Solução Aberta para uma Rede de Sondas de QoS na RCTS

Pedro Queirós
Centro Algoritmi
Univ. Minho, Portugal
a39845@alunos.uminho.pt

M. João Nicolau
Centro Algoritmi & DSI
Univ. Minho, Portugal
joao@dsi.uminho.pt

Alexandre Santos
Centro Algoritmi & DI
Univ. Minho, Portugal
alex@di.uminho.pt

***Resumo*—A RCTS tem em operação, desde há vários anos, um conjunto de sondas do tipo *appliance* destinadas a aferir de forma contínua a qualidade da rede, tanto em IPv4 como IPv6. Normalmente, pelo facto de serem *appliances*, consequentemente "fechadas", cria-se uma dependêndia de *hardware* e *software*. Apesar das *appliances* possuírem algumas características únicas ao nível de *hardware*, podem desenvolver-se soluções abertas (*open source*) equiparadas. Na realidade, vários sistemas de aferição de desempenho de redes, desenvolvidos em *open source*, tanto na GÉANT (rede académica europeia) como na Internet2 (rede académica americana), atingiram uma notoriedade que permite considerá-los como alternativas credíveis para as atuais sondas, se se provar a sua maturidade e estabilidade. Neste trabalho apresenta-se o desenvolvimento e teste de uma especificação de uma nova sonda para a RCTS, baseada em soluções *open source*.**

***Index Terms*—Sondas de QoS, Sincronização Temporal, One Way Delay, RCTS**

I. Introdução

O aumento da largura de banda possibilitou o surgimento de novos serviços, como o *video-on-demand*, e de uma nova cultura - *always-online* - onde o modelo *best-effort* da Internet já não se adequa. É necessário agora oferecer aos utilizadores uma experiência de utilização melhorada, que permita a utilização destes novos serviços na sua plenitude. Para isso, é também necessário medir constantemente a qualidade de serviço (em inglês: QoS) da rede.

É do interesse de um fornecedor de serviço de rede ter conhecimento dos problemas muito antes dos seus clientes os reportarem, tornando-se vital uma monitorização constante das ligações, de forma a detetar atempadamente avarias na rede, bem como conter os estragos que estas possam causar na mesma. Utilizando técnicas de monitorização do tráfego na rede, é possível estimar o comportamento da rede, saber se existem estrangulamentos na mesma, e mesmo detetar ataques a esta.

II. Estado da arte

Para efetuar esta monitorização, utilizam-se normalmente dois métodos de medição do tráfego: medição ativa e medição passiva.

A. Medição ativa

A medição ativa consiste na introdução de tráfego na rede que permita a medição fim-a-fim de parâmetros como o atraso, a perda de pacotes, a variação do atraso entre pacotes sucessivos de dados (*jitter*), a largura de banda disponível, entre outros. As medições ativas são feitas recorrendo a *software* específico, o qual permite construir pacotes que são transmitidos através da rede e posteriormente permitem a análise e cálculo de alguns dos parâmetros referidos anteriormente.

Uma destas ferramentas, e provavelmente a mais conhecida, é a ferramenta *ping* [1]. Fazendo uso do protocolo ICMP, é enviada uma mensagem *echo request* e o *echo reply* permite calcular o tempo RTT (*Round Trip Time*) e também informação sobre a perda de pacotes. Outra ferramenta que utiliza o protocolo ICMP é o *traceroute* [2]. Esta ferramenta permite determinar o caminho percorrido por uma mensagem na rede e o tempo de ida e volta para cada salto até ao destino. Ainda hoje, mais de duas décadas após o seu aparecimento, estas continuam a ser das ferramentas mais usadas para diagnosticar falhas de rede. Em 1997 foi criado um grupo de trabalho na IETF, denominado por IPPM [3], que se propôs a desenvolver métricas padrão para avaliar a qualidade, desempenho e confiança dos serviços de transporte de dados da Internet, em termos quantitativos. Este grupo de trabalho lançou documentos que especificam métricas e procedimentos para as determinar, como *one-way delay* ou *one-way loss*, obtidas a partir do OWAMP [4]. Estas medições efetuadas só num sentido permitem ao operador perceber em que sentido da ligação é que se encontra um problema, mas necessitam que os terminais de rede se encontrem sincronizados com alta precisão, p.ex. através de fontes externas de relógio como o GPS ou CDMA.

