego do utilizador é aceite. Esta fase garante que o tratamento fornecido ao tráfego está de acordo com o definido nos parâmetros de serviço.

Concluindo, o PBNM no contexto das NGN assegura que o tráfego recebe um tratamento diferenciado de acordo com o SLA especificado para cada utilizador. A admissão de novos serviços à rede é feita de forma sustentada, verificando que o utilizador está habilitado para requerer esses serviços e que a rede os poderá servir sem ocorrer degradação que já estão presentes. Simultaneamente, a aplicação de políticas e respectiva configuração dos elementos de rede torna-se um processo relativamente simples recorrendo à utilização de um *middleware*, designado por PDP.

2.4 O PBNM no IETF

Através da RFC 2753[18], o IETF propõe uma *framework* para controlo baseado em políticas orientado a decisões de controlo de admissão. Nesta *framework*, o IETF propõe a existência de dois elementos fundamentais, o PDP[17] e o PEP[18], tal como podem ser vistos na figura 2.2.

O PDP será o elemento responsável pela tomada de decisões de aplicação de políticas. A tomada de decisões por este elemento é orientada ao evento, isto é, a ocorrência de um determinado evento irá despoletar a tomada de uma decisão por parte deste elemento. A decisão consiste em configurações que serão aplicadas no elemento PEP. A interacção básica entre o PEP e o PDP começa pelo PEP. O PEP será o elemento responsável por despoletar os eventos utilizados pelo PDP para a tomada de decisões no que concerne a aplicação de políticas. É também responsável por, depois de receber a decisão do PDP, aplicar a decisão recebida. De notar que neste processo, o PDP poderá contactar elementos externos, por exemplo, o elemento responsável pela taxação.

[17]Policy Decision Point
[18]Policy Enforcement Point

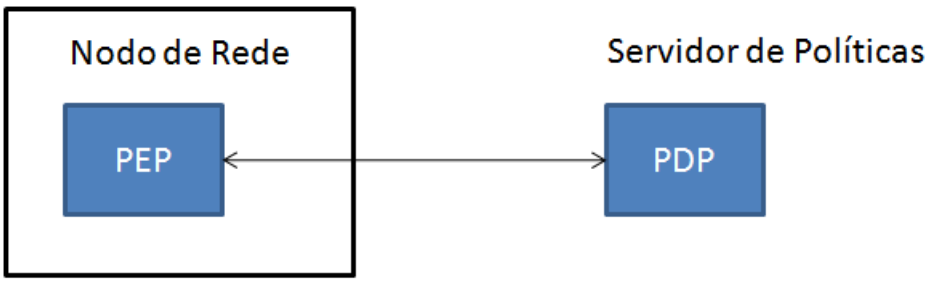

Figura 2.2: Arquitectura simples com os elementos primários da arquitectura do IEFT

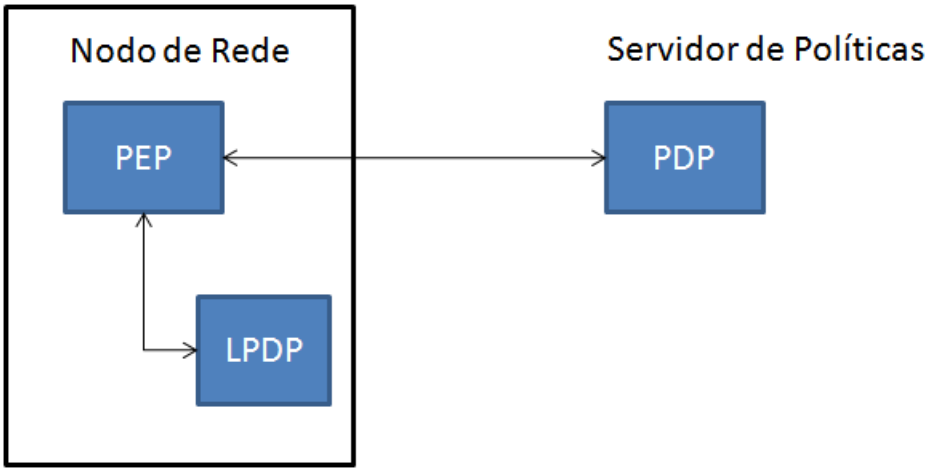

Figura 2.3: Arquitectura com inclusão de um LPDP

A *framework* proposta pelo IETF prevê a possível existência de um LPDP[19], tal como pode ser observado na figura 2.3. O LPDP será um elemento muito próximo do PEP que poderá ser utilizado para tomar decisões numa primeira instância, antes de consultar o PDP (caso este exista). No caso de, quer o LDPD, quer o PDP estarem presentes, embora o LPDP possa tomar decisões, estas serão sempre analisadas pelo PDP e, em última instância, a decisão entregue pelo PDP é a que prevalece. O LPDP poderá ser útil para cenários em que não exista conectividade temporária ou definitiva com um PDP.

O PDP é uma entidade remota que pode residir num servidor de políticas. No entanto o servidor de políticas pode ser externo, tal como exemplificado na figura 2.4. O servidor de políticas, ou repositório de políticas, contém a informação necessária para a tomada de decisões por parte do PDP (por exemplo, informação do utilizador referente ao perfil do mesmo, valores

[19]Local Policy Decision Point

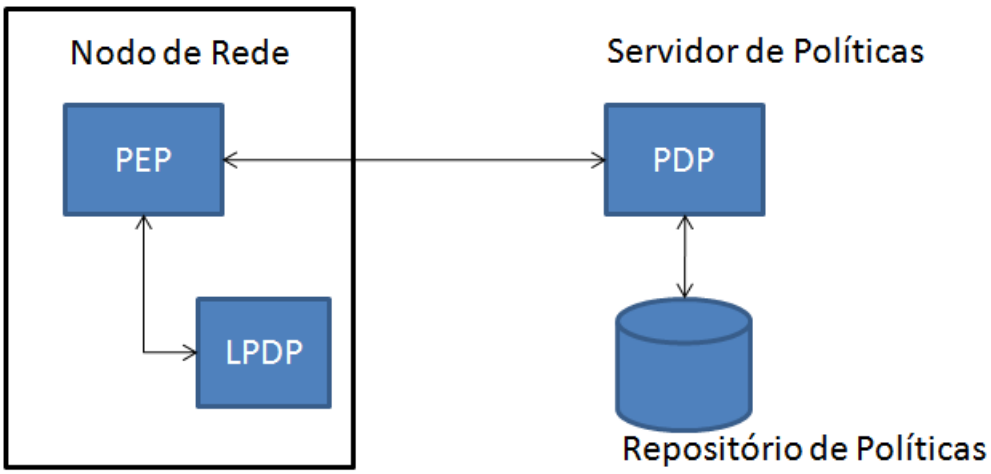

Figura 2.4: Configuração da arquitectura com inclusão de um LPDP e um Repositório de Políticas externo

máximos de largura de banda para o utilizador, etc).

A *framework* proposta pelo IETF não especifíca qualquer protocolo de comunicação concreto entre o PDP e o PEP. No entanto, na RFC 2748 é proposto um modelo cliente/servidor para suportar controlo de policiamento através de protocolos de sinalização de QoS, o protocolo COPS[20][19]. Este protocolo permite a troca de informação de policiamento entre o PDP e o PEP, tal como se pode observar na figura 2.5. Neste protocolo o PEP envia pedidos/actualizações/remoções para o PDP remoto, e este último responde de acordo com a decisão tomada para o evento específico[20]. Para fornecer garantias de entrega, o protocolo COPS opera sobre TCP, podendo utilizar extensões, como o TLS[21], para cifragem dos dados trocados. O PEP é sempre o responsável por iniciar a ligação inicial com o PDP.

O protocolo COPS suporta dois modelos de definição de políticas, *pull e push*[19]. No modo *pull*, a entrega de uma decisão de políticas é feita como reacção a um pedido prévio proveniente do PEP, isto é, ocorreu um evento para o qual o PEP gerou um pedido de decisão ao PDP. No modo *push*, a entrega de políticas é feita pelo PDP ao PEP sem que este último as tenha explicitamente requisitado, isto é, esta entrega de regras foi decidida de forma unilateral pelo PDP, como resposta, por exemplo a um evento interno deste.

[20]Common Open Policy Service
[21]Transport Layer Security

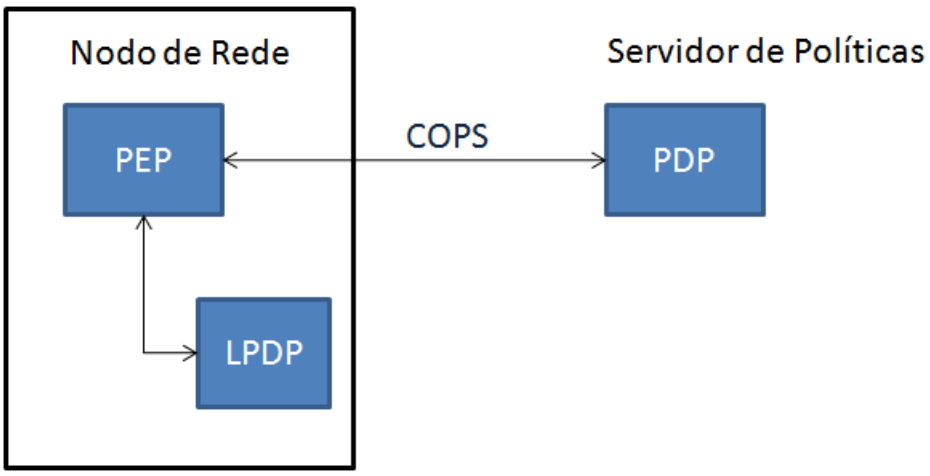

Figura 2.5: Arquitectura com uso do protocolo COPS para comunicacção entre o PEP e o PDP

2.5 O PBNM no ETSI-TISPAN

O TISPAN[22] teve origem no ano de 2003, resultado da fusão de dois antigos grupos, o TIPHON[23] e o SPAN[24]. Esta fusão teve o objectivo de definir uma visão, fundamentalmente europeia, para uma rede da próxima geração. Embora de âmbito alargado, o TISPAN tem concentrado os seus esforços na evolução da rede fixa, definindo camadas de componentes, funções e responsabilidades independentes entre si. Desta forma, alterações feitas a uma das camadas não irão ter impacto nas camadas adjacentes.

A figura 2.6 mostra as camadas e algumas entidades funcionais da arquitectura[21] preconizada pelo TISPAN.

O componente responsável pelo controlo e gestão de QoS encontra-se na camada de controlo e denomina-se por RACS[25][22]. O RACS é o elemento aglutinador das camadas de serviço e de transporte. O RACS apresenta uma única interface para a camada de serviço, a interface Gq'[23], abstraindo-se assim dos detalhes de topologia e tecnologia inerentes ao transporte e ao terminal. Este componente apresenta todas as características de um *Policy Server*.

A entidade funcional RACS é dividida em duas funções, o SPDF[26] e o A-RACF[27], como

[22]Telecoms & Internet Converged Services & Protocols for Advanced Networks
[23]Telecommunications and Internet Protocol Harmonization Over Networks
[24]Services and Protocols for Advanced Networks
[25]Resource Admission Control Sub-System
[26]Service Policy Decision Function
[27]Access-Resource Admission Control Function

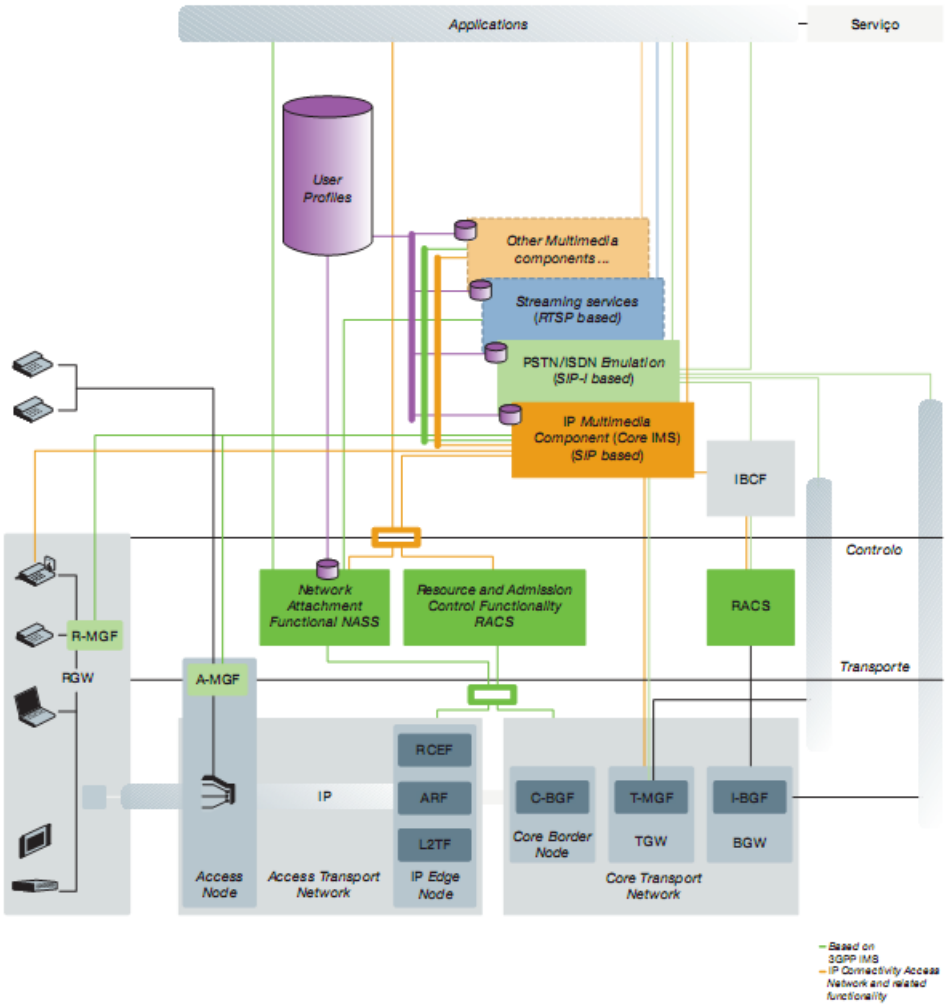

Figura 2.6: Arquitectura TISPAN

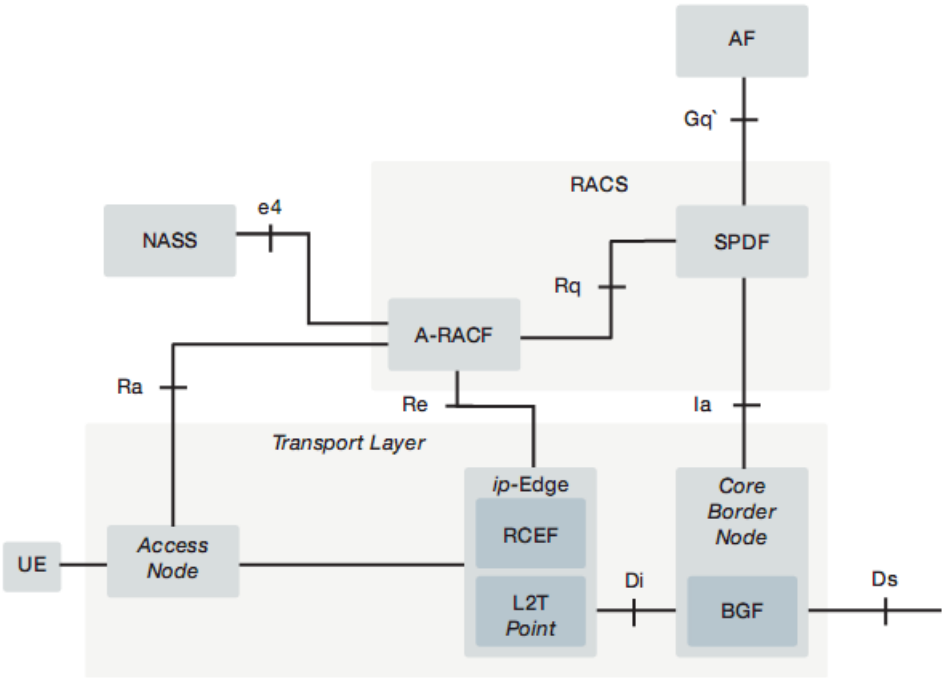

Figura 2.7: Arquitectura RACS

pode ser observado na figura 2.7.

Dentro do RACS, o SPDF é o responsável pela comunicação com a camada de serviço, através da interface Gq'. O SPDF verifica se os pedidos respeitam um conjunto de regras (políticas) definidas pelo operador para o serviço em questão. Se a sessão for aceite, o SPDF verifica se existe necessidade de contactar o A-RACF e/ou o BGF[28] para que seja executada a verificação do pedido ao nível da camada de transporte, isto é, verificar se existem recursos suficientes para dar resposta ao pedido formulado. O SPDF é também o responsável por efectuar as devidas diligências aquando da existência de NAT[29] ao nível do BGF.

O A-RACF, com base na informação da topologia, do estado da rede e da informação do utilizador, verifica se o pedido poderá ser aceite ou não. O A-RACF é, portanto, o elemento responsável pela admissão de novos fluxos de dados à rede. Em caso de aceitação do pedido, poderá haver a necessidade de configurar os elementos de transporte RCEF[30] para executarem acções específicas sobre o fluxo de dados, para que o comportamento desse fluxo de dados esteja

[28]Border Gateway Function
[29]Network Address Translation
[30]Resource Control Enforcement Function

de acordo com o determinado pelo RACS.

O modo de provisionamento das regras de policiamento na arquitectura TISPAN é exclusivamente através do modo *push*, isto é, sem haver um pedido prévio proveniente dos equipamentos da camada de transporte. Tipicamente, a instalação de regras posteriores ao registo de utilizador será motivada por pedidos provenientes da camada de serviço, através da interface Gq'. Na fase de registo do utilizador, poderão ser instaladas regras por omissão, definidas pelo operador.

2.6 O PBNM no 3GPP

O organismo de normalização 3GPP foi o primeiro organismo de normalização internacional a avançar para a definição de uma arquitectura para controlo de policiamento e QoS. O 3GPP identifica, claramente, diferentes elementos funcionais interligados entre si através dos denominados pontos de referência. Na sua primeira versão, a versão 5, a comunicação entre o PDP (instanciado na arquitectura pelo PDF[31]) e o PDP (instanciado na arquitectura exclusivamente pelo GGSN[32]), é feita através do ponto de referência Go[24]. As mensagens que fluem no ponto Go entre os dois componentes funcionais da arquitectura são trocadas através do protocolo COPS[24]. Nesta versão, o PDP é uma mistura entre um PDP puro e o P-CSCF[33], que garante a ligação à rede IMS[34][25, 26]. Na versão 6 já existe uma diferenciação entre o PDP e o AF[35]. Nesta versão, a comunicação entre estes dois elementos é feita por intermédio da interface Gq[27].

A versão 7 do 3GPP define a arquitectura *Policy and Charging Control* (PCC])[28]. A arquitectura PCC do 3GPP é, em si, uma harmonização e fusão da componente de policy e charging definidas na versão 6 do 3GPP (*Enhanced Policy Control* e *Flow Based Charging*)[29]. A partir da versão 6, e abrangendo igualmente a versão 7, a comunicação entre os diversos componentes da arquitectura deixou de ser feita pelo protocolo COPS e passou a ser feita por DIAMETER[30]. A harmonização dos componentes promovida pelo versão 7 da arquitectura PCC do 3GPP permitirá a optimização das interacções em tempo real com a rede de comutação de pacotes[29]. Além disso, esta versão traz também uma alteração muito importante, que é a independência da arquitectura relativamente à tecnologia de acesso. Desta forma a arquitectura poderá ser utilizada

[31] Policy Decision Function
[32] Gateway GPRS Support Node
[33] Proxy - Call Session Control Function
[34] IP Multimedia Subsystem
[35] Application Function

com qualquer rede de acesso, identificada na nomenclatura do 3GPP por IP-CAN[36][31]. Como se irá poder constatar mais à frente, o tipo de acesso IP poderá também ser usado para decisões de aplicação de políticas de QoS.

A arquitectura PCC unificada da versão 7 do 3GPP identifica claramente 6 componentes funcionais:

- PCRF - Policy Control and Charging Rules Function;
- PCEF (contido no GW[37]) - Policy and Charging Enforcement Function;
- AF - Application Function;
- SPR - Subscription Profile Repository;
- OCS - Online Charging System;
- OFCS - Offline Charging System.

A disposição dos diversos componentes da arquitectura PCC unificada pode ser vista na figura 2.8.

As três principais entidades do PCC no âmbito desta dissertação são o AF, o PCRF[38] e o PCEF[39]. O PCRF pode ser subdividido em dois componentes, o PDF[40] e o CRF[41] baseado em fluxos de dados que irão permitir definir como um *service data flow*[42] será tratado no PCEF, assim como garantir que o PCEF fará o mapeamento e tratamento do tráfego de acordo com o perfil do utilizador determinado pelo PCRF. Por outro lado, o PCEF actua como um PEP e como TPF[43]. O PEP consiste numa *gate enforcement* (uma *gate* ou está aberta ou está fechada, o que permitirá a passagem ou bloqueio selectivo de fluxos de dados do serviço) e uma função de aplicação de QoS, onde a informação de QoS autorizada é mapeado para os atributos de QoS específicos para um determinado IP-CAN. O TPF, tal como o PEP, também exerce uma função de *gating* (similar no funcionamento ao PEP, no sentido em que também permite a passagem ou

[36]IP-Connectivity Access Network
[37]Gateway
[38]Policy and Charging Rules Function
[39]Policy and Charging Enforcement Function
[40]Policy Decision Function
[41]Charging Rules Function
[42]Um *Service Data Flow* é um agregado de fluxos de pacotes que correspondem a um modelo para fluxos de dados do serviço
[43]Traffic Plane Function

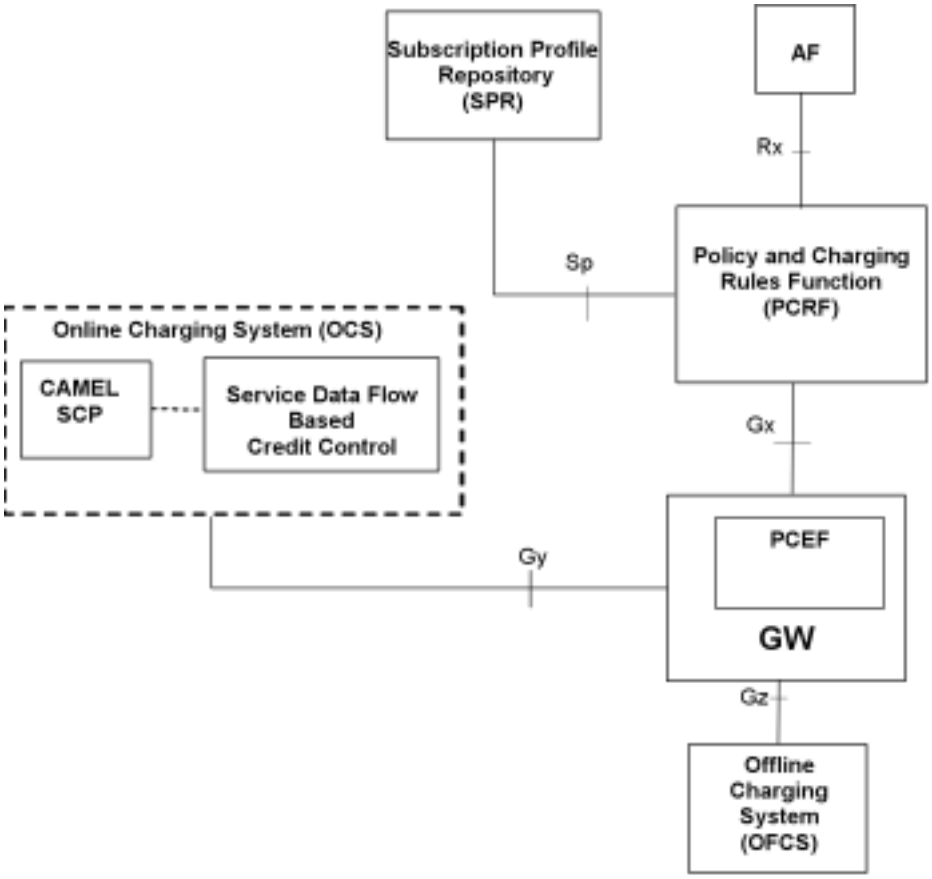

Figura 2.8: Arquitectura PCC

bloqueio selectivo de fluxos de dados do serviço), exercendo e forçando a aplicação de regras de taxação baseadas numa *Charging key* e na regra PCC (a qual pode ser estática e estar definida no GW ou então ser obtida a partir do PCRF, o que permite um controlo dinâmico sobre o comportamento do mesmo). O AF é um elemento que representa aplicações que requerem um controlo dinâmico sobre o Policiamento e QoS aplicados ao comportamento do plano de tráfego. De uma forma mais generalista, permite receber pedidos de requisitos QoS específicos de aplicações, que são posteriormente comunicados ao PCRF. Esses requisitos poderão depois ser utilizados para aplicação de regras de taxação apropriadas e para a aplicação de políticas locais de serviço pelo PCRF.

Como entidades que não se afiguram de capital importância no âmbito desta dissertação, temos também o SPR[44], o OCS[45] e o OFCS[46]. O SPR é uma entidade que contém informação dos subscritores/subscrições necessárias para o PCRF derivar as políticas de subscrição e as regras de taxação. Relativamente à componente de taxação, temos a componente de *Charging Online* (OCS) e a componente de *Charging Offline* (OFCS). A componente de *online charging* pressupõe uma pré autorização para utilização dos serviços de rede, indicando parâmetros, como por exemplo, o crédito disponível para o subscritor ao PCEF. No caso do OFCS, este é utilizado para guardar os parâmetros que serão alvo de taxação por parte do operador, sendo essa informação fornecida pelo PCEF.

A ligação entre o AF e o PCRF é concretizada através de uma interface Rx[32] (que resultou da combinação das partes mais importantes dos pontos de referência Gq e Rx definidos na versão 6) enquanto que, a ligação entre o PCEF e o PCRF é concretizada através da interface Gx[33] (resultado da combinação das componentes mais importantes dos pontos de referência Gx e Go, definidos na versão 6). A interface Rx permite o transporte de informação de sessão ao nível da aplicação, entre o AF e o PCRF. A interface Gx permite que o PCRF tenho um controlo dinâmico sobre o comportamento do PCEF. O principal objectivo desta interface é permitir que o PCRF possa provisionar regras PCC ao PCEF de forma a determinar o comportamento do plano de tráfego do utilizador. Estes pontos de referência serão analisados em maior detalhe mais à frente.

[44]Service Profile Repository
[45]Online Charging System
[46]Offline Charging System

2.6.1 PCRF

O *Policy and Charging Rules Function* é o componente central da arquitectura PCC. O PCRF é a entidade responsável pela tomada de decisões de políticas de QoS a aplicar a um determinado subscritor. Esta entidade reage a eventos provenientes dos outros elementos da arquitectura e baseia as suas decisões nesses mesmos eventos. Com base na arquitectura de *policy* do IETF, o PCRF instancia um *Policy Decision Point* (PDP). O PCRF fornece à rede controlo granular no que concerne à detecção de fluxos de dados do serviço, *gating* e controlo de QoS direccionados ao PCEF. O PCRF decide qual o tratamento que deve ser dado a um fluxo de dados do serviço ao nível do PCEF e também garante que o tratamento dado ao tráfego do utilizador pelo PCEF está em conformidade com o perfil do próprio utilizador.

Para a troca de informação entre o PCEF e o PCRF, a arquitectura introduz o conceito de regra PCC. Uma regra PCC consiste num conjunto de informação que permite a detecção de um fluxo de dados do serviço e que providencia parâmetros para controlo de policiamento e para controlo de taxação.

As regras PCC podem ser dinâmicas ou estáticas. Por regras PCC dinâmicas entendem-se as regras que são provisionadas pelo PCRF com toda a informação para detecção dos fluxos de tráfego e respectivos parâmetros de QoS a aplicar aos fluxos detectados. Estas regras são dinamicamente criadas e aprovisionadas pelo PCRF ao PCEF. As regras dinâmicas podem resultar igualmente de requisitos comunicados por um AF. As regras dinâmicas podem ser alteradas a qualquer altura, desde que um evento despoletado pelo PCEF ou pelo AF assim determinem. Quando existe uma alteração a uma regra PCC dinamicamente provisionada, apenas a nova informação deve ser provisionada. O PCEF, perante a omissão de informação que tinha sido previamente provisionada, deverá manter a informação anterior. As regras PCC estáticas são regras que existem a priori num PCEF e que não podem ser de forma alguma alteradas pelo PCRF. As duas únicas operações que o PCRF pode efectuar às regras PCC estáticas é a activação ou desactivação das regras para um determinado fluxo de dados do serviço. O PCRF referencia estas regras através do nome da regra. Desta forma o PCRF terá que conhecer previamente quais as regras PCC estáticas pré-configuradas no PCEF.

Para a troca de regras PCC entre o PCRF e o PCEF existem dois modelos distintos: o modo *pull* e o modo *push*. No modo *pull*, existe um pedido de regras PCC iniciado pelo PCEF, isto é, no modo *pull*, um evento originado no PCEF resultará num pedido de regras PCC ao PCRF. O modo *push* consiste numa entrega não solicitada de regras PCC pelo PCRF ao PCEF. Esta entrega não solicitada de regras PCC pode ser resultado de um evento interno ao nível do PCRF

(por exemplo, alterações de regras condicionadas à hora do dia), ou pode ser resultado de uma comunicação de requisitos por um AF externo (por exemplo, um servidor P-CSCF a comunicar requisitos para uma chamada SIP[47]).

Quando o PCRF recebe um pedido de QoS proveniente de um AF, o PCRF deve verificar se a informação de serviço fornecida pelo AF está em conformidade com as políticas globais do operador e se o perfil do utilizador prevê a aceitação dos parâmetros de serviço enviados pelo AF. A informação então recebida pelo PCRF será utilizada para derivar os parâmetros de QoS aceitáveis para o utilizador (que poderão ser diferentes dos pedidos pelo AF) que posteriormente se converterão numa ou mais regras PCC para aplicar ao utilizador e que serão entregues ao PCEF. No caso do PCRF não aceitar os parâmetros de serviço comunicados pelo AF, este indica ao AF que os parâmetros de serviço não são aceitáveis e pode indicar na resposta quais os parâmetros que estaria disposto a aceitar.

Para o PCRF tomar as decisões de regras PCC a instalar no PCEF, poderão existir vários parâmetros provenientes dos elementos da arquitectura PCC para a tomada da decisão. Parâmetros provenientes do PCEF podem-se ter, a título exemplificativo, o identificador do subscritor, o endereço IP do UE[48], os atributos do *bearer* IP-CAN, o tipo de pedido, etc. Atributos provenientes do SPR, podem-se ter, entre outros, a lista de serviços autorizados para o subscritor, a informação de QoS autorizada para o subscritor, a categoria do subscritor, etc. Por fim, o AF também fornece informação para decisão de quais as regras PCC a activar, tal como já pode ser constado anteriormente: Identificação do subscritor, tipo de média, formato do média, largura de banda, descrição dos fluxos, estado do fluxo (*gating*), etc.

2.6.2 PCEF

O PCEF é o componente da arquitectura PCC responsável pela detecção de fluxos de dados do serviço e por garantir a aplicação de políticas. O PCEF está tipicamente localizado ao nível do *gateway* de rede (GGSN no caso do GPRS[49]) e providencia a detecção de fluxos de dados do serviço, tratamento do tráfego do utilizador, controlo da sessão e alertas via *triggers*, tratamento de QoS, assim como outras funções relacionadas com taxação que não serão consideradas nesta dissertação.

O PCEF apenas deve permitir a passagem de fluxos de tráfego que estejam sobre o seu con-

[47] Session Initiation Protocol
[48] User Equipment
[49] General Packet Radio Service

trolo e cujas regras PCC que foram instanciadas para esses fluxos permitam a passagem efectiva dos mesmos. O PCEF deve igualmente forçar o QoS a aplicar aos fluxos de tráfego, consoante indicado pelas regras PCC. Isto é, para um determinado fluxo de tráfego o PCEF deverá aplicar *rate policing*, no caso de haver um limite superior estabelecido para um fluxo de dados do serviço. Apenas quando existe a indicação de aplicação de garantias de largura de banda numa regra PCC é que o PCEF deverá iniciar um processo de reserva de recursos para cumprir a garantia de largura de banda mínima.

O PCEF deve fazer o *enforcement* da QoS de três formas:

- QoS por QCI - Aplicação de regras de QoS consoante a classe de serviço;
- QoS ao nível do fluxo (PCC) - O QoS deve ser aplicado aos fluxos determinados pelos filtros de fluxo das regras PCC, e o QoS descrito na respectiva regra deve ser aplicado;
- QoS ao nível do *bearer* - Afecta um conjunto de fluxos de dados do serviço que estejam afectos a um mesmo *bearer*.

O operador poderá instalar regras pré-definidas no PCEF, que não sejam conhecidas pelo PCRF, de forma a garantir o funcionamento do serviço, mesmo quando o utilizador não possui regras PCC afectas a si mesmo. Este tipo de regras poderão ter informação "wildcarded" de forma a serem o mais genéricas possível. No caso extremo, em que não existe qualquer regra PCC activada para a sessão IP-CAN, então a sessão deverá ser descartada pelo PCEF. A modificação de uma sessão IP-CAN deverá levar a um pedido de novas regras PCC quando o PCRF indicou explicitamente, através da instalação de *triggers* de eventos no PCEF, de que deveria ser avisado para a reavaliação de regras PCC quando esse evento específico acontecesse. Desta forma, existem modificações ao nível da sessão IP-CAN que não irão despoletar um pedido de regras PCC.

2.6.3 AF

O *Application Function* é o elemento da arquitectura PCC que permite que entidades externas solicitem requisitos de QoS, que serão por sua vez comunicados ao PCRF. O AF oferece, efectivamente, a aplicações externas a capacidade de solicitarem requisitos de serviço, que por sua vez serão passados ao elemento PCRF para avaliação dos mesmos. Esses requisitos, fornecidos pelo AF ao PCRF, serão usados para as decisões do PCRF, no que diz respeito à aplicação de

regras PCC. Durante esta fase, o AF poderá requisitar que o PCRF o avise na eventualidade da ocorrência de algum evento que afecte a sessão de dados do utilizador ligado ao AF. Depois de o PCRF avaliar os requisitos comunicados pelo AF, o PCRF poderá aceitar o pedido sem qualquer alteração ou então, no caso do pedido não ser aceitável, comunicar quais os requisitos que este estaria disposto para aceitar ao AF.

Um exemplo de um AF é o P-CSCF, elemento constituinte das redes IMS[50], que poderá querer garantir determinado nível de QoS para, por exemplo, uma chamada SIP. Como é sabido, uma chamada SIP tem requisitos de largura de banda mínima, atraso e erros bem definidos. Desta forma, estes parâmetros devem ser conhecidos para que possam ser correctamente aplicados ao nível dos equipamentos de rede, para minimizar o risco de que a conversa se possa tornar ininteligível devido a uma eventual congestão da rede.

2.6.4 Ponto de referência Gx

O ponto de referência Gx é o ponto que liga o PCEF ao PCRF, tal como pode ser observado na figura 2.9.

Este ponto permite que o PCRF tenha um controlo dinâmico sobre o comportamento do PCEF, ao nível da aplicação de regras PCC. Este ponto permite a sinalização de decisões ao nível das regras PCC, decisões essas que governam o comportamento PCC. As funções suportadas por este ponto de referência incluem:

- Pedido de decisões PCC do PCEF ao PCRF;
- Provisionamento e remoção de decisões PCC do PCRF ao PCEF;
- Comunicação de eventos ocorridos ao nível do plano de tráfego (camada de transporte);
- Negociação do modo de estabelecimento do *bearer* IP-CAN;
- Término de uma sessão Gx, pelo PCEF ou PCRF.

Uma decisão PCC consiste em zero ou mais regras PCC e respectivos atributos do IP-CAN.

[50]IP Multimedia Subsystem

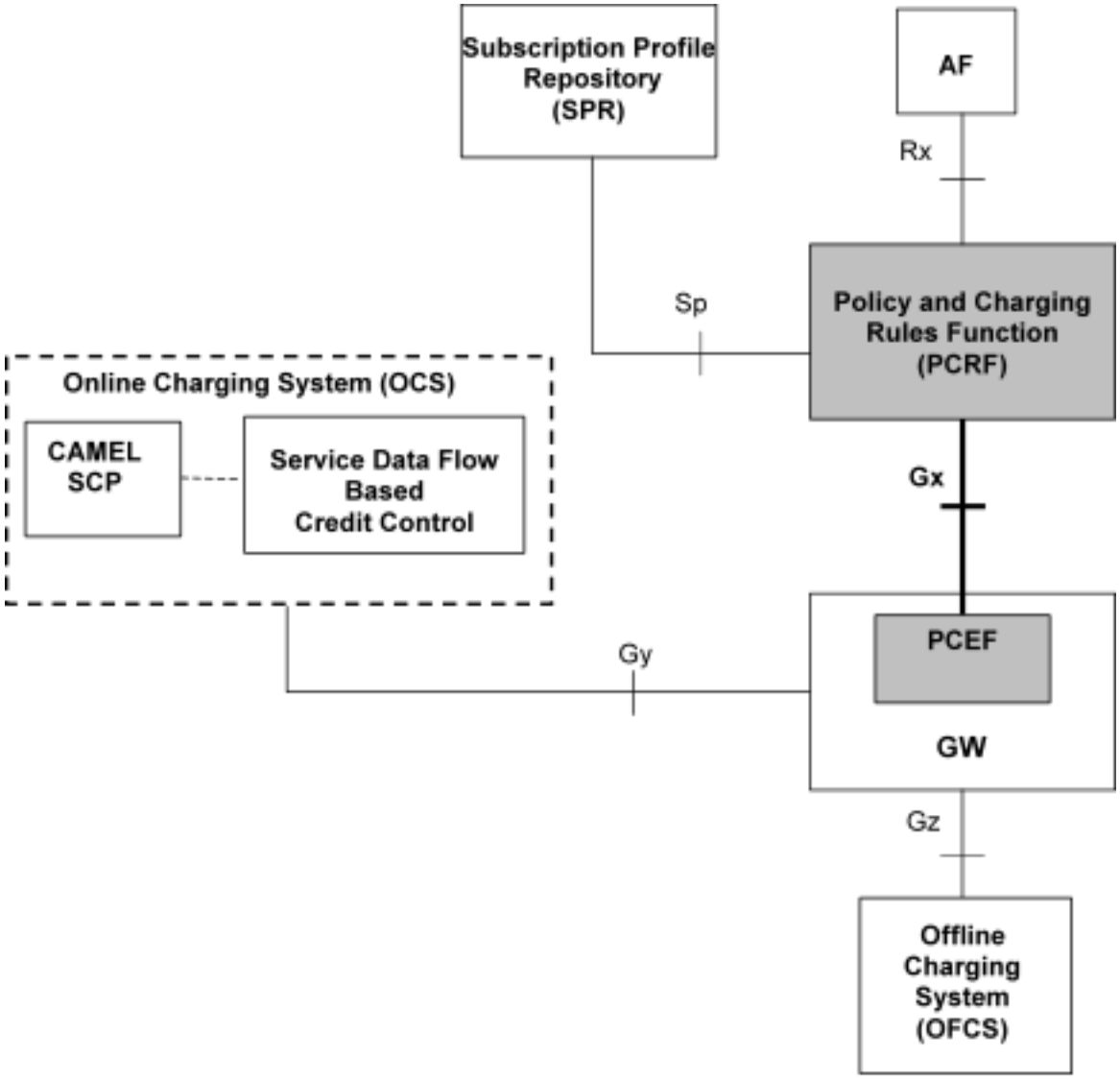

Figura 2.9: O ponto de referência Gx na arquitectura PCC

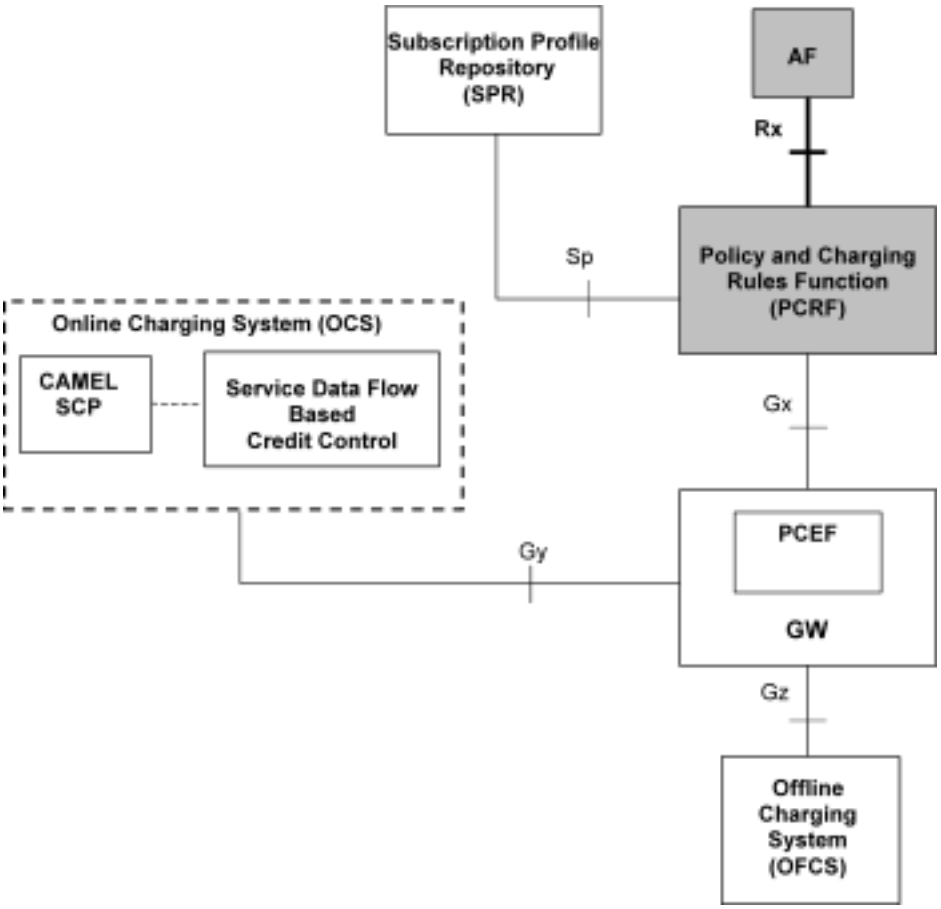

Figura 2.10: O ponto de referência Rx na arquitectura PCC

2.6.5 Ponto de referência Rx

O ponto de referência Rx é o ponto que liga o AF ao PCRF, tal como apresentado na figura 2.10.

Permite a troca de informação de sessão ao nível da aplicação entre o AF e o PCRF. Essa informação, entre outros, inclui os seguintes parâmetros: informação de filtros IP de forma a identificar os fluxos de dados do serviço e requisitos de largura de banda para a aplicação ou média para controlo de QoS.

2.7 Sumário

Neste capítulo foi apresentada e discutida a filosofia PBNM, a especificação de uma arquitectura genérica de *Policy* para controlo de admissão pelo IETF assim como também foi demonstrada a implementação concreta de uma arquitectura de *Policy and Charging* pelo organismo de normalização 3GPP e pelo ETSI-TISPAN. Deu-se igualmente bastante relevo aos componentes PCRF, PCEF e AF, que são os componentes mais importantes da arquitectura PCC e os responsáveis pela comunicação de necessidades de QoS, tomada de decisões e aplicação de políticas baseadas nas decisões tomadas pelo PCRF. O PCRF por sua vez, irá posteriormente comunicar a sua decisão aos equipamentos de rede que concretizam o elemento PCEF. Em [34] é possível ver uma tabela comparativa das diferentes implementações de *Policy Enforcement* dos diferentes organismos de normalização.

Capítulo 3

Desenho e implementação da solução

Neste capitulo serão demonstrados todos os passos conducentes ao desenvolvimento da solução proposta nesta dissertação. Começaremos por abordar a solução actual de controlo de tráfego existente na PT Inovação mostrando, de seguida, de que forma a nova solução de controlo de tráfego preconizada nesta dissertação representa uma evolução relativamente à solução actual. Será demonstrado de que forma o componente PACF concretiza um servidor de políticas genérico e de que forma a construção de um plugin RTDAP[1] e construção das respectivas políticas de suporte ao componente permitem, efectivamente, concretizar o componente PCRF[2] descrito na arquitectura PCC[3] do organismo 3GPP.

3.1 A solução actual de controlo de tráfego na PTIN

Actualmente, a PT Inovação possui uma solução de controlo de tráfego baseado em políticas que é uma funcionalidade do sistema IP-Raft. A funcionalidade *Policy Enforcement* da solução IP-Raft permite que o operador possa controlar, de uma forma integrada, a largura de banda a disponibilizar aos diferentes tipos de cliente, sejam estes pré-pagos ou pós-pagos, com origem na rede fixa ou móvel. Actualmente a solução concentra-se na utilização do equipamento Cisco SCE[4][35] que desempenhará a função de PCEF[5]. A comunicação com este equipamento PCEF

[1]Real Time Data Application Part
[2]Policy and Charging Rules Function
[3]Policy and Charging Control
[4]Service Control Engine
[5]Policy and Charging Enforcement Function

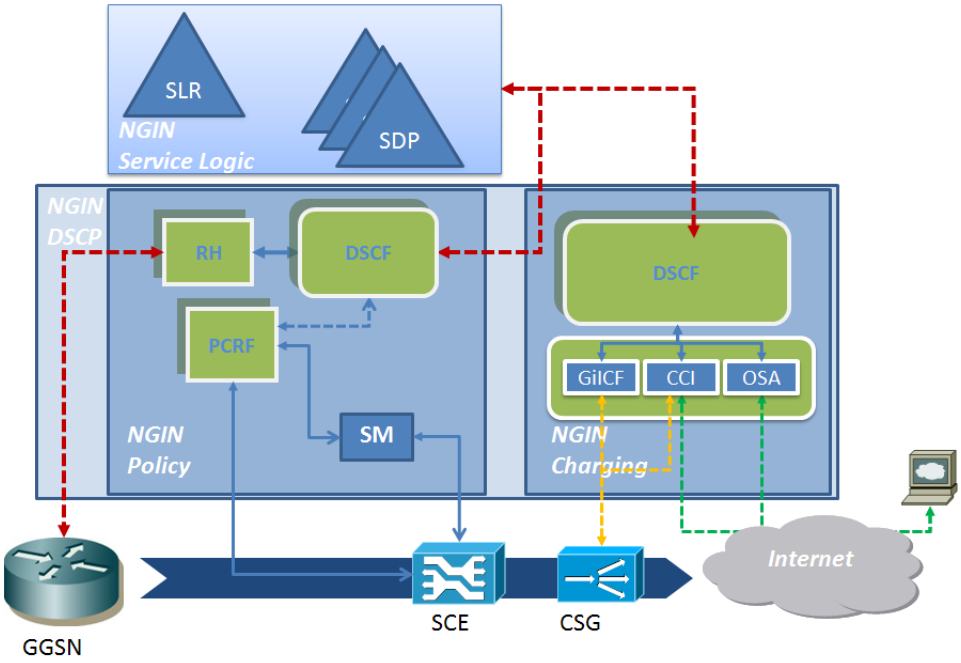

Figura 3.1: Solução actual de *Policy* na PT Inovação

é feita através de uma API[6] proprietária da Cisco, disponibilizada sob a forma de uma biblioteca Java[36]. Um esquema global da actual solução de *Policy Enforcement*, conjuntamente com a solução para tarifação de consumo de dados poderá ser observada na figura 3.1.