Foram encontradas três ferramentas que implementam o protocolo OWAMP: OWAMP [5] (tem o mesmo nome do protocolo), desenvolvida pela Internet2, QoSMet [6] (é baseada num *draft* do protocolo) e J-OWAMP [7], uma implementação em Java desenvolvida pelo Instituto de Telecomunicações de Aveiro. Em [8], os autores desenvolveram uma solução em *hardware* que gera informação de sincronização de relógio extremamente precisa para os pacotes de teste OWAMP, tanto no emissor como no receptor.

Outras métricas especificadas pelo IPPM dizem respeito à capacidade de uma ligação (largura de banda disponível quando não existe tráfego), largura de banda disponível (quando existe tráfego concorrente) e capacidade de transferência em bloco (*bulk transfer capacity*). Iperf [9], nuttcp

[10] e thrulay [11] são algumas das ferramentas que permitem calcular a largura de banda de uma ligação, usando UDP ou TCP, bem como outras métricas, tal como o atraso e a perda de pacotes.

B. Medição passiva

A medição passiva, por sua vez, não interfere com o tráfego na rede, consistindo apenas na análise do mesmo. O tráfego é capturado numa localização específica da rede, sendo armazenado e posteriormente processado, de forma a elaborar estatísticas sobre o mesmo. Esta análise do tráfego pode ser feita com vários níveis de granularidade, visto que os pacotes podem ser capturados ao nível da ligação (camada 2 da pilha OSI).

A medição passiva pode ainda ser distinguida, conforme se utilizem métodos baseados em *hardware* ou *software*. Os métodos baseados em *software* passam, na sua maioria, pela modificação dos sistemas operativos e *drivers* dos dispositivos de rede, de forma a possibilitar a captura do tráfego.

As ferramentas mais referidas na literatura para efetuar a captura do tráfego e respetiva análise são o *tcpdump/libpcap*, *tcptrace* e *Wireshark*. Em [12], os autores referem algumas limitações na utilização deste tipo de soluções em *hardware* convencional para analisar tráfego em ligações de alto débito (10 Gbps ou mais).

Os métodos baseados em *hardware* são desenhados para possibilitar a réplica dos dados que atravessam o canal de transmissão, de forma a duplicar os mesmos, permitindo assim que o tráfego seja dividido e processado de igual forma pela interface de rede e pelo *hardware* específico de monitorização. As diferenças entre os métodos baseados em *software* e *hardware* refletem-se essencialmente em custo e precisão dos resultados. Os dispositivos de rede comuns não são desenhados com a monitorização dos pacotes em mente, pelo que não manifestam bom desempenho quando é necessário efetuar a captura dos pacotes.

Foi com esta limitação em mente que os autores em [13] desenvolveram *hardware* específico para a captura de tráfego em redes de alta velocidade. Desde então, as placas de captura e processamento de tráfego Endace DAG [14] são hoje usadas em muitos projetos que requerem alta fiabilidade e sincronismo dos dados, garantindo 100% de captura de tráfego, mesmo em redes de alto débito (10 Gbps).

No entanto, mesmo utilizando métodos baseados em *hardware* específico para a captura do tráfego, é necessário garantir o armazenamento e processamento do tráfego capturado. Se esta captura for feita em troços de alto débito - 10 Gbps, p.ex. - a uma velocidade constante, podem obter-se 4.5 Terabytes de dados em apenas uma hora de captura (10 Gbps $\times$ 3600 segundos $\simeq$ 4.5 Terabytes). Além do armazenamento, o processamento desta quantidade de dados também não é trivial. O processamento pode ser feito em tempo real, se for crítico, ou mais tarde, permitindo a correlação do tráfego capturado dentro de uma determinada janela de tempo. Os autores em [15] sugerem um sistema com uma arquitetura que distribui as várias fases dos processos de captura e análise do tráfego, executando-as paralelamente em *hardware* convencional, tornando-se assim escalável e possibilitando a captura de tráfego em interfaces de alto débito.

A filtragem de pacotes permite observar apenas uma parte do tráfego, recorrendo a regras que permitem filtrar o tipo de tráfego a recolher (p.ex., todos os pacotes TCP), limitando o tamanho da informação a analisar ao considerar apenas os primeiros `N` bytes de dados do pacote, ignorando os restantes.

Um fluxo é uma sequência de pacotes que são trocados entre duas entidades, identificado recorrendo a uma chave, formada por alguns campos dos pacotes, como por exemplo o par `<IP orig., IP dest.>` ou `<porta#orig., porta#dest.>`, ou o `<protocolo#>` de transporte. Assim, todos os pacotes com uma chave idêntica são considerados como pertencentes ao mesmo fluxo, simplificando a análise. Neste contexto, a Cisco desenvolvou o NetFlow [16], que foi a norma *de-facto* durante muitos anos, antes do IETF organizar um grupo de trabalho denominado de IPFIX, que lançou uma norma de igual nomenclatura (IPFIX) [17], definindo como a informação de um fluxo IP deve ser formatada e transferida. A amostragem dos pacotes passa por recolher apenas alguns dos pacotes que chegam à interface de captura - por exemplo, recolher 5 em cada 100 pacotes. Existem vários métodos de amostragem, tais como aleatório simples, sistemático e estratificado, que são descritos e comparados em detalhe com novos métodos de amostragem em [18].