O Cisco SCE comunica com o elemento NGIN PCRF através da API proprietária da Cisco[36, 37], sendo o NGIN PCRF o elemento responsável pela mediação da interacção entre o Cisco SCE e a lógica de controlo de QoS[7] do operador. O NGIN PCRF não interage directamente com a lógica de controlo de QoS do operador, mas sim através do elemento DSCF[8], que é o responsável pelas interrogações necessárias para a obtenção do identificador de perfil para o respectivo utilizador. Esse identificador de perfil é determinado pela lógica de controlo de QoS do operador, através da análise da informação relativa a um utilizador da rede, como por exemplo, o MSISDN[9] e o endereço IP[10] deste. Esse identificador de perfil será posteriormente entregue pelo NGIN PCRF ao Cisco SCE que será capaz de o interpretar para determinar quais as regras de filtragem de tráfego e QoS que deverão ser aplicadas ao utilizador.

[6] Application Programming Interface
[7] Quality of Service
[8] Data Session Control Function
[9] Mobile Subscriber Integrated Services Digital Network Number
[10] Internet Protocol

A interacção entre o NGIN PCRF e o Cisco SCE poderá ser feita de duas formas distintas: comunicação directa entre eles, ou então poderá ser feita recorrendo a um mediador conhecido como Cisco SM[11]. Quando é utilizada comunicação directa entre o NGIN PCRF e o PCEF, o NGIN PCRF terá que possuir um conhecimento alargado de todos os SCEs. Por outro lado, a utilização do *Subscriber Manager* para comunicação do NGIN PCRF com o Cisco SCE, permite uma camada de abstracção adicional, no sentido em que, desta forma, o PCRF não precisa de conhecer explicitamente os equipamentos Cisco SCE que estão disponíveis. Simultaneamente, o Cisco SM permite abstrair os detalhes de configuração e gestão dos equipamentos Cisco SCE disponíveis na rede.

O Cisco SCE suporta dois modos de funcionamento distintos para autenticação dos clientes e aplicação de políticas, conferindo liberdade de escolha aos operadores, que desta forma poderão adaptar mais facilmente a solução à realidade das suas infra-estruturas. Esses dois modos são o modo *pull* e *push*. No modo *push* o componente que estabelece o diálogo com o Cisco SCE, é o responsável por fazer o provisionamento do identificador de perfil para um determinado cliente, o IP e o identificador do mesmo, sem que o Cisco SCE o tenha requisitado explicitamente. No modo *pull*, o processo de autorização do cliente é iniciado pelo Cisco SCE, ou seja, existe um pedido explícito do SCE ao NGIN PCRF para que este último lhe fornece as instruções necessárias para lidar com o tráfego do utilizador. Quando o PCRF recebe este pedido do SCE, terá que fazer as devidas diligências de forma a determinar qual o identificador de perfil a atribuir ao cliente. Em ambas as situações, o SCE perante o identificador de perfil irá aplicar ao tráfego do cliente as regras associadas a esse identificador.

Na figura 3.1 é possível igualmente verificar a existência de um elemento denominado por RH[12] que, dependendo do cenário, poderá fornecer sinalização RADIUS[13][20] ao DSCF, que o irá ajudar no fornecimento de um identificador de perfil para o utilizador que se encontra a iniciar sessão na rede.

Relativamente às questões de *Charging*, o "NGIN Charging" é um complemento à solução de Policy que implementa funcionalidades de taxação. O "NGIN Charging" irá interagir com o DSCF, possibilitando a contabilização do volume de tráfego, a respectiva taxação e mesmo a possibilidade de configuração de lógicas de negócio do operador com base no volume de dados do utilizador. Esta solução permite, entre outras, que se definam quotas de dados para os utilizadores e que, quando as quotas são atingidas, possa existir uma reacção do sistema, que pode

[11] Subscriber Manager
[12] Radius Handler
[13] Remote Authentication Dial In User Service

consistir na mudança de identificador de perfil associado ao utilizador. Para esse efeito, a solução "NGIN Charging" suporta vários equipamentos, desde que estes sejam capazes de reportar o volume de tráfego consumido por cada cliente. Um exemplo de um equipamento deste tipo é o Cisco CSG[38].

3.2 Requisitos da Solução

O desenho de uma solução de PBNM[14] tem sempre as suas especificidades. Os principais objectivos que criaram os traços gerais para o desenho da solução de *Policy* foram os seguintes:

- **Integração com os módulos do sistema DSCP**: A solução a ser desenvolvida deverá ser capaz de comunicar com alguns dos elementos do sistema DSCP[15][39], sub-sistema IP-RAFT[39], de forma a conseguir ter alguma interacção com a IN responsável pela integração da lógica de negócio do operador.

- **Normalização do comportamento do PACF com os restantes módulos DSCP**: Tendo o PACF[16] seguido outro processo de desenvolvimento, tornou-se veemente adaptar o funcionamento do mesmo para estar de acordo com o comportamento dos restantes módulos do sistema DSCP.

- **Modularidade**: Apesar do plugin desenvolvido ser direccionado para o funcionamento em ambientes DIAMETER Gx, deve possuir alguma flexibilidade para que, através da utilização de um *middleware*, possa dar resposta a pedidos endereçados por equipamentos de rede que eventualmente implementem um protocolo de comunicações diferente do DIAMETER.

- **Personalização**: As características fundamentais do produto devem ser facilmente personalizáveis, quer através de ficheiros de configuração, quer através da alterações de parâmetros de controlo na base de dados.

- **Orientado ao operador**: A tomada de decisões relativamente a regras PCC a aplicar, deverá ser sempre feita recorrendo às estruturas do PACF conhecidas como "políticas", permitindo desta forma um rápido desenvolvimento de soluções à medida do operador.

[14] Policy Based Network Management
[15] Data Service Control Point
[16] Policy and Admission Control Function

- **Restrições Temporais**: O produto deverá estar apto a aplicar diferentes regras PCC consoante o período horário, e deverá ter autonomia suficiente para perceber quando uma regra expira e quando outra deverá entrar em funcionamento, tudo de forma completamente automatizada. Esta característica permitirá, entre outras, a oferta de um serviço conhecido como *happy-hours*.

- **Regras PCC orientadas ao perfil**: A solução deverá ser capaz de relacionar um identificador de perfil com as respectivas regras PCC que a ele surgem associadas. Torna-se portanto, impreterível, que um utilizador possua sempre um identificador de perfil associado. Desta forma, é possível fazer a oferta de serviços diferenciados, personalizando o comportamento do PCEF perante o tráfego do cliente de acordo com o seu perfil.

- **Funcionamento Stand-Alone**: Este produto deverá possuir a capacidade de funcionar sem uma IN, isto é, sem ligação aos elementos que implementem a lógica de negócio do operador. Para este modo o operador deverá ser capaz de definir utilizadores junto da Base de Dados afecta à solução.

3.3 O Componente PACF

O PACF[17] é efectivamente um Servidor de Políticas Genérico[40]. O PACF é uma entidade capaz de gerir a rede, com base em políticas, que terá um comportamento dinâmico. A funcionalidade básica do PACF consiste em autorizar, ou não, a utilização de recursos da rede por parte de um utilizador, que anteriormente os tenha requerido.

O PACF é um servidor de políticas genérico orientado ao evento, o que significa que a tomada de decisões é sempre motivada pela ocorrência de um evento específico. Para decidir o que fazer perante um evento, o PACF utiliza o conceito de políticas, que são componentes do PACF que irão determinar qual o seu comportamento global. Tipicamente, a um evento, estarão associadas uma ou mais políticas, que ajudarão na tomada de decisões associadas ao evento que ocorreu. Estas políticas encontram-se definidas numa base de dados, e como tal poderão ser facilmente modificadas para dar resposta às necessidades específicas do operador.

O PACF, sendo um Servidor de Políticas Genérico, não possui nativamente qualquer capacidade de comunicação com elementos de rede externos. Essa capacidade é sempre facultada ao PACF através da construção de plugins específicos que implementem um determinado protocolo

[17]Policy and Admission Control Function

de comunicação. Para que o PACF possa utilizar os plugins, estes têm que ter uma interface bem definida, que seja reconhecida pelo mecanismo de carregamento de plugins do PACF. Através da utilização de plugins para comunicação com os elementos de rede, o PACF poderá ser adaptado para funcionar em cenários bastante distintos, desde cenários que utilizem DIAMETER[18] para comunicar com os elementos de rede, COPS[19], SNMP[20], RADIUS[21] ou um qualquer protocolo de comunicação que seja possível suportar. A adição de novas funcionalidades ao PACF não implica alterações ao nível do núcleo do programa. O seu núcleo é estático e independente dos protocolos de comunicação e das políticas a implementar. Portanto, a personalização do PACF é sempre feita através da adição de novos plugins e políticas, mas nunca alterando o seu núcleo.

Para conseguir toda esta abstracção que faz do PACF um Servidor de Políticas Genérico modular, o PACF baseia-se fortemente em modelos de dados abstractos facilmente extensíveis, que permitem suportar novas políticas sem grande esforço. A classe abstracta da qual derivam todas as políticas denomina-se por *policy*. Desta classe são estendidas as classes *policy rule*, *policy condition* e *policy action*. A classe *policy rule* agrega um conjunto de condições e acções associadas a uma política. Uma acção apenas será executada se o resultado da condição for verdadeiro. O modelo de dados das políticas pode ser visto na figura 3.2.

A grande flexibilidade do PACF reside também na capacidade de carregar plugins dinamicamente para comunicação com os elementos externos (por exemplo, *routers* ou *gateways*). Podem existir diversos plugins em simultâneo, sendo o kernel do PACF responsável pela escolha do plugin correcto para a entrega de mensagens. A entrega de mensagens aos diversos plugins é feita através de um *dispatcher*, não havendo por isso necessidade de que os plugins tenham conhecimento mútuo, o que a acontecer diminuiria fortemente a componente genérica do PACF. Nesse sentido, o PACF apresenta diversos blocos e responsabilidades que serão a seguir resumidos:

- Kernel - responsável pela interacção entre o Policy Server, a base de dados interna, o comportamento do PACF e o dispatcher de mensagens e eventos;
- Dispatcher - responsável pela recepção e envio de mensagens para os plugins correctos;
- Policy Server - responsável pelo carregamento dinâmico das políticas contidas na base de dados;

[18] O DIAMETER é um protocolo de redes de computadores para autenticação, autorização e contabilização (AAA)
[19] Common Open Policy Service
[20] Simple Network Management Protocol
[21] Remote Authentication Dial In User Service

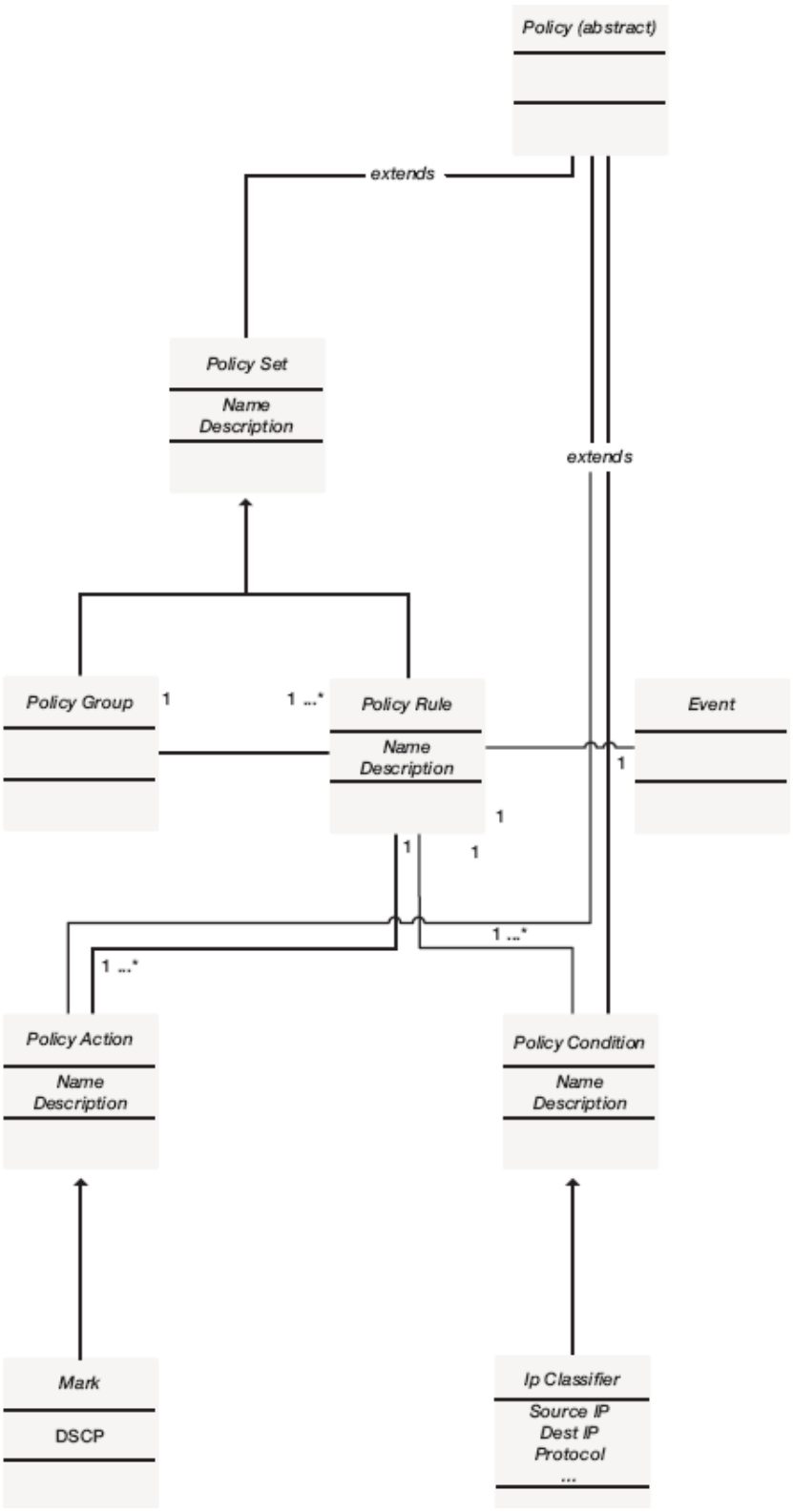

Figura 3.2: Modelo de dados estendido

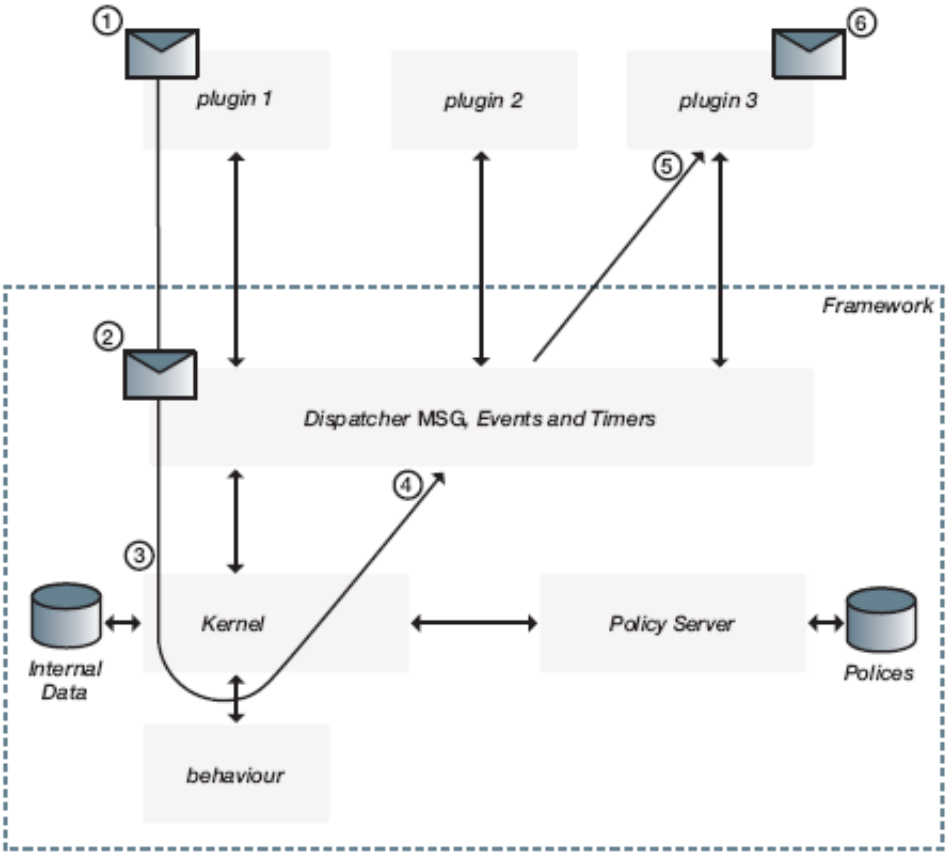

Figura 3.3: Exemplo do comportamento para uma mensagem

- Behaviour - Define o comportamento para as mensagens colocadas no dispatcher. Este parâmetro é parametrizado através da manipulação da base de dados;
- Internal Data - Base de dados que contém a informação necessária para o funcionamento do sistema. Inclui informação de sessões, utilizadores, perfis, regras associadas aos perfis, etc.

O comportamento definido pelo operador determina quais os plugins que serão contactados pelo *dispatcher* para o consumo das mensagens. Na figura 3.3 podemos ver o exemplo do fluxo de uma mensagem que chega ao PACF pelo plugin 1 e que é consumida pelo plugin 3, em que:

- 1 - Chegada de uma mensagem ao plugin 1. Validação da mensagem e construção da estrutura interna reconhecível pelo *dispatcher* e pelo kernel;

- 2 - A mensagem é colocada no *dispatcher* pelo plugin 1;
- 3 - O kernel constata que existe uma mensagem que precisa de ser processada, e com base no comportamento definido pelo operador aplica políticas e escolhe qual o plugin que irá consumir a mensagem;
- 4 - O kernel coloca a mensagem novamente no *dispatcher* a fim de ser consumida pelo plugin escolhido;
- 5 - O plugin constata que tem uma mensagem que precisa de ser processada;
- 6 - O plugin determina qual a acção a executar para a mensagem, se for caso disso, codifica a mensagem no formato da entidade que a irá receber, e finalmente entrega a mensagem (por exemplo, entrega a mensagem a um elemento de rede externo).

3.4 Enquadramento do PACF no DSCP

O PACF, por si só, não é um elemento funcional totalmente independente no que concerne ao fornecimento de regras PCC. Embora possua capacidades de funcionamento em modo independente (*Stand-Alone*), actualmente não existem equipamentos capazes de efectuar a função de PCEF que implementem o protocolo RTDAP[22] como protocolo de comunicação. Sendo assim poderemos considerar o PACF como um componente do sistema PCRF, que precisa de um *middleware* para traduzir as mensagens da rede para RTDAP e que também eventualmente precisará de um componente adicional para inquirir a lógica do operador. Nesse sentido, a solução PCRF no sistema NGIN Policy, como um todo, é capaz de fornecer a informação necessária ao PACF para que este possa tomar decisões quanto às regras PCC que deverá entregar ao PCEF para a sessão de um determinado utilizador. A arquitectura NGIN Policy com a inclusão do PACF poderá ser vista na figura 3.4.

Várias entidades podem ser identificadas nesta figura. De notar que, relativamente à figura 3.1, estes os PCRFs são diferentes do ponto de vista funcional. O NGIN PCRF é um *Gateway Protocolar*, entre o Cisco SCE e o DSCF. O PCRF da nova solução, que é composto pelo DSGW, DSCF e PACF, implementa as funcionalidades de um PCRF tal como descritas na norma 3GPP TS 29.212[33]. Dentro do PCRF, para além do PACF, destacam-se as seguintes entidades

[22]Real-Time Data Application Part

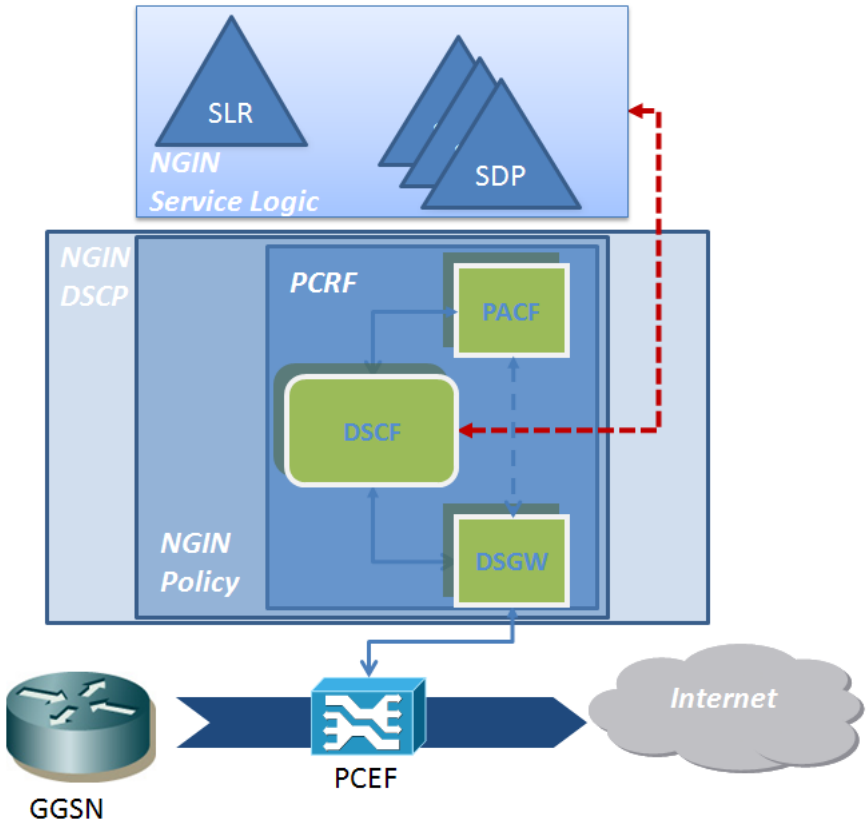

Figura 3.4: Enquadramento do PACF na arquitectura NGIN Policy

- **DSCF**: *Data Session Control Function* é o elemento central do sistema DSCP, que oferece ao sistema DSCP a capacidade de adaptação a diferentes cenários. No contexto deste trabalho, terá a função de efectuar questões à lógica do operador com o intuito de determinar qual o identificador de perfil associado a cada um dos utilizadores que inicia uma sessão de dados. É também responsável por determinar, através do recurso à lógica de controlo de QoS do operador, em que momento ou condições deverá ocorrer a mudança de perfil de um utilizador, consoante as políticas assumidas pelo operador.

- **DSGW**: *Diameter Signaling Gateway*, é um elemento que assume importância capital nesta solução. É o *middleware* responsável pela tradução de mensagens DIAMETER para RTDAP, assim como a operação inversa.

Como se pode ver na figura 3.4, a solução poderá funcionar sem a existência de um DSCF. Sem um DSCF é assumido que o PACF disporá de toda a informação necessária para relacionar os utilizadores com os respectivos perfis, de forma a fornecer uma decisão de regras PCC para entrega ao PCEF.

Um componente que faz parte da especificação da arquitectura PCC mas que não foi abordado no âmbito deste trabalho foi a componente de taxação. Embora a arquitectura PCC preveja que a componente de taxação possa ser directamente abordada no momento de entrega das regras PCC ao equipamento PCEF, na PT Inovação já existem outras soluções, nomeadamente o NGIN Charging ,que são perfeitamente capazes de efectuar essa tarefa, não havendo portanto necessidade de replicar as mesmas características nesta solução.

3.5 Desenvolvimento de um plugin RTDAP para o PACF

O PACF, pelo facto de ser um servidor de políticas genérico, não possui capacidade nativa de comunicação com elementos exteriores. Como foi anteriormente referido, para que o PACF possa comunicar com elementos de rede externos, terão que ser construídos plugins, os quais serão integrados com o mesmo. Estes plugins irão facultar ao PACF a capacidade de comunicar com os elementos de rede externos.

Dentro da arquitectura IP-Raft surgiu a necessidade de construir um plugin para o PACF, que lhe facultasse a capacidade de comunicações através do protocolo RTDAP com o elemento

DSCF[23] e/ou DSGW[24] da arquitectura IP-Raft. A opção pelo protocolo RTDAP e não pela comunicação directa em DIAMETER prende-se com vários factores. Em primeiro lugar, todo o subsistema DSCP utiliza RTDAP para a comunicação entre os diversos componentes. Em segundo lugar, a utilização do RTDAP permitiu utilizar o componente DSCF como *middleware*. Sendo o responsável pela consulta da lógica de controlo de QoS do operador para obtenção do identificador de perfil possibilita, igualmente, a introdução de várias lógicas de serviço. E em terceiro lugar, como existe um componente capaz de fazer a tradução directa entre DIAMETER e RTDAP e vice-versa (o DSGW), deixou de ser um problema o facto de que não existirem equipamentos de rede que desempenham o papel de PCEF que implementem RTDAP como protocolo de comunicações. No entanto, será de destacar a introdução do DSCF como *middleware*, que permitiu tirar alguma da complexidade associada a um PCRF isolado. Globalmente, poderemos olhar para o PACF, DSCF e DSGW como um PCRF completo. Um diagrama da solução poderá ser visto na figura 3.5.

A escolha pelo RTDAP como protocolo de comunicação entre o PACF e os restantes componentes da arquitectura tem também um outro objectivo: ao se escolher um protocolo próprio da PT Inovação que possui alguma facilidade em receber "traduções" de mensagens de outros protocolos, ganha-se automaticamente a capacidade de comunicar com outras entidades para além de entidades DIAMETER, desde que haja um *middleware* capaz de fazer a tradução. Como o plugin é configurável ao nível do tipo de parâmetros dentro de uma mensagem, desde que o *middleware* adapte as mensagens provenientes de um qualquer equipamento de rede para que estas, estejam de acordo com a estrutura que o libRTDAP espera, à partida não será necessário fazer quaisquer alterações ao código do libRTDAP e das respectivas políticas de apoio.

3.5.1 Descrição das entidades do PACF

Dentro da arquitectura do PACF, a entidade "RTDAPtoKernel" é a entidade responsável por todas as comunicações que envolvam o protocolo RTDAP. Nesse sentido, o libRTDAP possui várias entidades funcionais que desempenham funções bem específicas, tal como pode ser observado na figura 3.5.

A entidade "RTDAPtoKernel" é a responsável por gerir todas as outras entidades e por implementar a interface que permite que o libRTDAP comunique e conheça os métodos de interacção com o kernel do PACF. A entidade "RtdapHandleEvents", ao receber um mensagem proveniente

[23]Data Session Control Function
[24]Diameter Signaling Gateway

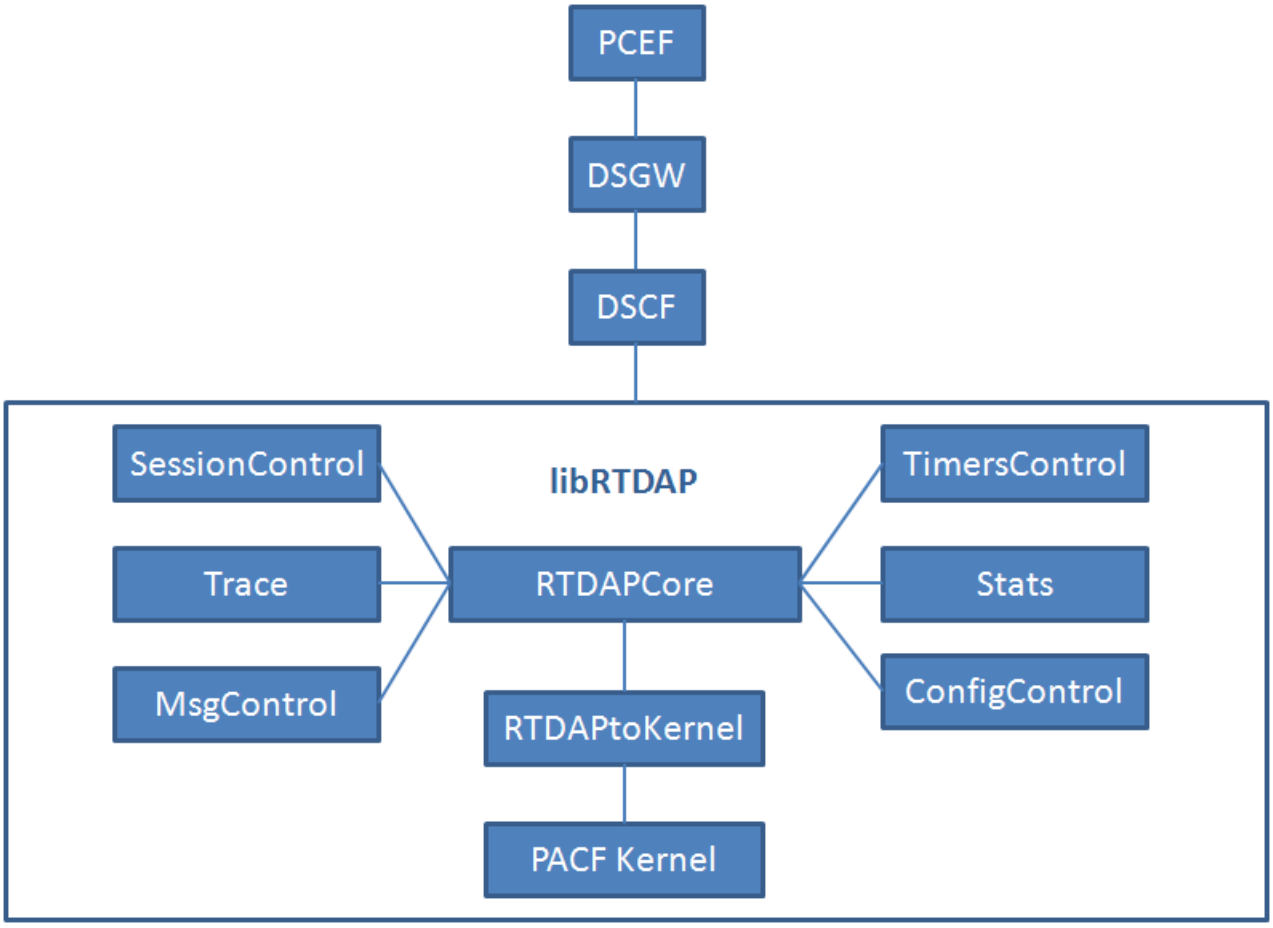

Figura 3.5: Entidades funcionais da nova solução PCRF

da rede, faz a descodificação dessa mensagem para construir uma mensagem que seja compatível com a estrutura esperada pelo kernel do PACF, coloca essa mensagem na entidade "RTDAPtoKernel", ficando depois essa entidade responsável pela colocação final da mensagem no kernel do PACF, para que a mesma seja analisada pelas diversas políticas. No entanto, para que o kernel saiba quais as políticas que deve aplicar à mensagem, a entidade "RTDAPtoKernel", antes de colocar a mensagem no kernel do PACF para processamento, irá associar-lhe um evento e um tipo de mensagem. Estes dois parâmetros, evento e tipo de mensagem, são parâmetros configuráveis no ficheiro de configuração de mensagens do libRTDAP. Portanto, quando houver necessidade de efectuar modificações a algum destes parâmetros, apenas será necessário alterá-los no ficheiro de configuração.

Existe uma entidade específica cuja tarefa é exclusivamente fazer o carregamento e validação de ficheiros de configuração. Essa entidade, ou bloco funcional, denomina-se por "ConfigControl" na arquitectura do plugin libRTDAP. Os ficheiros de configuração seguem um formato Secção e respectivo conjunto de parâmetros associados a essa mesma secção. Os parâmetros para cada secção correspondem a um par parâmetro/valor. Todos os parâmetros esperados pela entidade "ConfigControl" devem estar impreterivelmente presentes no ficheiro de configuração. Caso isto não aconteça, o carregamento do plugin pára resultando numa situação de erro, impossibilitando o arranque com sucesso do PACF. Com esta abordagem, todos os parâmetros são passiveis de alteração, fornecendo um carácter extremamente configurável à solução.

A entidade "RTDAPtoKernel" é a responsável pela mediação entre a entidade responsável pela recepção e envio de mensagens RTDAP, a "RTDAPCore", e entre a colocação e recepção das mensagens de, e para, o kernel do PACF. Este processo permite que as operações necessárias sobre as mensagens que irão permitir fornecer uma resposta à rede possam ser efectuadas.

A entidade que desempenha um papel central para o plugin libRTDAP, entidade essa que é responsável pela gestão e bom funcionamento de todas as outras entidades, é a "RTDAPCore". Esta entidade tem a tarefa de toda a gestão de comunicações RTDAP, controlo de restrições horárias, alocação de sessões, controlo do estado das sessões, descodificação das mensagens provenientes da rede, construção das mensagens para envio para a rede, validação das mensagens que são recebidas ou que estão prestes a ser enviadas, entre outras.

Num típico fluxo de mensagens, a entidade "RtdapHandleEvents" começa por receber uma mensagem proveniente da rede. Essa mensagem virá segundo o formato do protocolo RTDAP e, como tal, terá que ser analisada de forma a que se possa construir uma mensagem numa estrutura específica que será posteriormente utilizada pelo kernel do PACF. O kernel do PACF será

igualmente o responsável por, numa fase posterior, entregar a mensagem às políticas correctas que instanciou de acordo com o evento subjacente à mensagem. Antes de ser feita a descodificação completa da mensagem, terá que ser alocada um identificador interno para esta sessão, que manterá em memória os parâmetros fundamentais para a gestão da própria sessão. Para essa tarefa existe a entidade "SessionControl", que é responsável pela gestão das sessões internas do libRTDAP. Este bloco implementa um mecanismo de "Round-Robin" de forma a fornecer identificadores de sessões válidos. Quando o libRTDAP recebe um pedido de uma nova sessão (ou seja, um "INITIAL_REQUEST"), o bloco "SessionControl" procura o primeiro identificador de sessão que não esteja alocado, e atribuí esse identificador à nova sessão, passando a sessão para o estado "ALLOCATED". Quando não existem mais identificadores de sessão disponíveis, é gerada uma mensagem de erro. O número de sessões que esta entidade é capaz de controlar é manipulado por um ficheiro de configuração. Quando ocorre um "TERMINATION_REQUEST", este bloco efectua a desafectação da sessão, marcando-a como "FREE" e deixando este identificador disponível para ser aplicado a uma nova sessão.

Depois de se alocar a sessão e de se atribuir um identificador interno, será necessário fazer a descodificação concreta da mensagem. No sentido de fazer a descodificação da mensagem em RTDAP para uma mensagem numa estrutura própria do PACF, é necessário perceber à priori qual a estrutura esperada da mensagem. Teremos que ter sempre presente que o protocolo RTDAP é um protocolo sem "tags" e como tal a ordem dos parâmetros é determinante para se perceber o significado de cada parâmetro da mensagem. Para tudo isto, terá que ser utilizado um ficheiro de configuração que determina qual o formato da mensagem. Cada mensagem é identificada por um "opCode" específico que torna o processo de identificação da estrutura da mensagem relativamente simples. A cada mensagem vem igualmente associado o tipo de evento que a mensagem deverá despoletar no kernel do PACF, o tipo de operação, um temporizador e o tipo da mensagem. O tipo de operação poderá ser um "REQUEST" ou uma "ANSWER" e terá repercussões na possível aplicação de um temporizador interno. Tipicamente, quando há o envio de uma mensagem de forma não solicitada do PACF para a rede, é activado um temporizador com o objectivo de controlar o tempo de resposta da rede a esse pedido não solicitado. Desta forma, quando o temporizador expira, o estado dessa sessão entrará em erro uma vez que não recebeu a confirmação da rede num período de tempo espectável.

O evento associado a uma mensagem permitirá que o plugin libRTDAP perceba quais as políticas que deverá instanciar para tratar a mesma. O tipo de mensagem identifica que tratamento deve ser dado à mensagem junto do plugin. Nesta fase, uma mensagem poderá ser um "INITIAL_REQUEST", "MODIFICATION_REQUEST" ou "TERMINATION_REQUEST". Para

que a entidade "RTDAPtoKERNEL" possa posteriormente conhecer de que forma deve formatar a mensagem para o envio para a rede, é procurada a estrutura de uma mensagem no ficheiro de configuração das mensagens que, ao tipo de mensagem que deu origem ao pedido, tenha associado o sufixo "_RETURN" para indicar que esta é a mensagem de saída para o respectivo pedido. Desta forma, uma mensagem proveniente da rede é identificada através de um tipo comum, sendo a mensagem de resposta identificada pela adição do sufixo "_RETURN" (por exemplo, "INITIAL_REQUEST_RETURN". Seguidamente, e por limitações subjacentes ao próprio protocolo RTDAP, surge o número de parâmetros esperados. Sabendo o número de parâmetros esperados, é explicitado qual o nome das variáveis que serão usados para cada um dos parâmetros. Todos os nomes das variáveis são, nesta fase, precedidos de um "$". A ausência de um prefixo que identifique o valor como uma variável fará com que aquele valor seja interpretado como uma constante. Os nomes das variáveis serão posteriormente utilizados pelo motor de codificação/descodificação do libRTDAP de forma a fazer uma correspondência entre os parâmetros que chegam da rede e os parâmetros que deverão existir numa mensagem interna do PACF. Possui a limitação de que os parâmetros devem vir devidamente ordenados, uma vez que o protocolo RTDAP não possui suporte para "tags" que os identifiquem, sendo desta forma impossível enviar esses mesmos parâmetros seguindo uma ordem aleatória.

Havendo ocorrido a recepção de uma mensagem RTDAP, terá que haver a natural associação dos parâmetros aos respectivos nomes das variáveis. Esta tarefa é efectuada pela entidade "MsgControl". Aqui ocorrerá a descodificação da mensagem que levará a que seja criada uma estrutura interna que fará a associação das variáveis declaradas no ficheiro de configuração com o respectivo valor que lhes foi atribuído. Este passo será extremamente importante para o procedimento seguinte, que consiste em validar a estrutura da mensagem, assim como os parâmetros por ela transportados.

Depois da entidade "MsgControl" fazer a descodificação da mensagem, terá que ser feito um "parsing" dessa mesma mensagem de forma a determinar a validade da mesma. Para esse efeito, foram seguidos os princípios subjacentes a uma mensagem DIAMETER "Credit-Control-Request" para pedidos do tipo "INITIAL/MODIFICATION/TERMINATION_REQUEST". Um dos princípios associados a uma mensagem DIAMETER CCR é o de que existem parâmetros obrigatórios e parâmetros opcionais. Portanto, terão que existir dois momentos de avaliação dos parâmetros que venham na mensagem proveniente da rede: por um lado, analisar se todos os parâmetros obrigatórios estão presentes na mensagem e, por outro lado, analisar se os parâmetros não obrigatórios que se encontrem presentes estão correctamente estruturados.

Os parâmetros que têm que estar obrigatoriamente presentes numa mensagem do tipo "REQUEST" são os seguintes:

- **Session-Id**: Identificador utilizado para manter o contexto de sessão;
- **Auth-Application-Id** Identificador único que revela qual o tipo de interface envolvida (neste caso poderá ser Gx);
- **Origin-Host** Identifica o ponto de origem da mensagem;
- **Origin-Realm** Identifica o domínio de origem da mensagem;
- **Destination-Realm** Identifica o domínio de destino da mensagem;
- **CC-Request-Type** Indica qual o tipo de "Credit Control Request", que poderá ser "initial", "modification" ou "termination";
- **CC-Request-Number** Identificador único que permite fazer um mapeamento entre os pedidos e respectivas respostas.

A ausência de algum destes parâmetros numa mensagem do tipo "REQUEST" invalida imediatamente a mensagem e gera uma mensagem de erro enviada para a rede no sentido de avisar o originador que ocorreu uma situação anómala com a mensagem por ele enviada. Nesta situação o estado da sessão é actualizado para um estado de erro. Parâmetros como o "Session-Id" permitem identificar claramente uma sessão de um determinado utilizador na rede. Os parâmetros obrigatórios poderão ser usados para se fazer já uma pré-validação da mensagem. Isto é, poderá ser definido que este PCRF terá um comportamento "X" perante mensagens provenientes do domínio "A", mas poderá ter um comportamento diferente perante mensagens provenientes do domínio "B". Este comportamento poderá ser facilmente implementado através da construção de políticas personalizadas para o PACF.

Depois de se fazer a validação dos parâmetros obrigatórios que deverão surgir numa mensagem do tipo "REQUEST", é necessário proceder à validação dos parâmetros não obrigatórios. Um dos parâmetros não obrigatórios na norma que praticamente se torna obrigatório nesta implementação é o "Subscriber-Id". Este parâmetro traz o identificador do utilizador, username, que será depois utilizado internamente para manter o controlo de toda a informação associada a esse utilizador. Um parâmetro que não surge na norma mas para que a presente implementação é de importância capital, é o "Profile-Id". Quando este parâmetro surge numa mensagem

do tipo "REQUEST", significa que esse parâmetro foi obtido pelo componente DSCF através de uma consulta à lógica de controlo de QoS do operador. Este parâmetro é de importância capital para identificar quais as regras PCC que deverão ser aplicadas ao respectivo utilizador. Como se verá mais à frente durante a explicação da política "Profile Definition", a ausência deste parâmetro na mensagem em situações bem controladas será perfeitamente aceitável. Outros parâmetros não obrigatórios importantes que poderão surgir numa mensagem do tipo "REQUEST" poderão ser o "QoS-Information" que traz a informação de qual o QoS disponibilizado pela rede para o utilizador; o "IP-CAN-Rat-Type" que identifica o tipo de acesso rádio do utilizador; o "Event-Trigger" que identifica qual o evento que levou a um novo pedido de regras PCC (este parâmetro surge tipicamente como parâmetro informativo numa mensagem do tipo "MODIFICATION_REQUEST"), assim como vários outros parâmetros que terão menor importância relativamente aos parâmetros supra citados. Como explicado anteriormente, o *parsing* dos parâmetros não obrigatórios consiste em validar a boa formação dos mesmos.

As mensagens do tipo "ANSWER" são também sujeitas à validação de parâmetros. Neste tipo de mensagens, a estrutura que lhes dá origem é uma estrutura interna proveniente do PACF, depois de aplicadas todas as políticas à mensagem "REQUEST". Para as mensagens do tipo "ANSWER" existem naturalmente também dois momentos de avaliação distintos: um que valida os parâmetros obrigatórios e outro que valida os parâmetros não obrigatórios. No caso dos parâmetros obrigatórios, os parâmetros que são validados são os seguintes:

- Session-Id
- Auth-Application-Id
- Origin-Host
- Origin-Realm
- CC-Request-Type
- CC-Request-Number

Estes parâmetros já foram anteriormente explicados e como tal não é necessário abordá-los novamente. Relativamente aos parâmetros não obrigatórios, mais uma vez, a validação que se faz é puramente para verificar a boa formação dos parâmetros. No entanto, para uma mensagem do tipo "ANSWER", existem alguns parâmetros não obrigatórios que são deveras importantes: "Result-Code", para indicar qual o estado da resposta relativamente ao pedido que

foi feito, "Event-Trigger", para promover a instalação de "triggers" de eventos[25], "Charging-Rule-Remove" para promover a remoção ou desactivação de regras PCC presentemente activas, "Charging-Rule-Install" para promover a instalação de regras PCC dinâmicas ou a activação de regras PCC estáticas, o "QoS-Information" que indica qual o QoS global por classe de QoS para aquela sessão, entre outros parâmetros que não serão tão importantes quanto os atrás mencionados.

Durante a construção dos *parsers*, surgiu um desafio, relacionado com as limitações subjacentes ao protocolo RTDAP: o protocolo RTDAP é um protocolo unidimensional contrariamente ao protocolo DIAMETER que é multidimensional. As implicações que este facto representa traduzir-se-ão na capacidade que o protocolo DIAMETER tem de possuir vários AVPs dentro de outros AVPs, conhecidos como AVPs do tipo "Grouped". Outro problema que surgiu está relacionado com o facto de que alguns dos AVPs poderiam surgir um número de vezes aleatório. Dadas as particularidades do protocolo RTDAP, a aleatoriedade pura no número de parâmetros não é possível. Para colmatar estas situações, algumas medidas tiveram que ser tomadas. Relativamente à questão do número aleatório de parâmetros, esse número deixa de ser aleatório e passa a existir um limite superior fixo para os parâmetros que podem surgir numa mensagem, seja ela do tipo "REQUEST" ou "ANSWER". Entre 0 e esse limite, os parâmetros poderão surgir normalmente. O limite para cada um dos AVPs que poderão surgir em número aleatório é controlado através de configuração. Quanto aos AVPs que podem surgir dentro de outros AVPs, a solução encontrada consiste em concatenar o nome dos diferentes AVPs através da utilização de um ponto, pela ordem em que surgem agrupados, por exemplo, "QosInformation.QoSClassIdentifier". Para os AVPs do tipo "Grouped", que podem surgir em número aleatório, foi adoptada uma notação em forma de *array*, onde se identifica o número de cada AVP *grouped*, por exemplo, "SubscriberId[0].SubscriberIdNumber". Esta solução eliminou por completo o problema da tradução dos AVPs para parâmetros RTDAP levantado pela característica unidimensional deste último. Actualmente, a única limitação associada a esta abordagem é a de que os parâmetros que poderão surgir em número aleatório, terão que possuir um tecto superior que define o número máximo de ocorrências desse parâmetro.

Depois de ter ocorrido toda a validação dos parâmetros de uma mensagem, duas situações ocorrerão, consoante se trate de um "REQUEST" ou de uma "ANSWER". Se se tratar de um pedido proveniente da rede, portanto, um "REQUEST", depois do *parsing* da mensagem será

[25]Um "Event-Trigger" é um parâmetro que quando enviado para o PCEF lhe indica que, para aquela sessão, sempre que ocorrer um evento na rede que corresponda ao "Event-Trigger" instalado anteriormente pelo PCRF, o PCRF deverá ser contactado no sentido de reavaliar as regras PCC instaladas para aquela sessão em particular.

criada uma estrutura interna, denominada "msg". Esta estrutura interna transportará os parâmetros provenientes da rede, e é inteligível pelo kernel do PACF. Depois da mensagem estar no kernel do PACF, este irá entregar a mensagem às diversas políticas que irão operar os procedimentos necessários para posteriormente a mensagem poder ser entregue novamente à rede. Depois da mensagem ter passado por todos os plugins, esta será novamente entregue ao plugin que lhe deu origem, o libRTDAP. Depois de estar novamente no libRTDAP, ir-se-á proceder, dentro da entidade "RTDAPCore", ao *parsing* da mensagem proveniente do kernel, não só para verificar a validade da mensagem, mas também para construir uma mensagem RTDAP com os parâmetros que vinham na mensagem devolvida pelo kernel do PACF. Mais uma vez existe um ficheiro de configuração com a descrição do formato da mensagem a entregar à rede. O "opCode" da mensagem de saída é encontrado através de um processo em que se procura qual a mensagem que terá um tipo igual à mensagem original proveniente da rede, onde é acrescentado o sufixo "RETURN". Deve-se, porém, verificar se a operação associada a essa mensagem é do tipo "RESPONSE" (equivalente a "answer"). A mensagem de resposta, resultante de um "REQUEST" normal é denominada por "Credit Control Answer". Com todos estes mecanismos, a associação de mensagens pergunta/resposta é também controlada por configuração.

O fluxo típico de uma mensagem RTDAP desde que entra no "libRTDAP" até que sai pelo mesmo pode ser visto na figura 3.6.