C. Projetos internacionais

Enquanto algumas das ferramentas descritas anteriormente foram derivadas de projetos de métricas de QoS, outras foram diretamente utilizadas em alguns projetos internacionais, que são brevemente referidos de seguida.

Surveyor [19] Este projeto visava a medição do desempenho das ligações de uma WAN, utilizando métricas bem definidas. Neste projeto foram usadas as métricas de *one-way delay* e perda de pacotes para caracterizar as ligações entre as organizações participantes. Em 1999 este projeto contava com medições em 41 locais, maioritariamente Universidades e Centros I&D.

RIPE Network Coordination Centre Test Traffic Measurement [20] Um projeto semelhante ao Surveyor, destinado a todos aqueles que desejam testar as suas ligações com outros clientes deste serviço. Disponibiliza um serviço comercial que é gerido pelo RIPE NCC.

RIPE Network Coordination Centre Atlas [21] Este trata-se de outro projeto do RIPE NCC, que utiliza sondas próprias (construídas pelo RIPE NCC) para efetuar medições. Estas sondas podem ser compradas e instaladas por qualquer operador, funcionando como um dispositivo *plug-and-play*. O objetivo desta rede de sondas é elaborar vários mapas a nível mundial (latência, conetividade, etc.) para avaliar o estado da rede que liga as sondas. Depois de devidamente instaladas, estas sondas fazem vários testes com a rede de sondas ativas no projeto Atlas (em Agosto de 2012 existiam já 1750 sondas espalhadas por todo o Mundo): testes de *traceroute*, DNS, *round trip time*, entre outros.

perfSonar [22] perfSonar é uma arquitetura orientada aos serviços que permite a monitorização do desempenho das redes, facilitando a resolução de problemas em ligações fim a fim que atravessam várias redes. Especifica um conjunto de serviços e define um protocolo de comunicação entre os vários serviços, permitindo a qualquer pessoa interessada implementar os mesmos. Esta ferramenta pode ser vista como *middleware*, permitindo que várias implementações de serviços diferentes comuniquem entre si, alargando assim o espectro de medições que podem ser efetuadas entre os utilizadores desta arquitetura. O perfSonar foi desenvolvido através de uma colaboração internacional entre a Internet2, ESnet, RNP e a GÉANT.

PingER [23] Utilizando a ferramenta *ping*, este projeto da IEPM visa testar a conetividade das ligações da Internet. Começando por se focar na comunidade HEP, atualmente conta com medições para mais de 700 locais em mais de 160 países e disponibiliza os dados das suas medições online.

Etomic [24] Este projeto pretende, à semelhança de outros, criar uma infraestrutura de medições ativas capaz de operar com uma resolução temporal elevada (aproximadamente 10 ns) e globalmente sincronizada. O objetivo destas medições é fornecer uma visão geral de alta resolução das rápidas mudanças do tráfego na rede. Entre os participantes deste projeto encontram-se várias universidades e institutos de investigação europeus.

Archipelago [25] A CAIDA [26] trata-se de uma colaboração entre organizações comerciais, governamentais e de investigação, com o objetivo de promover uma maior cooperação no desenvolvimento e manutenção de uma infraestrutura escalável e robusta para a Internet. Muitas das ferramentas desenvolvidas em projetos de investigação são depois usadas em projetos da CAIDA em larga escala. Um desses projetos é o Archipelago (ou Ark), uma infraestrutura de medições ativas que pretende reduzir o esforço necessário ao desenvolvimento e implementação de medições em larga escala.

D. Sumário

Apesar de se referirem algumas das ferramentas mais usadas, seria impossível analisar, ou sequer descrever, todas as ferramentas desenvolvidas para a medição, ativa ou passiva, de parâmetros de QoS. A CAIDA fornece uma lista extensa, agrupando as ferramentas conforme a sua função na medição dos parâmetros. Existe atualmente um vasto leque de opções para efetuar medições passivas e ativas. A investigação nesta área é muito rica e produz com frequência novos dados e direções que permitem a evolução das métricas e das ferramentas utilizadas para as obter.