A figura 3.6 representa o cenário típico de pergunta resposta da rede. Uma mensagem proveniente da rede entra no libRTDAP e é processada. Depois de processada pelo libRTDAP é colocada no kernel do PACF, para que este a possa colocar nas políticas correctas. As políticas promoverão alterações à mensagem, transformando-a, passo a passo, numa mensagem com os parâmetros que deverão ser entregues à rede como resposta ao pedido efectuado. Depois de todas as políticas terem promovido as suas alterações junto da mensagem, o kernel do PACF entregará a mensagem ao plugin original, neste caso o libRTDAP. Depois da mensagem estar novamente no libRTDAP este terá que determinar qual o "opCode" da mensagem para identificar a sua estrutura RTDAP, assim como terá que validar os parâmetros da mensagem. Depois de confirmada a boa formação da mensagem, esta é entregue à rede para que chegue ao elemento de rede responsável pela aplicação das políticas, o PCEF.

Dada a necessidade de mapear os parâmetros que vêm contidos na mensagem RTDAP para uma estrutura interna do PACF de forma a que a mensagem RTDAP seja inteligível ao mesmo, os *parsers* esperam encontrar nomes de variáveis fixos e bem conhecidos atribuídos aos diversos parâmetros provenientes na mensagem. Isto significa que existe uma tabela de equivalências

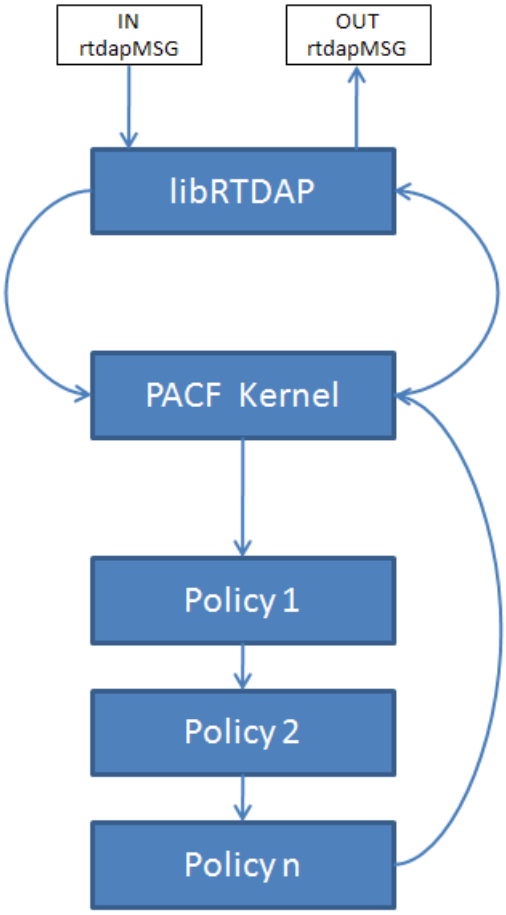

Figura 3.6: Fluxo típico de mensagens no PACF

entre o nome dos AVPs e o nome final que receberão na mensagem RTDAP (neste caso, os nomes de variáveis atribuídos aos diversos parâmetros da mensagem RTDAP no ficheiro de configuração das mensagens). Poderemos ver na tabela 3.1, alguns AVPS e respectivos parâmetros equivalentes em RTDAP.

Tabela 3.1: Exemplo de tradução de AVPs para parâmetros RTDAP

Nome do AVP	Mapeamento RTDAP
Session-Id	$SessionId
Auth-Application-Id	$AuthAppId
Origin-Host	$OriginHost
Subscription-Id.Subscription-Id-Number	$SubsId[x].SubsIdNumber
IP-Can-Type	$IPCanType

A existência de uma tabela de equivalências permite também usufruir de uma vantagem conseguida com a escolha do RTDAP para o protocolo de comunicação por excelência. Através da existência de um *middleware* que faz a tradução das mensagens de um qualquer protocolo de comunicação para RTDAP e através da existência de uma tabela de equivalências entre os parâmetros do protocolo de comunicações nativo dos equipamentos de rede e entre os parâmetros internos da estrutura do PACF, é possível obter uma solução bastante escalável e adaptativa na medida em que não está limitada a um protocolo de comunicações específico. Isto é uma mais valia para um mundo onde os processos de normalização nem sempre são rapidamente adoptados.

3.5.2 Mecanismo de controlo temporal

Uma das características mais importantes da implementação deste PCRF é o controlo temporal de regras PCC. O controlo temporal de regras PCC consiste em manter um controlo fino das regras que foram instaladas ou que aguardam o momento da activação. Na arquitectura geral do plugin libRTDAP, a entidade ou bloco responsável pelo controlo temporal das regras denomina-se por "TimersControl". Nesta entidade existem estruturas que irão manter controlo sobre quais as regras presentemente activadas, quais as regras à espera de serem activadas, e quais as regras cuja validade estará prestes a expirar. Como se poderá constatar pela afirmação anterior, existe subjacente ao controlo temporal das regras uma noção de estado. Isto é, cada regra manterá qual o próximo estado para o qual deverá transitar. Mas antes de se apresentar a questão do

controlo das regras propriamente dito, é primeiro necessário perceber de que forma a entidade "TimersControl" toma conhecimento das regras que deverá manter sobre controlo.

Como se irá constatar mais à frente, a determinação de quais as regras a aplicar à sessão de um utilizador é feita pela política conhecida como "Profile Definition". Esta política irá determinar quais as regras PCC que serão instaladas para a sessão de dados do utilizador assim como também irá determinar quais as regras que deverão ficar em espera para eventual activação futura. Estes dois tipos de regras serão, depois, entregues à entidade "TimersControl" que fará o controlo dessas regras.

Depois da entidade "TimersControl" receber as regras que deverão ser controladas para uma sessão de dados de um utilizador, o primeiro passo a efectuar é verificar se alguma das regras já se encontra presentemente sob monitoria. A entidade "TimersControl" não monitoriza cada sessão de dados independentemente, mas sim, monitoriza as regras aplicadas a todas as sessões de dados em simultâneo, uma vez que uma mesma regra é aplicada independentemente a cada sessão de dados. Isto significa que a regra conhecida por "regra1" terá um comportamento eventualmente distinto perante dois utilizadores diferentes no que concerne ao tratamento dado ao tráfego do utilizador, embora possuam o mesmo *template* de regras e as mesmas restrições temporais.

As regras que irão ser controladas, para além do nome, terão outra informação associada que será utilizada consoante o estado para o qual a regra deverá transitar no próximo evento que a afecte. Em primeiro lugar, cada uma das regras possui um campo que identifica o estado para o qual a regra irá transitar quando houver uma modificação horária que se aplique a essa regra. Isto significa que o próximo estado identifica qual a hora que será utilizada para controlar a aplicabilidade da regra. Existem três estados possíveis para uma regra: "INSTALL", "REMOVE" e "EXPIRED". O estado "EXPIRED" significa que a validade de uma regra expirou totalmente, tipicamente porque a data de término da aplicação da regra foi atingida, e como tal a regra deixará de ser controlada e, caso esteja afecta a alguma sessão, será removida. Os outros dois estados, dependendo da data de aplicabilidade da regra, significam, respectivamente, que a regra no próximo evento deverá ser instalada ou removida. Portanto, quando o próximo estado para uma regra é "INSTALL", a hora para aplicabilidade da regra a controlar será a hora de início da regra (start hour). Caso o próximo estado da regra seja "REMOVE", então a hora da regra que será controlada será a hora de fim da regra (stop hour). De momento, o controlo horário apenas tem granularidade à hora. Para além de controlo horário, as regras estão também sujeitas a controlo por data. Isto significa que uma regra apenas será controlada caso a data actual esteja no intervalo de início e fim de data de aplicabilidade da regra. Se uma regra a ser controlada

ultrapassar a data de aplicabilidade, isto é, o "stop date" for ultrapassado, a regra é considerada expirada e como tal é agendada para remoção de todas as sessões que a possuam. Existe um outro parâmetro associado aos parâmetros das regras, que é a identificação de que se trata de uma regra estática ou dinâmica. Este parâmetro é importante na medida em que, permite saber com exactidão quantas consultas será necessário fazer à base de dados de forma a entregar as regras PCC à rede. Desta forma, para uma regra estática, não é necessário fazer consultas à BD[26] uma vez que, para uma regra estática, toda a informação que é necessário saber é o nome da regra. Por outro lado, se se tratar de uma regra dinâmica, já será necessário fazer consultas à BD com o intuito de determinar quais as características associadas a essa regra PCC, nomeadamente filtros de tráfego e QoS a aplicar ao tráfego identificado por esses filtros.

Quando uma regra é afectada, isto é, é determinado que essa regra deverá ser instalada ou removida (poderá ser mais do que uma, pois não existe um limite para o número de regras que poderão ser afectadas em simultâneo), é determinado o que se deve fazer com essa regra (gerar o AVP Charging Rule Remove, ou então o AVP Charging Rule Install, dependendo de qual é o estado da regra), a BD é consultada para determinar quais as sessões que serão afectadas pela mudança de estada da(s) regra(s), e é gerada uma mensagem RAR[27] que levará os respectivos AVPs de forma a assinalar as alterações necessárias à sessão a que a mensagem RAR se destina. Cada RAR já leva todas as alterações necessárias à sessão como consequência da mudança horária, isto é, caso mais do que uma regra seja afectada que possa gerar vários AVPs Charging Rule Install, para a mesma sessão, apenas uma mensagem RAR é enviada de forma a minimizar possíveis sobrecargas na rede.

3.6 Desenvolvimento das políticas de suporte ao plugin RT-DAP

Na nomenclatura do PACF, políticas são porções de código que irão analisar uma mensagem proveniente de um PCEF e irão tomar decisões em concordância com o estabelecido pelo operador. Por exemplo, uma política poderá consistir em fazer o registo do utilizador na base de dados do PACF. As políticas a aplicar a uma determinada mensagem são deduzidas com base no evento associado à mensagem. Desta forma, na base de dados de políticas do PACF, a um determinado evento estarão associadas as políticas que deverão ser dinamicamente instanciadas. Depois

[26] Base de Dados
[27] Re-Auth-Request

de instanciadas, as diferentes políticas terão funções distintas perante a mensagem recebida da rede. Sendo um sistema modular e dinâmico, novas políticas poderão ser criadas para modificar o comportamento associado a um evento.

O mecanismo de políticas consiste fundamentalmente em analisar o evento que deu origem à mensagem proveniente da rede, consultar a base de dados e verificar quais as políticas que o PACF deverá instanciar para esse evento. A ordem pela qual as políticas são registadas na base de dados para um determinado evento é determinante para a aplicação das mesmas pelo PACF. Uma vez que as políticas poderão ter dependências das políticas anteriores, este é um factor determinante para o correcto comportamento global do PACF perante uma mensagem proveniente da rede. Este mecanismo de políticas permite uma grande modularidade e permite, igualmente, que soluções adaptadas para o operador possam ser facilmente criadas, sem necessidade de efectuar alterações ao núcleo do PACF ou do próprio plugin libRTDAP. E aqui reside o verdadeiro poder do mecanismo de políticas. Quando um operador tem uma qualquer necessidade específica, por exemplo, um operador decide que quer ter controlo rigoroso de quais os domínios de origem aceitáveis para uma sessão de dados, poderá optar por criar uma política cujo objectivo primordial é verificar o domínio de origem do utilizador e validá-lo consoante as suas necessidades. Outra possibilidade consiste em realizar controlo de admissão do utilizador à rede. Através de uma política, o operador poderá implementar um mecanismo de CAC[28] consoante as suas necessidades, onde vai controlando o QoS atribuído a um utilizador, podendo chegar a um ponto em que o utilizador já não terá autorização para fazer pedidos de mais recursos. Em suma, o mecanismo de políticas é extremamente poderoso e versátil, permitindo adaptar o comportamento global do PACF para variados cenários, consoante as necessidades do operador.

No âmbito deste projecto foram criadas duas políticas distintas: "Register User" e "Profile Definition". Estas duas políticas serão explicadas em maior detalhe nas secções seguintes.

3.6.1 Register User

A primeira política a ser desenvolvida denomina-se "Register User". Tal como o nome indica, este política será responsável pela gestão do registo de novos utilizadores na base de dados, que iniciam sessão na rede. Este política tem dois modos de funcionamento distintos que se traduzem num comportamento totalmente distinto da política. Esses modos são o modo persistente ou o modo não persistente. De acordo com o tipo de modo de funcionamento escolhido,

[28] Call Admission Control

o registo/remoção do utilizador da BD terá um comportamento distinto. Em qualquer um dos modos de funcionamento, é guardada alguma informação base relativa ao utilizador, informação essa que assume um carácter possivelmente menos temporário, no sentido em que não é tão susceptível a mudanças no decorrer da sessão. Essa informação consiste tipicamente no nome de utilizador, identificador do perfil de utilizador, tipo de acesso IP-CAN[29], informação sobre a persistência ou não do utilizador (por omissão, os utilizadores não são guardados em modo persistente), informação sobre a conectividade do utilizador (este campo é particularmente importante para as situações em que um utilizador é persistente e a política está a operar em modo persistente), o QoS inicial fornecido pela rede, o(s) endereço(s) IPv4 e/ou IPv6 e a data/hora actual do início de sessão do utilizador na rede. De notar que, por motivos de confidencialidade, não será incluído nesta dissertação o modelo de dados de suporte desta política.

O modo persistente assume-se como uma decisão de desenho essencial para esta política. Fundamentalmente, o modo persistente consegue fazer com que o PACF seja completamente autónomo de uma IN[30] do operador. Desta forma, todas as tomadas de decisão são feitas a nível local do PACF. O modo persistente, como regista os utilizadores novos que vão iniciando sessão na rede, permite que, quando o utilizador iniciar uma nova sessão na rede e mesmo para um cenário em que tenha sido perdida a ligação com a IN, o PACF forneça uma decisão de quais as regras PCC e respectivo QoS a aplicar ao cliente. No modo completamente *stand-alone*, onde nunca existiu uma interacção com uma IN, o operador terá que definir os utilizadores manualmente na Base de Dados. Neste modo, é fundamental a existência de um identificador de perfil associado ao utilizador, com o intuito de correctamente determinar as regras PCC e QoS a aplicar a esse mesmo utilizador.

A política "Register User" espera, potencialmente, três eventos distintos para uma sessão:

- *INITIAL_REQUEST*
- *MODIFICATION_REQUEST*
- *TERMINATION_REQUEST*

Uma nova sessão deverá sempre iniciar com o evento "INITIAL_REQUEST", e deverá ser descartada sempre que não cumpra este requisito. Perante o evento "INITIAL_REQUEST" a política "Register User" terá um comportamento diferente consoante esteja a funcionar em modo

[29]IP-Connectivity Access Network
[30]Intelligent Network

persistente ou em modo não persistente. Se estiver a funcionar em modo persistente e caso o utilizador não exista na Base de Dados, o utilizador será inserido na BD de forma persistente. Esta operação é igual para o modo não persistente, exceptuando o facto de que o utilizador é inserido em modo não persistente (um utilizador é persistente ou não consoante o valor de uma *flag* na linha de registo do utilizador). Quando a funcionar em modo persistente e verificando-se que o utilizador já existe na BD, aquilo que se efectua consiste em actualizar a informação relevante do utilizador, nomeadamente o identificador de perfil, o QoS e o endereço atribuído, a rede IP-CAN de acesso e a hora actual, isto é, praticamente toda a informação precisa de ser actualizada. Se a mensagem enviada pela rede não trouxer um identificador de perfil associado (acontece quando não existe uma IN associada) então o identificador de perfil já disponível na BD é o que será utilizado para efeitos de atribuição de regras PCC e QoS.

Alterações à informação de uma sessão actualmente em curso apenas serão efectuadas com o evento "MODIFICATION_REQUEST". Nesta fase inicial, apenas está previsto que a modificação de uma sessão consiste tipicamente na alteração do identificador de perfil do utilizador. Esta alteração levará a que nas políticas que venham a ser instanciadas posteriormente a esta política, a informação de regras PCC e QoS a atribuir tenha que ser reavaliada.

Por fim, existe o evento "TERMINATION_REQUEST". Este evento ocorre quando um utilizador termina a sua sessão de dados. Perante um pedido de término de sessão, o comportamento da política "Register User" é diferente consoante o modo de funcionamento da política seja persistente ou não persistente. No caso de a política estar a operar em modo persistente, o utilizador não será removido, mas o seu estado será actualizado para "offline". Todos os outros dados são mantidos intactos. Caso o política se encontre a operar em modo não persistente então toda a informação associada ao utilizador é removida da BD. Um pedido de término de sessão apenas é aceite e processado caso o utilizador se encontre actualmente em estado "online".

3.6.2 Profile Definition

A política "Profile Definition" é a política responsável pela consulta da BD de forma a associar regras PCC ao identificador de perfil atribuído ao utilizador. A BD possui um modelo de dados no qual, para cada identificador de perfil, estão associadas regras PCC e QoS global a aplicar à sessão de um utilizador. De notar que, por motivos de confidencialidade, não será incluído nesta dissertação o modelo de dados de suporte desta política. Cada regra PCC caracteriza-se por possuir filtros para detecção de tráfego, QoS a aplicar ao tráfego detectado por essa regra,

qual a precedência dessa regra sobre as outras (ordenação por prioridade), qual o estado a aplicar à regra (se deve autorizar ou negar os fluxos identificados pelos filtros de tráfego) e quais as restrições horárias a aplicar à mesma regra.

Ao nível da QoS granular por cada regra PCC, essa QoS apenas será aplicada aos fluxos identificados pelos filtros de tráfego da regra. Desta forma, por cada identificador de perfil, a existência de várias regras permite ter QoS personalizada para fluxos de tráfego distintos, permitindo uma grande granularidade no tratamento que é dado ao tráfego individual do utilizador. A existência de restrições horárias permite implementar, entre outros, as típicas "happy hours", mas não ficando restringido a isso. As restrições horárias também possuem restrições de data, permitindo, por exemplo, activar uma regra PCC numa determinada data, durante um período de tempo bem definido. Quando a BD é inquirida acerca de quais as regras que deve activar, as regras recebidas para activação serão aquelas cujas restrições temporais assim o permitam. Estas regras serão posteriormente associadas a uma mensagem do tipo RAR, que será entregue à rede para que as regras PCC sejam activadas para a sessão correcta. As regras cujas restrições horárias façam com que sejam colocadas em modo "stand-by" serão posteriormente entregues a uma estrutura do libRTDAP que fará o controlo dessas regras. Desta forma permite determinar o período de activação das mesmas, assim como controlo da eventual necessidade de desactivar regras que estejam presentemente activas para uma ou mais sessões consoante o período de validade dessas regras.

As regras PCC devolvidas depois da consulta da BD podem assumir dois tipos distintos, tal como previsto na arquitectura PCC: estáticas ou dinâmicas. As regras estáticas são relativamente simples de activar uma vez que aquilo que é necessário fornecer à rede é apenas o nome da regra. No entanto, para as regras estáticas são precisos alguns cuidados, nomeadamente garantir que as regras estáticas já estão definidas no equipamento PCEF, e que o nome indicado pelo PCRF à rede corresponde ao nome atribuído à regra no PCEF. No caso das regras PCC dinâmicas, a rede terá que ser previamente construída indicando os filtros de fluxos de dados do serviço, para detecção granular de tráfego, a QoS a aplicar ao tráfego detectado pelos filtros, a prioridade da regra e o estado da *gate*. A regra terá obrigatoriamente que ter um nome associado e que seja único. O nome de uma regra dinâmica poderá ser igual ao de uma regra estática caso aquilo que se pretenda seja que a regra dinâmica se sobreponha à regra estática (as regras dinâmicas têm preferência sobre as regras estáticas quando estas possuem o mesmo nome).

Depois de determinadas quais as regras a instalar para uma determinada sessão, essa informação será posteriormente entregue à rede, chegando ao PCEF, que será a entidade responsável

por aplicar as regras ao tráfego do utilizador associado à respectiva sessão. De forma a manter um controlo mais apurado do estado do utilizador, as regras que foram instaladas e/ou que foram colocadas em *standby* são também colocadas em BD numa tabela de sessões. Uma vez que o PACF já possuía uma tabela de sessões para a interface Rx/Gq', essa tabela foi estendida de forma a suportar informação de sessão IP-CAN, tal como descrita para a arquitectura PCC. Desta forma, para cada sessão IP-CAN fica associada uma série de informação, nomeadamente o identificador da sessão, o "application id", o domínio e máquina de origem da mensagem, o domínio de destino da mensagem, o identificador único do cliente, as regras que foram instaladas para a sessão do utilizador, a informação de QoS global da sessão do utilizador e informação adicional utilizada para a manutenção interna da informação de sessão, mais concretamente o identificador interno do libRTDAP para a sessão. Desta forma, quer as políticas quer o librtdap são capazes de manter a informação de sessão actualizada durante o período de existência da mesma.

Tal como a política "Register User", a política "Profile Definition" espera três estados ou eventos distintos por cada sessão: "INITIAL_REQUEST", "MODIFICATION_REQUEST", "TERMINATION_REQUEST". O comportamento da política para cada um dos estados é distinto. Para um "INITIAL_REQUEST" a política regista toda a informação de sessão necessária, tal como anteriormente explicado. Para um "MODIFICATION_REQUEST" esta política não só é responsável por fazer a actualização da informação de sessão, assim como também é responsável pela identificação das alterações que são necessárias efectuar à sessão junto do PCEF. Quando existe uma "MODIFICATION_REQUEST" onde ocorreu uma alteração do identificador de perfil associado ao utilizador, alguns passos têm que ser seguidos:

- Identificar quais as regras actuais da sessão que serão afectadas e prepará-las para eventual remoção junto do PCEF;
- Identificar quais as regras a aplicar à sessão com base no novo identificador de perfil;
- Enviar os respectivos "Charging Rule Install" e "Charging Rule Remove" para actualizar o comportamento do PCEF perante a sessão do utilizador.

Depois de seguidos estes passos, o PCEF já estará a aplicar as regras PCC correspondentes ao identificador de perfil associado, tendo as regras do perfil anterior deixado de ser aplicadas.

Por fim, para o estado "TERMINATION_REQUEST" a política limita-se a remover toda a informação da sessão que está agora a terminar. Como um "TERMINATION_REQUEST" é um pedido proveniente da rede e final, sendo também um pedido meramente informativo afim do

PCRF actualizar a informação interna que mantinha para o estado da sessão, a política não terá que fornecer informação específica para remoção de regras, sendo o PCEF o responsável por identificar as regras PCC associadas à sessão que agora se encontra a terminar.

3.7 Sumário

Neste capítulo foi demonstrado o trabalho desenvolvido para criar uma implementação do PCRF. Através da utilização de um servidor de políticas genérico, da criação de um plugin que fosse capaz de interpretar e perceber o protocolo RTDAP, da criação de políticas que determinam o comportamento do servidor genérico perante as mensagens recebidas pelo plugin RTDAP e, em conjunto com os componentes DSCF e DSGW, foi efectivamente possível criar uma implementação do componente PCRF.

Capítulo 4

Cenários de utilização e Testes

Neste capítulo serão apresentados alguns cenários de utilização do componente PACF[1] em conjunto com o plugin RTDAP[2] assim como testes aos nível da aplicação de regras PCC e respectivo resultado da aplicação dessas regras. Os cenários que seguidamente se apresentarão são cenários que estarão muito próximos daquilo que a solução fará na realidade, próximo do que será pedido por um qualquer operador de telecomunicações. No campo dos testes propriamente ditos, serão analisados os resultados obtidos através da realização de testes sintéticos, uma vez que não foi possível ter acesso a equipamento real em tempo útil.

4.1 Cenários de Utilização

Seguidamente serão demonstrados alguns cenários de utilização possíveis com a solução desenvolvida. De referir que, tipicamente, estes cenários surgem numa sequência lógica, ou seja, princípio, meio e fim. Por exemplo, terá que existir sempre um início de sessão e respectivo término. Poderá, no decorrer de uma sessão, ocorrer a necessidade de fazer alguma modificação às condições da sessão, pedido esse que poderá ser originado no PCEF[3] ou dentro da própria solução PCRF[4].

Olhando um pouco para a arquitectura da solução, a entidade PCRF é constituída pelo

[1]Policy and Admission Control Function
[2]Real-Time Data Application Part
[3]Policy and Charging Enforcement Function
[4]Policy and Charging Rules Function

DSGW[5], DSCF[6] e PACF, que aqui são abstraídos sob o nome comum PCRF. Para cada um dos cenários de utilização serão explicadas, com algum detalhe, as diferentes fases do diagrama de sequência de cada cenário.

4.1.1 Início de sessão

O diagrama de sequência típico para o início de sessão de um utilizador na rede poderá ser visto na figura 4.1.

Como se pode verificar na figura 4.1, o evento que desencadeia um inicio de sessão IP-CAN[7] é um pedido de estabelecimento do primeiro *bearer* IP-CAN. Este pedido é endereçado ao PCEF, que por sua vez irá desencadear um pedido de regras PCC[8] direccionado ao PCRF. O pedido ao PCRF é feito através de um "CC-Request", onde o tipo de pedido é estabelecida para "INITIAL" pelo PCEF. O PCEF enviará toda a informação necessária para a tomada de decisão de regras PCC ao PCRF. Analisando agora os elementos constituintes do PCRF, sabe-se que este é constituído pelo DSGW, DSCF e PACF. Numa primeira fase, o PCEF liga-se ao DSGW. A função do DSGW será fazer a tradução das mensagens DIAMETER para mensagens em RTDAP, e vice-versa, quando a comunicação se efectua no sentido contrário. Seguidamente o DSGW passará os parâmetros já traduzidos para RTDAP ao elemento DSCF. Este elemento terá uma importância capital para este cenário. Quando o DSCF recebe a mensagem proveniente do DSGW, e se estiver configurado para isso, irá questionar a lógica de controlo de QoS[9] do operador sobre qual o identificador de perfil que deverá ser atribuído ao utilizador que está a iniciar a sessão. O DSCF recebe a resposta da lógica de controlo de QoS do operador e incorpora esse identificador na mensagem que irá seguir para o PACF. De notar que o DSCF poderá não fazer nenhuma inquirição à lógica do operador caso o produto esteja a funcionar em modo *stand-alone* ou mesmo nem sequer estar presente. Nesta situação, a decisão de qual o identificador de perfil a aplicar ao utilizador será sempre tomada pelo PACF. Depois de entregue a mensagem ao PACF propriamente dito, há uma primeira fase de descodificação de mensagem, sendo depois essa mensagem colocada numa fila especial do PACF para que esta possa ser analisada pelas políticas configuradas pelo operador.

Como se pode verificar pelo diagrama de sequência, a primeira fase, depois da mensagem

[5]Diameter Signaling Gateway
[6]Data Session Control Function
[7]IP-Connectivity Access Network
[8]Policy and Charging Control
[9]Quality of Service

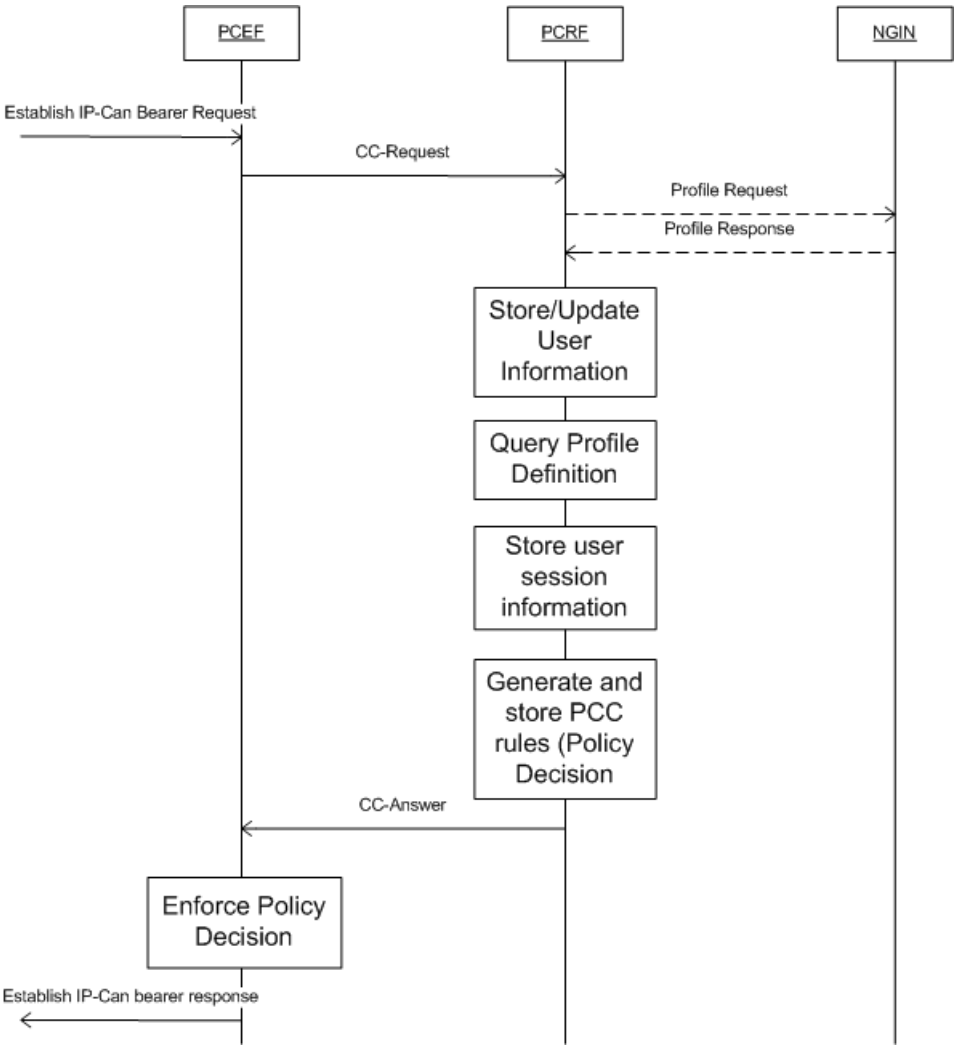

Figura 4.1: Diagrama de sequência do início de sessão de um utilizador na rede

estar colocada no kernel do PACF, corresponderá ao "Store/Update User Information". Esta fase passa-se toda na política "Register User" e o comportamento da política irá depender da configuração global estabelecida pelo operador. Como este produto poderá existir sem a presença de uma IN[10] e poderá funcionar em modo persistente, no que concerne ao registo dos utilizadores, haverá a necessidade de já estarem os utilizadores definidos na BD[11] e, como tal, apenas deverá ser feita uma actualização à informação do utilizador que se considere relevante. Dependendo do modo de funcionamento da solução, a informação do utilizador poderá ser mantida ou não, após o término da sessão do mesmo.

A segunda fase corresponde a inquirir a BD no sentido de determinar qual a informação de perfil associada ao utilizador. Esta fase passa-se na política "Profile Definition" e nesta altura tem que estar impreterivelmente presente um identificador de perfil associado ao utilizador. Desta forma, após a estrutura que mantém toda a informação necessária à tomada de decisão relativamente às regras PCC a aplicar ao utilizador deixar a política "Register User", trará sempre um identificador de perfil para o utilizador. A não existência deste identificador traduzir-se-á num erro grave. Com este identificador de perfil procede-se, seguidamente, a uma consulta à BD com a finalidade de recolher as regras para activação.

Depois de determinadas quais as regras a aplicar, é guardada a informação de sessão do utilizador onde, entre outros, se incluem a lista de regras activadas e/ou à espera de processamento para esse utilizador. Poderão existir regras à espera de processamento devido ao mecanismo de controlo temporal implementado na solução, que permite o agendamento para activação e/ou desactivação de regras pelo PCRF no equipamento PCEF. De referir que as regras PCC associadas a um identificador de perfil são completamente configuráveis pelo operador. A sua estrutura segue de perto as indicação dadas para o AVP[12] Charging-Rule-Definition nos documentos de especificação da norma[33]. De igual forma, são suportadas regras dinâmicas e regras estáticas. De lembrar, no entanto, que as regras estáticas têm que já existir previamente configuradas no PCEF. A existência destas estruturas para guardar as definições das regras PCC que compõe os perfis de utilizador tornará simples, mas bastante poderoso, todo o processo de criação de regras PCC à medida, nomeadamente no que concerne às capacidades de personalização das regras PCC, tendo impacto na facilidade com que são definidos os diferentes perfis de utilizador.

De referir que no momento do registo da sessão do utilizador, para além das regras, também é registado qual o valor do QoS global atribuído à sessão, e é também registado o somatório da

[10]Intelligent Network
[11]Base de Dados
[12]Attribute-Value Pair

QoS garantida que possa ter sido fornecida por alguma das regras PCC. O registo da QoS garantida permitirá, futuramente, implementar mecanismos para controlo de admissão. Também é registado o conjunto de *Event Triggers* eventualmente instalados para o utilizador. Estes *triggers* são muito importantes pois permitem saber quais os eventos que darão origem a um novo pedido de regras PCC pelo PCEF, no decorrer de uma sessão.

Depois de obtida a informação de regras PCC e QoS associada ao identificador de perfil, é finalmente gerada a informação que será entregue à rede para que as regras seleccionadas para a sessão do utilizador sejam aplicadas com sucesso. Nesta fase, o PCEF, depois de receber as indicações de regras PCC entregues pelo PCRF, deverá garantir que o tráfego do utilizador é sujeito à acção dessas regras. Depois disto, o PCEF poderá fazer a confirmação do estabelecimento de sessão do utilizador e, como tal, o utilizado poderá a partir daí utilizar os recursos da rede de acordo com as regras PCC instaladas no PCEF.

4.1.2 Modificação da sessão

Este cenário de utilização diz respeito à modificação de uma sessão em curso. No decorrer de uma sessão IP-CAN, poderá ocorrer algum evento que irá provocar um novo pedido de regras PCC. Isto significa que o evento que deu origem ao pedido poderá ter feito com que a informação anteriormente utilizada para derivar as regras PCC se tenha tornado obsoleta. O diagrama de sequência típico para uma modificação de sessão poderá ser visto na figura 4.2.

Para que haja um novo pedido de regras PCC ao PCRF, o PCEF terá, primeiro, que determinar quais as condições que deram origem às regras actualmente em utilização já não se verificarem completamente. Havendo a determinação dessas condições, o PCEF envia um "CC-Request" do tipo "Modification" para o PCRF (Um "CC-Request" é muitas vezes abreviado para CCR[13]). Seguindo a lógica do diagrama de sequência, dentro do PCRF, a entidade DSCF irá inquirir a lógica de controlo de QoS do operador no sentido de determinar se existe a necessidade de modificar as regras PCC do utilizador. Tipicamente esta alteração será comunicada através da alteração do identificador de perfil do utilizador. Através da análise dos novos dados enviados pelo PCEF, a lógica de controlo de QoS do operador analisará qual o melhor identificador de perfil para o utilizador em questão. De referir que, perante a ausência de uma IN, essa decisão terá que ser tomada pelo PACF. Havendo a análise pela IN, a mensagem seguirá então para o PACF. Mais uma vez, depois da descodificação da mensagem, esta será entregue à política "Register

[13]Credit Control Request

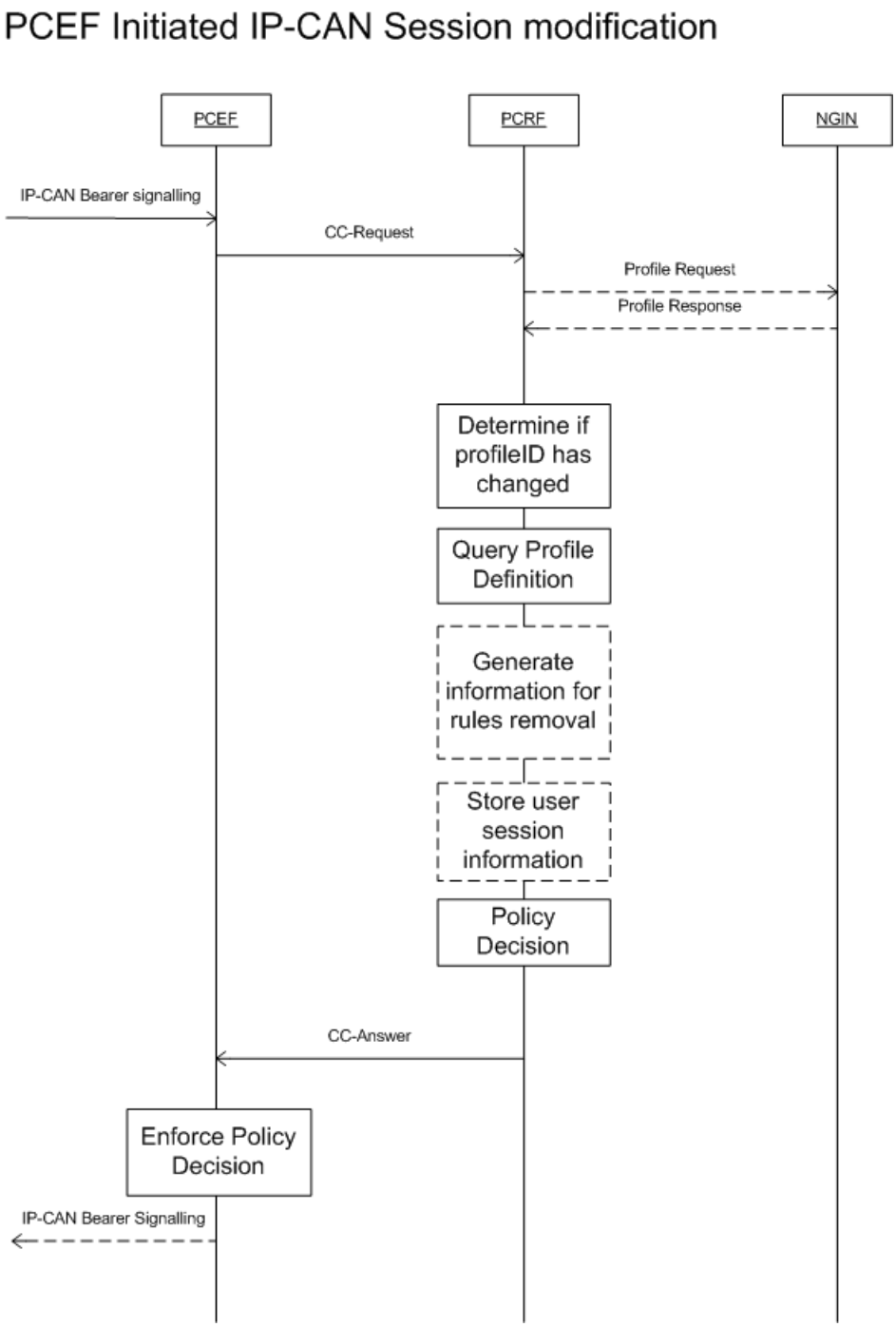

Figura 4.2: Diagrama de sequência da modificação de uma sessão de um utilizador na rede

User". Aqui será feita a actualização da informação relevante do utilizador. Eventualmente, o PCRF irá verificar se o evento que deu origem ao "Modification Request" foi requisitado pelo próprio PCRF, mas tipicamente esta verificação deverá ser sempre feita pelo PCEF, isto é, o PCEF não deverá reportar eventos que não tenham sido explicitamente requisitados pelo PCRF.

Depois de ter sido feita a actualização da informação de registo do utilizador, a mensagem com a informação da sessão será entregue agora à política "Profile Definition". Uma das primeiras verificações que é feita é sobre a eventual alteração do perfil de utilizador. Caso tenha ocorrido uma alteração de perfil, que é de facto aquilo que será expectável quando é comunicada uma modificação na sessão do utilizador, alguns passos terão que ser efectuados. O primeiro passo será determinar quais as regras que deixaram de fazer sentido para o utilizador actual. A actualização do identificador de perfil irá eventualmente fazer com que as regras associadas ao perfil anterior deixem de ter validade. Desta forma, recorrendo ao estado da sessão do utilizador, são determinadas quais as regras que não fazem parte do novo perfil atribuído ao utilizador, e como tal deverão ser marcadas para remoção. Portanto, numa primeira fase determinam-se quais são as regras para remoção. Depois de determinadas as regras que serão removidas, será necessário consultar a BD para determinar quais as novas regras associadas ao novo identificador de perfil que deverão ser instaladas no equipamento PCEF. Determinadas as regras a instalar e/ou a remover, terá que haver o natural registo da nova informação da sessão na tabela de sessões da BD. Terá que haver também uma actualização do QoS garantido atribuído para efeitos da controlo de admissão. Os mecanismos de controlo de admissão mais avançados serão implementados numa versão posterior da solução.

Depois de determinadas as regras a instalar/remover e feitas as devidas alterações à BD, haverá a entrega da mensagem novamente ao plugin libRTDAP, onde será feita a validação da mensagem para entrega à rede. A mensagem gerada será um "CC-Answer" do tipo "Modification", onde irá a informação das regras a instalar e/ou remover. Esta mensagem poderá eventualmente levar também informação de "Event-Triggers" novos para instalar no PCEF, caso essa necessidade seja determinada pelo PACF[14]. Depois do PACF enviar a mensagem CCA[15] para o PCEF, este terá que garantir a aplicação das novas regras ao tráfego do utilizador, para que o tratamento dado ao tráfego deste esteja de acordo com o perfil atribuído.

[14]Esta funcionalidade será implementada numa versão futura do PACF
[15]Credita Control Answer

4.1.3 Modificação da sessão iniciada pelo PCRF

Este cenário de utilização retrata uma entrega de regras PCC à rede sem que esta as tenha explicitamente requisitado. Este tipo de cenário apenas é possível para uma sessão em curso, ou seja, pelo menos terá que ter havido um "CC-Request" do tipo "Initiation" que deu início a uma sessão IP-CAN para que este tipo de mensagem seja passível de ser utilizado. A entrega de regras PCC não solicitadas à rede é possível através da utilização do modo *push*, tal como explicitado na norma[33]. O diagrama de sequência típico para uma modificação de sessão iniciada pelo PCRF poderá ser visto na figura 4.3.

Como se pode verificar pelo diagrama, o evento que despoleta os procedimentos relativos a uma modificação de sessão iniciada pelo PCRF é o *trigger* interno no PCRF. Este *trigger* interno poder-se-á dever a várias situações, mas na versão actual da solução está geralmente relacionada com a ocorrência de um evento horário. Como já foi anteriormente explicado, o PACF implementa um mecanismo de eventos horários responsável pelo controlo de políticas baseado em hora/data. Este mecanismo permite que uma regra PCC apenas seja aplicada durante um determinado período de tempo, assim como permite agendar a aplicação futura de outras regras PCC. Também permite agendar a remoção dessas mesmas regras.

Numa modificação de sessão iniciada pelo PCRF, depois de ocorrer um evento dentro do próprio PCRF que motiva essa modificação, ocorrerão um certo número de verificações de forma a determinar quais os procedimentos a tomar. Depois da ocorrência do evento, que na solução actual corresponderá a um evento horário, é necessário verificar quais são as sessões afectadas por esse evento. Isto é importante para identificar quais as sessões que irão ser alvo de actualização de informação. Na solução actual, verificar quais são as sessões afectadas consiste em determinar quais as sessões que tinham a regra que transitou de estado devido a um evento horário. Depois de determinada essa informação será necessário gerar as novas regras para instalação e/ou remoção, para serem entregues posteriormente ao equipamento PCEF. As regras que sofrerão alterações são então determinadas e a informação de sessão dos diversos utilizadores que possam ter sido afectados é actualizada, nomeadamente as regras activas/não activas para cada sessão, assim como o QoS garantido atribuído para cada sessão. Em simultâneo, a regra que foi afectada pelo evento horário irá sofrer uma transição de estado para o próximo estado espectável da regra, que poderá ser "INSTALL", "REMOVE" ou "EXPIRED". Seguidamente é gerada a mensagem que será entregue à rede sob a forma de um "Re-Auth-Request". Esta mensagem, também conhecida como "RAR", leva toda a informação necessária para a instalação e/ou remoção de regras PCC, consoante anteriormente determinado, assim como também poderá levar

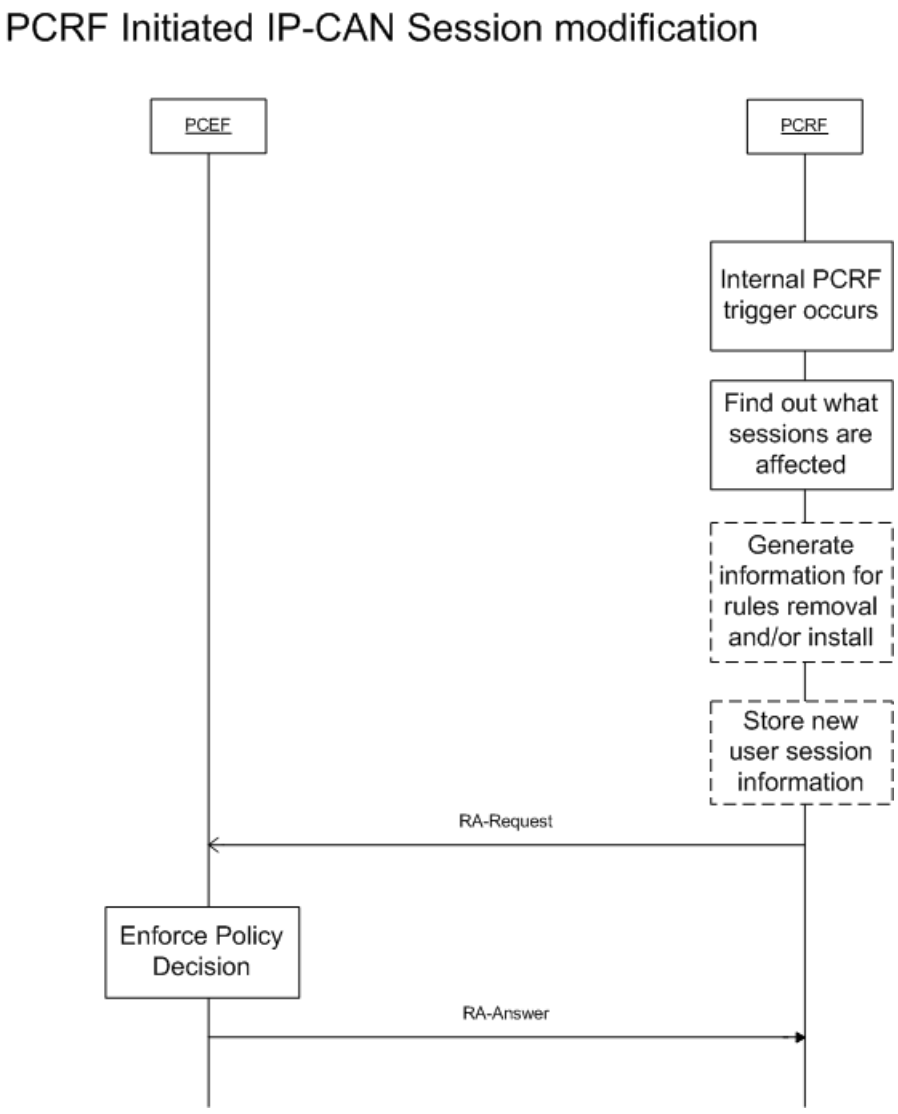

Figura 4.3: Diagrama de sequência da modificação de uma sessão de um utilizador iniciada pelo PCRF

informação de *Event Triggers*. O PCEF recebendo esta mensagem deverá proceder à instalação das regras para a sessão afectada de forma a garantir que o tratamento do tráfego do utilizador está de acordo com o perfil a ele atribuído. O PCRF inicia, entretanto, um temporizador para aguardar a resposta do PCEF de que este recebeu e aplicou as regras PCC, tal como instruído pelo PCRF. Ocorrendo a aceitação do RAR[16] e a aplicação das respectivas regras, o PCEF envia a mensagem RAA[17] a confirmar a boa recepção e aplicação das regras, ou então, poderá também ser usada para reportar erros, como por exemplo, reportar porque razão alguma das regras não foi eventualmente aceite. Um exemplo prático para utilização deste mecanismo poderá ser a implementação concreta das denominadas *happy hours*.

De referir por fim que, para esta solução, não existe intervenção directa das chamadas "políticas" do PACF. No entanto, o mecanismo de controlo temporal é bastante personalizável, tendo suporte para múltiplas restrições horárias por regra, conferindo-lhe uma grande flexibilidade. O mecanismo de restrição horária foi concebido já a pensar numa abrangência mais generalista.

4.1.4 Término da sessão

Este cenário representa uma situação de término de sessão. Muito genericamente, este cenário consiste num aviso enviado pelo PCEF ao PCRF de que a corrente sessão irá terminar. Este aviso tem o objectivo de dar tempo ao PCRF para que este possa tomar todas as devidas diligências para manter o modelo de sessões na BD consistente. O diagrama de sequência típico para o evento término de sessão é o que se pode verificar na figura 4.4.

[16]Reauthentication Request
[17]Reauthentication Answer

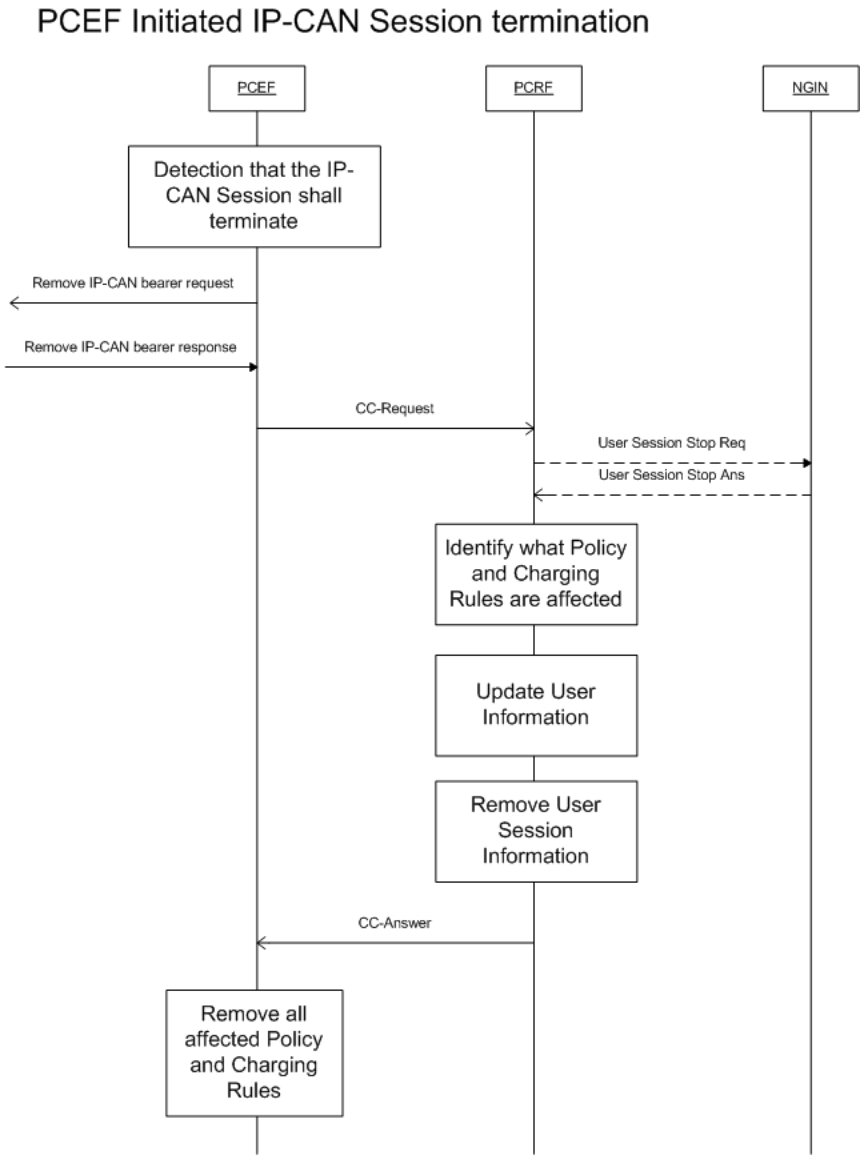

Figura 4.4: Diagrama de sequência do término de uma sessão IP-CAN

Como se pode ver neste cenário, o PCEF começa por determinar que a actual sessão IP-CAN deverá terminar, ocorrendo depois a remoção dos bearers IP-CAN da sessão. Seguidamente, o PCEF contacta o PCRF alertando-o para o término da sessão IP-CAN através de um CC-Request do tipo "Termination". Neste caso, este pedido será então tratado pelo PACF. O PACF recebendo o pedido de término de sessão irá instanciar as políticas necessárias para tratar esse pedido. Neste caso e de acordo com a solução desenvolvida, serão novamente instanciadas as duas políticas mais comuns, a "Register User" e a "Profile Definition". Numa primeira fase irá

haver o tratamento da informação do utilizador. A forma como a informação do utilizador será abordada está directamente relacionada com o método de funcionamento do PACF, isto é, ser persistente ou não. Caso o PACF se encontre em modo persistente, a informação do utilizador não será removida, sendo a sua conectividade IP[18] actualizada para "offline". Desta forma, toda a informação relevante para que se possa determinar o perfil do utilizador numa futura conexão onde não haja ligação a uma IN é mantida. Caso o PACF esteja a funcionar em modo não persistente, então, toda a informação do utilizador é removida quando é recebida a indicação de término de sessão.

Numa segunda fase, será consultada a política "Profile Definition". Aqui o primeiro ponto a verificar é, quais as regras PCC que serão afectadas. Este ponto poderá eventualmente ter impacto para o mecanismo de restrições temporais, uma vez que algumas das regras poderão não estar afectas a mais nenhuma sessão, não fazendo sentido continuar a haver controlo das mesmas. Depois de determinadas quais as regras afectas e de comunicada essa situação aos componentes relevantes, terá que ocorrer a remoção da informação de sessão do utilizador. Este passo é muito importante, uma vez que irá garantir a integridade da BD de sessões. Quando um utilizador termina a sessão, a sua informação de sessão deverá ser sempre removida. Caso isto não aconteça, numa futura nova sessão poderão ser criadas inconsistências (QoS atribuída mal calculada, por exemplo).

Em paralelo com os dois passos anteriores, o PCEF trata de remover a informação das regras PCC aplicadas à sessão do utilizador que agora se encontra a terminar. Nesta situação o PCEF não precisa que o PCRF lhe envie essa informação uma vez que o PCEF mantém o mapeamento de quais as regras que estão afectas à sessão que se encontra em estado de término. No entanto, o PCEF aguarda a confirmação do PCRF de que a informação de estado da sessão e do utilizador foi correctamente removida e/ou actualizada pelo PCRF. Depois destes passos considera-se a sessão IP-CAN do utilizador terminada com sucesso.

4.2 Teste da solução

De seguida serão mostrados alguns testes sintéticos que foram feitos à solução. Infelizmente, como não foi possível ter acesso a um PCEF em hardware real, foi montado um cenário de testes que utiliza geradores de mensagens para se testar as funcionalidades do produto. Foram testados dois cenários fundamentais:

[18] Internet Protocol

- Cenário 1: Este cenário envolve os três tipos de CC-Request possíveis, "Initial", "Modification" e "Termination". Este cenário consiste basicamente no envio de uma mensagem CCR-Initial, seguida de uma CCR-Modification onde ocorreu a mudança de perfil, finalizando com um CCR-Termination, onde se dá o término da sessão.

- Cenário 2: Neste cenário ir-se-á testar a funcionalidade das mensagens RAR, que nesta solução se encontram intimamente ligadas com o mecanismo de restrições horárias.

De referir também que a dupla PACF e DSCF incorporam um mecanismo de *keep-alive* através da troca periódica de mensagens. Este mecanismo permite garantir que as duas entidades se encontram a comunicar com sucesso. Na figura 4.5 pode ser visto um exemplo do registo da troca de mensagens do mecanismo *keep-alive*.

```
2009-10-07 16:51:30.015 librtdap     INF 1 | RP: ==> PING
2009-10-07 16:51:30.015 librtdap     INF 1 | RP: <== PONG
2009-10-07 16:51:33.184 librtdap     RUN S | libRtdap is running...
```

Figura 4.5: Mecanismo de *keep-alive*

4.2.1 Cenário de Testes 1 - CCR & CCA

Este cenário de testes envolve as mensagens CCR e CCA. Será demonstrada uma sessão que passa pelos três possíveis estados, desde o estabelecimento da sessão, até uma modificação da sessão que implica a alteração do identificador de perfil, até ao término da sessão.

CCR&CCA-Initial

O envio de uma mensagem CCR-Initial marca o estabelecimento de uma nova sessão. A mensagem terá que trazer a informação necessária para que essa sessão seja tratada de forma unívoca. Nesta fase, o PCEF envia uma mensagem para o PCRF indicando o início de uma nova sessão. No caso desta solução em particular, o agente que recebe a mensagem proveniente do PCEF é, em última instância, o PACF. Um exemplo de uma mensagem proveniente de um PCEF com destino ao PACF pode ser vista na figura 4.6.

```
2009-10-07 16:51:28.768 librtdap     INF 2 | We have received a message...
2009-10-07 16:51:28.768 librtdap     INF 2 | Received a new session
2009-10-07 16:51:28.769 librtdap     INF 1 | RECV: ===> RP_OPER[0] CallId[3], OpCode = [11]

        ------------------------------------------------------------------------------
                 OperatioName[GxInitialRequest]
        ------------------------------------------------------------------------------
         RP     | IN  |   0 | $SessionId       | 16781342
         RP     | IN  |   1 | $AuthAppId       | 16777238
         RP     | IN  |   2 | $OriginHost      | DSGW.gxinterfaceRTDAP.ptin
         RP     | IN  |   3 | $OriginRealm     | pe.teste
         RP     | IN  |   4 | $DestinationRealm| pe.ptin
         RP     | IN  |   5 | $CCRType         | 1
         RP     | IN  |   6 | $CCRNumber       | 10
         RP     | IN  |   7 | $SubsId[0].SubsId| 0
         RP     | IN  |   8 | $SubsId[0].SubsId| 9600101000
         RP     | IN  |   9 | $FrammedIPAddr   | 192.168.1.2
         RP     | IN  |  10 | $QoSInfo.QoSClass| 0
         RP     | IN  |  11 | $QoSInfo.MaxReqBa| 999999
         RP     | IN  |  12 | $QoSInfo.MaxReqBa| 888888
         RP     | IN  |  13 | $ProfileID       | 1
        ------------------------------------------------------------------------------

2009-10-07 16:51:28.771 librtdap     INF 1 | Got a new message from the itQueue
2009-10-07 16:51:28.771 librtdap     INF 1 | INITIAL_REQUEST
2009-10-07 16:51:28.771 librtdap     INF 1 | InitialRequest
```

Figura 4.6: Exemplo de um CCR-Initial

Um dos parâmetros que permitirá que o libRTDAP proceda à descodificação da mensagem é o "OpCode". O "OpCode" permite que o libRTDAP carregue, a partir do ficheiro de configuração, as variáveis RTDAP para que estas sejam confrontadas com os parâmetros da mensagem que entretanto chegou da rede. Naturalmente, todas as variáveis RTDAP correspondentes aos AVPs obrigatórios estão presentes na mensagem RTDAP. Esta situação será análoga para todas as outras mensagens. Consegue-se determinar que se está perante um CCR-Initial uma vez que a variável RTDAP $CCRType vem com o valor 1, que é um mapeamento directo do valor do AVP CC-Request-Type. Outra informação relevante que é possível encontrar nesta mensagem de exemplo é o identificador do utilizador, o seu endereço IP, a QoS requisitada pela rede para uma classe de QoS específica e o profile-ID a atribuir ao utilizador. De referir que o profile-ID é uma implementação interna e não faz parte das mensagens CCR padrão. A informação proveniente na mensagem, nomeadamente o profile-ID, será utilizada para derivar a QoS a aplicar ao utilizador e respectivas regras PCC. Para determinar então a QoS e regras PCC a aplicar ao utilizador, a mensagem que foi analisada pelo libRTDAP terá que ser também analisada pelas diversas políticas do PACF para que essa informação seja determinada. De forma a que o PACF

saiba quais são as políticas a aplicar à mensagem deste utilizador, o PACF terá que saber qual o evento despoletado por esta mensagem. Esta informação também se encontra em ficheiro de configuração, e pode ser determinada através do "OpCode" da mensagem. Na figura, é possível verificar que o evento despoletado por esta mensagem é o "INITIAL_REQUEST" (o nome particular do evento não é importante, sendo, isso sim, muito importante, as políticas configuradas na BD que estão associadas ao evento em questão). Na BD, este evento está configurado para instanciar três políticas: "Register User", "Profile Definition" e "Action Null". As primeiras duas políticas já foram anteriormente explicitadas em grande detalhe. A política "Action Null" até agora não sofreu qualquer tipo de explicação particular uma vez que a sua utilidade é de amplitude reduzida. Esta política limita-se a entregar a mensagem ao plugin que lhe deu origem. Na figura 4.7 podemos ver um registo do kernel do PACF a instanciar as diversas políticas para serem aplicadas a esta mensagem.

```
2009-10-07 16:51:28.775 PACF_KERNEL  INF 5 | Executing the following action: RegisterUser
2009-10-07 16:51:28.775 PACF_KERNEL  INF 5 | Execute action: RegisterUser
2009-10-07 16:51:28.809 PACF_KERNEL  INF 5 | Executing the following action: ProfileDefinition
2009-10-07 16:51:28.809 PACF_KERNEL  INF 5 | Execute action: ProfileDefinition
2009-10-07 16:51:28.854 PACF_KERNEL  INF 5 | Executing the following action: ActionNull
2009-10-07 16:51:28.854 PACF_KERNEL  INF 5 | Execute action: ActionNull
```

Figura 4.7: Políticas a serem instanciadas pelo kernel

Apesar de não haver muita informação relevante, é possível ver o PACF a instanciar as diversas políticas, pela ordem indicada na BD, para que estas possam analisar a mensagem proveniente da rede e, dessa forma, possam tomar decisões relativas às regras PCC a activar. Apenas haveria mais informação por parte das políticas caso acontecessem situações anómalas durante o tratamento das mensagens. Depois de analisada pelas diversas políticas, a mensagem será novamente entregue ao plugin librtdap para que seja feita a validação da mensagem e respectiva conversão para parâmetros RTDAP. Um exemplo de uma mensagem CCA-Initial pode ser vista na figura 4.8.

```
2009-10-07 16:51:28.872 librtdap     INF 1 | SEND: <=== RP_OPER[1] CallId[3], OpCode = [12]

-------------------------------------------------------------------------------
         OperatioName[GxInitialResponse]
-------------------------------------------------------------------------------
 RP    | OUT |   0 | $SessionId       | 16781342
 RP    | OUT |   1 | $AuthAppId       | 16777238
 RP    | OUT |   2 | $OriginHost      | dscp1
 RP    | OUT |   3 | $OriginRealm     | ptinovacao.pt
 RP    | OUT |   4 | $CCRType         | 1
 RP    | OUT |   5 | $CCRNumber       | 10
 RP    | OUT |   6 | $ResultCode      | 2001
 RP    | OUT |   7 | $EventTrigger[0] |
 RP    | OUT |   8 | $EventTrigger[1] |
 RP    | OUT |   9 | $ChgRuleInst[0].C| rule1
 RP    | OUT |  10 | $ChgRuleInst[0].C|
 RP    | OUT |  11 | $QoSInfo[0].QoSCl| 1
 RP    | OUT |  12 | $QoSInfo[0].MaxRe| 4096
 RP    | OUT |  13 | $QoSInfo[0].MaxRe| 1024
 RP    | OUT |  14 | $ChgRuleInst[1].C|
 RP    | OUT |  15 | $ChgRuleInst[1].C|
 RP    | OUT |  16 | $ChgRuleRem[0].Ch|
 RP    | OUT |  17 | $ChgRuleRem[0].Ch|
-------------------------------------------------------------------------------
```

Figura 4.8: Exemplo de um CCA-Initial

Como se pode ver na mensagem, mais uma vez os parâmetros obrigatórios estão todos presentes. Em relação aos parâmetros não obrigatórios, nota-se que nem todos estão presentes. Isto não representa de forma alguma um problema, e o elemento responsável pela tradução de RT-DAP para DIAMETER, o DSGW, consegue interpretar perfeitamente esta situação, ignorando os parâmetros vazios (ou seja, não converte os parâmetros vazios para AVPs). Nesta resposta, poderemos verificar que o CCR-Initial foi processado com sucesso uma vez que leva o parâmetro "$Result-Code" com o valor 2001. Consegue-se verificar igualmente que será instalada uma regra PCC, denominada "rule1", que é uma regra estática. É também possível verificar quais os valores de QoS que serão aplicadas à sessão, sendo neste caso determinados os valores máximos de *download e upload* para QoS de classe 1 (trata-se meramente de um exemplo). Estes valores foram determinados com base no identificador de perfil do utilizador, que internamente é utilizado pela política "Profile Definition" para determinar quais as regras PCC a aplicar à sessão e respectiva QoS. Esta informação será posteriormente convertida pelo DSGW para DIAMETER, sendo a partir daí entregue ao equipamento PCEF que se encarregará de aplicar as regras PCC ao tráfego do utilizador, assim como a QoS global determinado para a sessão. Em simultâneo, internamente, a entidade "TimersControl" poderá ser contactada para fazer o controlo das regras PCC que serão instaladas para a sessão.

CCR&CCA-Modification

O envio de uma mensagem CCR-Modification determina que houve alguma alteração à sessão, o que irá eventualmente determinar que as regras PCC actualmente aplicadas à sessão já não são válidas. Como explicado anteriormente, uma mudança às condições da sessão actual irão, por ventura, provocar uma alteração ao identificador de perfil do utilizador. A figura 4.9 mostra o resultado do envio de uma mensagem CCR-Modification por parte da rede.

```
2009-10-07 16:51:29.100 librtdap     INF 2 | We have received a message...
2009-10-07 16:51:29.101 librtdap     INF 1 | RECV: ===> RP_OPER[6] CallId[3], OpCode = [17]

        ------------------------------------------------------------------------------
                 OperatioName[GxModificationRequest]
        ------------------------------------------------------------------------------
         RP    | IN  |   0 | $SessionId       | 16781342
         RP    | IN  |   1 | $AuthAppId       | 16777238
         RP    | IN  |   2 | $OriginHost      | itanium
         RP    | IN  |   3 | $OriginRealm     | pe.ptin
         RP    | IN  |   4 | $DestinationRealm| pe.ptin
         RP    | IN  |   5 | $CCRType         | 2
         RP    | IN  |   6 | $CCRNumber       | 11
         RP    | IN  |   7 | $SubsId[0].SubsId| 0
         RP    | IN  |   8 | $SubsId[0].SubsId| 9600101000
         RP    | IN  |   9 | $FrammedIPAddr   | 192.168.1.2
         RP    | IN  |  10 | $QoSInfo.QoSClass| 0
         RP    | IN  |  11 | $QoSInfo.MaxReqBa| 999999
         RP    | IN  |  12 | $QoSInfo.MaxReqBa| 888888
         RP    | IN  |  13 | $ProfileID       | 2
        ------------------------------------------------------------------------------

2009-10-07 16:51:29.102 librtdap     INF 1 | Got a new message from the itQueue
2009-10-07 16:51:29.102 librtdap     INF 1 | MODIFICATION_REQUEST
2009-10-07 16:51:29.102 librtdap     INF 1 | InitialRequest
```

Figura 4.9: Exemplo de um CCR-Modification

Como se pode verificar em relação à figura 4.6 que representa o resultado do um CCR-Initial, não existem alterações significativas relativamente à estrutura da mesma. Importa também frisar que os parâmetros recebidos nas mensagens são configuráveis, isto é, o plugin librtdap tem capacidade para tratar todos os parâmetros RTDAP correspondentes a todos os AVPs estabelecidos na norma TS 29.212. As duas alterações significativas que se poderão constatar entre a mensagem CCR-Initial e esta CCR-Modification dizem respeito às variáveis "$CCRNumber" e "$ProfileID". A variável "$CCRNumber" na mensagem CCR-Modification foi naturalmente incrementada em uma unidade relativamente à mensagem anterior. No entanto, o parâmetro que vai determinar as alterações às regras PCC aplicadas à sessão é o parâmetro correspondente à

variável "$ProfileID". Verifica-se que o Profile-ID foi alterado do identificador *1* para o identificador *2*. Internamente, o profile-ID 2 tem associada uma única regra PCC, denominada "rule2".

Não será mostrada novamente a sequência de instanciação das políticas por parte do PACF uma vez que a figura 4.7 continua a ser válida. Aquilo que é possível dizer é que a política "Register User" irá fazer o registo das alterações da informação do utilizador e que a política "Profile Definition" será a responsável pela construção das regras PCC. No caso da política "Profile Definition", nesta situação em que é confrontada com um "MODIFICATION_REQUEST", a política terá que analisar quais as regras PCC actualmente instaladas para esta sessão e que deixam de ser válidas e quais as novas regras PCC que deverão ser instaladas para a sessão. Através da consulta da informação de sessão e das regras associadas ao identificador de perfil *2*, a política é capaz de determinar que a regra que deverá ser removida da sessão actual é a regra "rule1" e a nova regra que deverá ser instalada para a sessão é a regra "rule2". A política altera a mensagem de forma a reflectir estas alterações e posteriormente a mensagem é devolvida ao plugin libRTDAP para ser entregue à rede. A mensagem CCA-Modification que será entregue à rede poderá ser vista na figura 4.10.

```
2009-10-07 16:51:29.154 librtdap     INF 1 | SEND: <=== RP_OPER[7] CallId[3], OpCode = [18]

        ------------------------------------------------------------------------------
                 OperatioName[GxModificationResponse]
        ------------------------------------------------------------------------------
         RP     | OUT |    0 | $SessionId       | 16781342
         RP     | OUT |    1 | $AuthAppId       | 16777238
         RP     | OUT |    2 | $OriginHost      | dscp1
         RP     | OUT |    3 | $OriginRealm     | ptinovacao.pt
         RP     | OUT |    4 | $CCRType         | 2
         RP     | OUT |    5 | $CCRNumber       | 11
         RP     | OUT |    6 | $ResultCode      | 2001
         RP     | OUT |    7 | $EventTrigger[0] |
         RP     | OUT |    8 | $EventTrigger[1] |
         RP     | OUT |    9 | $ChgRuleInst[0].C| rule2
         RP     | OUT |   10 | $ChgRuleInst[0].C|
         RP     | OUT |   11 | $QoSInfo[0].QoSCl| 1
         RP     | OUT |   12 | $QoSInfo[0].MaxRe| 4096
         RP     | OUT |   13 | $QoSInfo[0].MaxRe| 1024
         RP     | OUT |   14 | $ChgRuleInst[1].C|
         RP     | OUT |   15 | $ChgRuleInst[1].C|
         RP     | OUT |   16 | $ChgRuleRem[0].Ch| rule1
         RP     | OUT |   17 | $ChgRuleRem[0].Ch|
        ------------------------------------------------------------------------------
```

Figura 4.10: Exemplo de um CCA-Modification

Portanto, na figura 4.10 podem-se ver reflectidas as alterações à sessão deste utilizador tal

como determinadas pela política "Profile Definition". No que concerne aos valores de QoS relativamente ao CCA-Initial constata-se que não sofreram alterações e que de momento também não estão a ser instalados "Event-Triggers" específicos. Esta mensagem será posteriormente entregue ao PCEF que se encarregará de aplicar as regras PCC descritas na mesma ao tráfego do utilizador. Em simultâneo, a entidade "TimersControl" será abordada no sentido de actualizar a informação das regras PCC que estão a ser controladas para determinação de eventos temporais.

CCR&CCA-Termination

A recepção de uma mensagem CCR-Termination por parte do PACF assinala o fim de uma sessão. Esta mensagem tem origem no PCEF, depois de este determinar que a sessão IP-CAN deve ser terminada. A mensagem CCR-Termination, tal como é recebida pelo PACF pode ser vista na figura 4.11.

```
2009-10-07 16:51:29.280 librtdap     INF 2 | We have received a message...
2009-10-07 16:51:29.281 librtdap     INF 1 | RECV: ===> RP_OPER[2] CallId[3], OpCode = [13]

        ------------------------------------------------------------------------------
                 OperatioName[GxTerminationRequest]
        ------------------------------------------------------------------------------
         RP    | IN  |   0 | $SessionId       | 16781342
         RP    | IN  |   1 | $AuthAppId       | 16777238
         RP    | IN  |   2 | $OriginHost      | itanium
         RP    | IN  |   3 | $OriginRealm     | pe.teste
         RP    | IN  |   4 | $DestinationRealm| pe.ptin
         RP    | IN  |   5 | $CCRType         | 3
         RP    | IN  |   6 | $CCRNumber       | 12
         RP    | IN  |   7 | $SubsId[0].SubsId| 0
         RP    | IN  |   8 | $SubsId[0].SubsId| 9600101000
         RP    | IN  |   9 | $FrammedIPAddr   | 192.168.1.2
        ------------------------------------------------------------------------------

2009-10-07 16:51:29.282 librtdap     INF 1 | Got a new message from the itQueue
2009-10-07 16:51:29.282 librtdap     INF 1 | TERMINATION_REQUEST
2009-10-07 16:51:29.282 librtdap     INF 1 | InitialRequest
```

Figura 4.11: Exemplo de um CCR-Termination

A mensagem CCR-Termination tem, naturalmente, menos informação do que as mensagens CCR anteriores. Esta mensagem tem como finalidade principal alertar o PACF de que a sessão actual do utilizador está a terminar. Para que se processe o término de uma sessão, mais uma vez as políticas anteriormente referidas serão novamente instanciadas. Pela figura 4.11, podemos constatar que o nome de utilizador cuja sessão está a terminar é o "9600101000". Dependendo

do modo de funcionamento do PACF, o "Register User" irá proceder à remoção da informação deste utilizador ou então irá manter a informação, alterando apenas o valor da conectividade do utilizador para '*offline*. Este comportamento está relacionado com o facto de que o PACF se encontre ou não a funcionar em modo persistente. Passada esta fase, será instanciada a política "Profile Definition" que se encarregará de remover a informação relativa ao estado da sessão do utilizador. Depois de passar por estas duas políticas, a mensagem será novamente entregue ao plugin libRTDAP para que esta a possa entregar ao PCEF. A mensagem CCA-Termination que irá seguir para o PCEF poderá ser vista na figura 4.12

```
2009-10-07 16:51:29.333 librtdap     INF 1 | SEND: <=== RP_OPER[3] CallId[3], OpCode = [14]

-------------------------------------------------------------------------------
         OperatioName[GxTerminationResponse]
-------------------------------------------------------------------------------
 RP    | OUT |   0 | $SessionId       | 16781342
 RP    | OUT |   1 | $AuthAppId       | 16777238
 RP    | OUT |   2 | $OriginHost      | dscp1
 RP    | OUT |   3 | $OriginRealm     | ptinovacao.pt
 RP    | OUT |   4 | $CCRType         | 3
 RP    | OUT |   5 | $CCRNumber       | 12
 RP    | OUT |   6 | $ResultCode      | 2001
-------------------------------------------------------------------------------
```

Figura 4.12: Exemplo de um CCA-Termination

Como se pode verificar pela figura, um dos parâmetros mais importantes será o parâmetro "$ResultCode" que irá indicar ao PCEF que a informação relativa ao utilizador e à sessão foram tratadas com sucesso pelo PCRF. Em simultâneo, a entidade "TimersControl" receberá a informação de quais as regras que estavam afectas à sessão no sentido de apurar se essas regras deverão deixar de sofrer controlo temporal.

4.2.2 Cenário de Testes 2 - RAR & RAA

Neste cenário de testes, o objectivo era testar as capacidades do PACF de enviar mensagens RAR e de receber a respectiva resposta RAA. Neste caso, o PACF, devido a um evento interno, envia uma mensagem RAR para a rede a solicitar a instalação de novas regras PCC e a eventual remoção de outras regras PCC que já estavam instaladas. O resultado desta operação no PACF poderá ser visto na figura 4.13.

```
2009-10-06 14:52:05.158 librtdap     INF 1 | SEND: <=== RP_OPER[4] CallId[1], OpCode = [15]

        ------------------------------------------------------------------------------
                 OperatioName[GxReAuthReq]
        ------------------------------------------------------------------------------
         RP     | OUT |    0 | $SessionId       | 16781342
         RP     | OUT |    1 | $AuthAppId       | 16777238
         RP     | OUT |    2 | $OriginHost      | dscp1
         RP     | OUT |    3 | $OriginRealm     | ptinovacao.pt
         RP     | OUT |    4 | $DestinationRealm| ptinovacao.pt
         RP     | OUT |    5 | $DestinationHost | dsgw
         RP     | OUT |    6 | $ReAuthReqType   | 0
         RP     | OUT |    7 | $ChgRuleInst[0].C| rule1
         RP     | OUT |    8 | $ChgRuleRem[0].Ch| rule2
        ------------------------------------------------------------------------------

2009-10-06 14:52:05.158 librtdap     INF 1 | Timer will be inserted for CallId[1] - [10]secs.
2009-10-06 14:52:05.260 librtdap     INF 2 | We have received a message...
2009-10-06 14:52:05.260 librtdap     INF 1 | RECV: ===> RP_OPER[5] CallId[1], OpCode = [16]

        ------------------------------------------------------------------------------
                 OperatioName[GxReAuthAnswer]
        ------------------------------------------------------------------------------
         RP     | IN  |    0 | $SessionId       | 16781342
         RP     | IN  |    1 | $OriginHost      | dscp1
         RP     | IN  |    2 | $OriginRealm     | ptinovacao.pt
         RP     | IN  |    3 | $ResultCode      | 2001
        ------------------------------------------------------------------------------
```

Figura 4.13: Resultado da execução de um RAR e respectiva resposta RAA

Como se pode verificar, vários parâmetros são enviados na mensagem. Inicialmente vão os parâmetros obrigatórios para uma mensagem do tipo RAR. O último parâmetro obrigatório é identificado na figura pela variável "$ReAuthReqType". De seguida é colocado um Charging-Rule-Install e um Charging-Rule-Remove que significam, respectivamente, a instalação de uma regra (neste caso, a regra "rule1") e a remoção de uma outra regra (neste caso, a regra "rule2"). Este tipo de cenário poder-se-á dever à ocorrência de um evento horário em que haja necessidade de operar alterações nas regras PCC instaladas para uma ou mais sessões. Esta mensagem será enviada para o DSCF, que por sua vez irá enviar a mesma mensagem para o DSGW, que irá promover a tradução da mesma para DIAMETER. Portanto, cada um dos parâmetros enviados na mensagem RTDAP serão traduzidos para um AVP DIAMETER. A forma como se processa esta tradução é através de um ficheiro de configuração do DSGW que determina os mapeamentos entre os parâmetros RTDAP e respectivos AVP's DIAMETER. Desta forma, o RAR será entregue ao equipamento PCEF para aplicação ao tráfego do utilizador. Havendo a boa recepção e aceitação da mensagem RAR por parte do PCEF, este envia uma resposta de volta indicando essa

situação. Como se pode verificar na figura 4.13, o PCEF respondeu com o "Result-Code" 2001, o que indica "DIAMETER_SUCCESS", ou seja, a mensagem foi recebida, interpretada e aplicada com sucesso. "Result-Codes" diferentes de 2001 poderão vir a ter, no futuro, um tratamento específico por parte do PACF.

4.3 Sumário

Neste capítulo, a solução desenvolvida foi efectivamente colocada em teste. Primeiro começou-se por identificar os cenários tipo que terão vantagem em possuir um PCRF para fazer controlo de políticas de QoS. Neste caso, foram identificados quatro cenários tipo. Seguidamente, esses cenários foram colocados à prova através da realização de testes, já com a solução esperada e já envolvendo todos os elementos que darão identidade ao componente PCRF (PACF+DSCF+DSGW). Os testes, embora não tenham utilizado equipamento real devido à indisponibilidade do mesmo, mostram-se muito promissores e demonstram de que forma, um ISP poderá vir a ganhar um maior controlo da sua infra-estrutura de rede através da aplicação de regras PCC.

Capítulo 5

Conclusões

Ao longo de toda esta dissertação, foram apresentadas, discutidas e explicadas as diversas contribuições de forma pormenorizada. Neste capítulo serão discutidas as principais contribuições de cada capítulo para este trabalho de uma forma mais sumária, destacando-se os pontos de maior relevo. Serão igualmente apresentadas perspectivas de trabalho futuro que contribuirão em boa medida como mais valias para a solução agora apresentada.

5.1 Principais contribuições

A grande contribuição desta dissertação resultou da implementação de uma solução de policiamento de tráfego que segue de perto as linhas definidas pelo organismo 3GPP para controlo de tráfego. Uma vez que foram seguidas as linhas orientadoras na norma 3GPP[1] TS 29.212[33] a solução é capaz de usar diversos equipamentos de rede, desde que estes utilizem uma interface Gx. Além disso, esta solução suporta a diferenciação de utilizadores por perfil de utilização, permitindo a implementação por parte dos operadores de telecomunicações de modernos modelos de negócio. De seguida, será feito um resumo das principais contribuições por capítulo para o desenvolvimento desta solução.

[1] 3rd Generation Partnership Project

5.1.1 Estado da arte

No capítulo 2, começou-se por fazer uma abordagem genérica à gestão de redes baseada em políticas, através da filosofia PBNM[2], particularizando-se depois, mais um pouco, para a visão do PBNM segundo alguns organismos de normalização internacionais, nomeadamente o IETF[3], o ETSI-TISPAN[4] e o 3GPP. Na abordagem que foi feita ao PBNM, mostrou-se o porquê de haver a necessidade de implementar a gestão de uma rede assente em políticas. A eficaz gestão de uma rede permitirá, entre outros, que a rede se torne num centro de lucro efectivo para o operador. Questionando abordagens à problemática da QoS[5], nomeadamente colocando em causa a típica solução do sobre-aprovisionamento da rede para resolver o problema da gestão de QoS, o PBNM mostra claramente que, a simples adição de mais largura de banda ao núcleo de uma infra-estrutura de rede não é solução para a gestão de QoS. Ao invés, dotando a rede de capacidades de configuração dinâmicas, capazes de interagir com os diversos equipamentos, recorrendo para isso a modelos de dados que descrevem a interacção com os mesmos, será possível vender serviços cada vez melhores e cada vez mais personalizados a um conjunto de clientes cujas exigências perante os operadores têm vindo cada vez mais a aumentar. Estas capacidades dinâmicas da rede permitirão também que o núcleo da rede se torne num centro de lucro para o operador. A utilização de regras de negócio para motivar a configuração, e a própria organização da rede, permitirão que os serviços diferenciados que o operador apresenta junto dos clientes, possam ser transparentemente transpostos para os equipamentos que irão suportar esses mesmos serviços. Um operador, ao utilizar métodos de gestão de rede ligados à filosofia PBNM, fará igualmente que a sua rede se torne mais segura, capaz e eficaz, diminuindo os tempos de falta de serviço e, consequentemente, proporcionando uma experiência mais satisfatória ao utilizador final.

De seguida, ainda no capítulo 2, foram abordadas as visões de alguns organismos de normalização internacional relativas à implementação da filosofia PBNM nas suas áreas de acção. Foram abordadas as visões do IETF, do ETSI-TISPAN e do 3GPP. Aquilo que se destaca claramente nas três abordagens é que todas utilizam um elemento central para fazer o provisionamento de políticas nos equipamentos de rede. Esta entidade é a responsável por inquirir os elementos necessários para a sua tomada de decisão. Focando, em particular, na abordagem 3GPP, esta entidade é o PCRF[6], que fará a ligação entre os elementos responsáveis pelo fornecimento das políticas para

[2]Policy Based Network Management

[3]Internet Engineering Task Force

[4]European Telecommunications Standards Institute - Telecommunications and Internet Converged Services and Protocols for Advanced Networking

[5]Quality of Service

[6]Policy and Charging Rules Function

os utilizadores e o equipamento responsável pela aplicação dessas mesmas políticas, o PCEF[7]. Estas interacções permitiram que um operador possa oferecer serviços diferenciados aos seus utilizadores, assim como garantir o funcionamento equilibrado da sua estrutura de rede, impedindo eventuais "abusos" por parte de algum utilizador. Permite também assumir um modelo de negócio baseado na qualidade ao invés do típico modelo de negócio baseado na quantidade.

5.1.2 Desenvolvimento da Solução

No capítulo 3, foi descrito todo o caminho para a criação de uma solução que implemente as capacidades de um PCRF, tal como definido na norma 3GPP TS 23.203[28]. Como se pode verificar, a entidade PCRF será constituída por vários componentes que assim implementarão a funcionalidade de um PCRF. De todos os componentes destaca-se o PACF[8], juntamente com o plugin libRTDAP e respectivas políticas de suporte, que implementam a funcionalidade preconizada nesta dissertação. Neste capítulo começou-se por fazer uma apresentação da solução actual da PT Inovação, apresentou-se o componente PACF e o porquê da escolha deste componente para base da solução proposta. Seguidamente foram descritas todas as fases de desenvolvimento do libRTDAP e respectivas políticas de suporte. Esta abordagem, um plugin mais as políticas de suporte, permite que a solução seja extremamente personalizável. Estando o plugin responsável pela recepção e envio das mensagens, assim como a respectiva descodificação e codificação das mesmas, o trabalho de análise do conteúdo das mensagens fica delegado para as políticas de suporte, permitindo assim que, alterações aos modelos de decisão do PACF possam ser feitas através da modificação destas políticas ou adição de novas políticas que implementarão novas funcionalidades e métodos de análise da mensagem e construção da respectiva resposta para entrega à rede. Mostrou-se também que é possível proceder à implementação de um mecanismo de restrições horárias genérico capaz de dar resposta a vários cenários distinto. Por todas estas razões é possível extrapolar que o PACF é um componente bastante personalizável, preparado para ser adaptado às diversas necessidades dos mais vários operadores.

5.1.3 Cenários e Testes

No capítulo 4 foram apresentados os cenários de utilização expectáveis para esta solução, assim como testes concretos que foram realizados à solução. No caso dos cenários, foram identificados

[7]Policy and Charging Enforcement Function
[8]Policy and Admission Control Function

cenários cuja aplicabilidade junto de um operador de telecomunicações seja expectável. Foram identificados quatro cenários tipo que, dependendo do modo de funcionamento do PACF poderão apresentar algumas variações. Desde logo, o PACF em modo standalone dispensará a consulta da lógica do operador para obter um identificador de perfil para aplicar ao utilizador. Os três primeiros cenários dizem respeito ao funcionamento típico de pedido originado na rede e resposta dada pelo PACF a essa mesma pergunta. O último cenário identificado estará, porventura, mais intimamente relacionado com o mecanismo de restrições horárias, uma vez que este cenário prevê a realização de alterações a uma sessão em curso sem que haja um pedido explícito do PCEF nesse sentido.

Naturalmente, os diversos cenários do capítulo 4 deram origem a realização de alguns testes. A viabilidade destes cenários foi testada recorrendo a alguns testes sintéticos. Como se pode constatar pela realização dos diversos testes, qualquer um dos cenários é viável. A funcionalidade e viabilidade dos diversos cenários ficou portanto comprovada, sendo constatado que o PACF tem a capacidade de dar resposta a qualquer um dos quatro cenários agora preconizados. De notar, no entanto, que o PACF, dada a sua estrutura modular e extremamente personalizável, poderá ser rapidamente estendido para suportar novos cenários dentro do âmbito da norma 3GPP TS 23.203[28] ou mesmo cenários mais particulares dentro das necessidades do sistema IP-Raft.

5.2 Trabalho Futuro

Dadas as características do projecto desenvolvido, existem alguns pontos que poderão ser tomados em consideração para trabalho futuro e que trarão, inegavelmente, mais valia à solução para que esta se posicione no mercado como uma solução mais funcional e estável, nomeadamente:

- GRE: O GRE[9] será um componente a desenvolver, cuja principal função será tomar a decisão de qual o perfil a aplicar a um determinado utilizador. Actualmente, o perfil de utilizador é tipicamente devolvida pela lógica do operador. No futuro, com um motor de regras genérico, as decisões serão tomadas de acordo com as condições definidas pelo operador. Desta forma, a informação do utilizador recebida da rede será utilizada para eficazmente derivar o perfil do utilizador, de forma completamente automática, dispensando a utilização da lógica do operador e permitindo que o PACF seja um produto completamente autónomo e capaz de tomar decisões baseadas nos diversos parâmetros que lhe são

[9]Generic Rules Engine

fornecidos.

- Integração com o AF[10]: Num mundo em que os serviços all-IP estão cada vez mais presentes, nomeadamente serviços VoIP[11] ou mesmo VoD[12], dotar o PACF da capacidade de comunicação com a camada de serviço do operador será um objectivo indispensável. Desta forma, necessidades de QoS por parte das aplicações poderão ser comunicadas ao PACF para que este valide a aplicação desses pedidos e que proceda junto do PCEF à reserva de recursos.

- Implementação de um mecanismo de CAC[13]: Para que o PACF possa oferecer QoS a pedido às aplicações que assim o requeiram, terá que haver naturalmente um controlo dos recursos alocados, não só por sessão, mas também será necessário haver um conhecimento global de qual é o estado actual da rede. Torna-se portanto necessário proceder à implementação de um mecanismo de controlo de admissão dos vários pedidos, garantido que o pedido se encontra de acordo com o perfil do utilizador, que o utilizador tem direito a pedir os requisitos de QoS específicos, e que as condições actuais da rede não serão degradadas como consequência desse pedido.

Todas estas ideias contribuirão para a manutenção de uma solução mais equilibrada, capaz de responder às necessidades e desafios actuais e futuros que os operadores de telecomunicações possam vir a enfrentar.

[10] Application Function
[11] Voice over IP
[12] Video-on-Demand
[13] Call Admission Control

Bibliografia

[1] F. Ricciato. Unwanted Traffic in 3G Networks. *ACM SIGCOMM Computer Communication Review*, 36(2), April 2006.

[2] S. Baudet, C. Besset-Bathias, P. Frêne, and N. Giroux. QoS implementation in UMTS networks. *Alcatel Telecommunications Review*, 2001.

[3] Internet Engineering Task Force. URL: www.ietf.org.

[4] European Telecommunications Standards Institute - Telecommunications and Internet converged Services and Protocols for Advanced Networking. URL: http://www.etsi.org/tispan/.

[5] 3rd Generation Partnership Project. URL: www.3gpp.org.

[6] J. Strassner. *Policy-Based Network Management*. Morgan Kaufman Publishers Inc., September 2003.

[7] D. C. Verma. Simplifying Network Administration Using Policy-Based Management. *IEEE Network*, March/April 2003.

[8] V. P. Kumar, T. V. Lakshman, and D. Stiliadis. Beyond best effort: Router architectures for the differentiated services of tomorrow‘s internet. *IEEE Communications Magazine*, May 1998.

[9] D. Soldani. *QoS and QoE Management in UMTS Cellular Systems*. Wiley, August 2006.

[10] M. J. Karam and F. A. Tobagi. On traffic types and service classes in the Internet. *Global Telecommunications Conference, 2000. GLOBECOM '00. IEEE*, 1(548-554), 2000.

[11] A. Gupta, D. O. Stahl, and A. B. Whinston. The Economics of Network Management. *Communications of the ACM*, 42(9):57–63, 1999.

[12] D. C. Verma. *Policy-Based Networking: Architecture and Algorithms*. Sams, 2000.

[13] X. Xiao and Lionel M. Ni. Internet QoS: A Big Picture. *IEEE Network*, 13(2):8–18, 1999.

[14] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. An Architecture for Differentiated Services. RFC 2475 (Informational), December 1998. Updated by RFC 3260.

[15] R. Braden, D. Clark, and S. Shenker. Integrated Services in the Internet Architecture: an Overview. RFC 1633 (Informational), June 1994.

[16] Xiaorong Li, Hoong Maeng Chan, T. Hung, and S. J. Turner. Design of an SLA-Driven QoS Management Platform for Provisioning Multimedia Personalized Services. *Advanced Information Networking and Applications - Workshops, 2008. AINAW 2008. 22nd International Conference*, pages 1405–1409, March 2008.

[17] C. Esteve Rothenberg and A. Roos. A review of policy-based resource and admission control functions in evolving access and next generation networks. *Springer Science+Business Media*, 2008.

[18] R. Yavatkar, D. Pendarakis, and R. Guerin. A Framework for Policy-based Admission Control. RFC 2753 (Informational), January 2000.

[19] D. Durham, J. Boyle, R. Cohen, S. Herzog, R. Rajan, and A. Sastry. The COPS (Common Open Policy Service) Protocol. RFC 2748 (Proposed Standard), January 2000.

[20] K. Chan, J. Seligson, D. Durham, S. Gai, K. McCloghrie, S. Herzog, F. Reichmeyer, R. Yavatkar, and A. Smith. COPS Usage for Policy Provisioning (COPS-PR). RFC 3084 (Proposed Standard), March 2001.

[21] ETSI-TIPSAN: NGN Release 1; Release definition. In: ETSI TR 180 001 V1.1.1, March 2006.

[22] ETSI-TISPAN: Resource and Admission Control Sub-System (RACS): Functional Architecture. In: ETSI ES 282 003 V2.0.0, May 2008.

[23] ETSI-TISPAN: DIAMETER protocol for session based policy set-up information exchange between the Application Function (AF) and the Service Policy Decision Function (SPDF); Protocol specification. In: ETSI TS 183 017 V1.1.1, March 2006.

[24] 3GPP: Policy control over Go interface. In: 3GPP TS 29.207 V6.5.0, September 2005.

[25] 3GPP: Policy and charging control architecture IP Multimedia Subsystem (IMS) - Stage 2. In: 3GPP TS 23.228 V7.14.0, December 2008.

[26] G. Camarillo and Miguel-Angel García-Martín. *The 3G IP multimedia subsystem (IMS); merging the Internet and the cellular worlds*. John Wiley & Sons, 2004.

[27] 3GPP: Policy control over Gq interface. In: 3GPP TS 29.209 V6.7.0, June 2007.

[28] 3GPP: Policy and charging control architecture. In: 3GPP TS 23.203 V7.11.0, June 2009.

[29] V. Y.H. Kueh and M. Wilson. Evolution of policy control and charging (pcc) architecture for 3gpp evolved system architecture. In *Vehicular Technology Conference*. Fujitsu Laboratories of Europe Ltd (FLE),, 2006.

[30] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, and J. Arkko. Diameter Base Protocol. RFC 3588 (Proposed Standard), September 2003.

[31] 3GPP: Vocabulary for 3GPP Specifications. In: 3GPP TS 21.905 V8.8.0, March 2008.

[32] 3GPP: Policy and Charging Control over Rx reference point. In: 3GPP TS 29.214 V7.7.0, March 2009.

[33] 3GPP: Policy and Charging Control over Gx reference point. In: 3GPP TS 29.212 V7.8.0, March 2009.

[34] C. MC Daid. Overview and Comparison of QoS Control in Next Generation Networks. Palowireless.

[35] Cisco Systems. Cisco SCE 2000 Series Service Control Engine. *Data Sheet*, October 2006.

[36] Cisco SCMS SCE Subscriber API Programmer Guide. Release 3.1, May 2007.

[37] Cisco SCMS SM Java API Programmer Guide. Release 3.1, May 2007.

[38] Cisco. Subscriber Control and Billing with the Cisco Content Services Gateway. White Paper, March 2003.

[39] PT Inovação. IP-Raft Policy and Charging Control System. Data sheet.

[40] R. Azevedo, F. Fontes, J. Loura, and A. Oliveira. Desenvolvimento de um Policy Server aplicado a um Cenário de Convergência de Redes. *Saber & Fazer*, (5):73–83, December 2007.