III. Rede de Sondas: Soluções Open Source

Nesta secção apresenta-se a metodologia e resultado dos testes laboratoriais efetuados para determinar a viabilidade da utilização de soluções abertas (Open Source) para uma rede de sondas de QoS na RCTS. Inicialmente, pretendeu-se aferir a viabilidade operacional de reutilização do *hardware* das *appliance* QoSMetrics (pré-existentes) usando uma instalação personalizada, com sistema operativo CentOS, de baixa complexidade e pouco exigente em termos de recursos. Define-se depois uma arquitetura genérica, que pode utilizar tanto a *appliance* QoSMetrics como *hardware* genérico, integrando outras ferramentas Open Source, nomeadamente o perfSonar, para estabelecer uma rede de sondas e o respetivo sistema de monitorização de QoS.

A. Enquadramento e Análise de Cenários de Reutilização

Desde 2006 que a FCCN possui um parque de sondas na RCTS, o qual permite o controlo e medição da qualidade de serviço que a FCCN oferece aos utilizadores desta rede - nomeadamente Universidades, Laboratórios de Investigação e Institutos Politécnicos. O parque de sondas é constituído por duas sondas principais, localizadas nos nós centrais da rede (Lisboa e Porto), de maior capacidade, e por várias sondas de menor capacidade localizadas nos pontos de interligação da RCTS com as instituições académicas. No total, 26 sondas espalhadas pelos maiores pólos de Ensino Superior em Portugal - apresentadas na figura 1 - efetuam medições que permitem avaliar o estado das ligações da rede [27].

Figura 1: Mapa da distribuição das sondas

O *hardware* da maioria das sondas é composto por chassi de *rack*, tamanho 1U, com processador a 1.6 GHz, 512 MB RAM. Possui ainda um adaptador de cartões CF para IDE, produto da QoSMetrics, que utiliza um cartão Compact Flash de 2 GB como dispositivo de armazenamento primário.

Durante este estudo, foram realizadas com sucesso novas instalações de *software* aberto, quer ao nível de SO, quer a nível aplicacional, nas sondas / *appliance* QoSMetrics pré-existentes. Foi demonstrado que o *hardware* atualmente à disposição da FCCN poderia ser reutilizado, por forma a instalar uma nova solução de medição de parâmetros de qualidade de serviço. Foram no entanto identificadas e documentadas limitações relacionadas com a utilização de cartões CF (sem

suporte para DMA) como dispositivo de armazenamento primário; recomendar-se-ia a inclusão de um disco SSD e ainda a expansão da memória RAM para o máximo suportado pela placa mãe, 2 GB.

Como boa caraterística do *hardware* da *appliance* convirá referir a existência de uma boa solução proprietária para a sincronização temporal. Quando lidamos com medições que exigem uma precisão ao nível do nanosegundo ou do microsegundo, não podemos ignorar os fatores de atraso presentes nos métodos de sincronização temporal através de uma rede de pacotes. É um facto que a sincronização temporal absoluta sobre uma rede de pacotes é impossível: existem atrasos não determinísticos que irão sempre influenciar a sincronização. O importante é estar consciente destas limitações e utilizar os métodos que melhor se adequam ao nível de exatidão que pretendemos obter com a sincronização e que são, ao mesmo tempo, economicamente eficientes.

Uma exatidão fiável na ordem dos microsegundos requer *hardware* com um propósito específico, que geralmente inclui uma relação explícita entre o *hardware* que produz um evento e o relógio que gera o carimbo temporal associado a esse evento.

No levantamento de plataformas de metrologia em rede feito em [28], os autores analisam os métodos aqui apresentados e constatam que a maioria das soluções existentes, ou são precisas e eficientes, mas caras e não escaláveis, ou são imprecisas e com fraco desempenho, mas de fácil implementação. Assim, concluem, projetar uma infraestrutura para permitir a medição da qualidade de serviço de uma rede, que seja precisa, pouco dispendiosa e fácil de implementar, utilizando as soluções atualmente existentes, permanece um desafio.

Neste contexto, este trabalho pretende explorar uma solução de arquitetura aberta, tanto a nível de *hardware* como de *software*, capaz de dar resposta satisfatória à rede de sondas para o sistema de monitorização de QoS da rede académica.

B. Arquitetura para Solução Aberta

Para a definição de uma arquitetura de medição de parâmetros de QoS, consideraram-se duas premisas: a topologia da rede e a solução atualmente em utilização, ambas referidas em III-A. As várias instituições servidas pela RCTS encontram-se, na sua maioria, ligadas diretamente a dois nós principais: Lisboa e Porto. Assim sendo, justifica-se a realização de medições entre os nós principais e a periferia da rede para avaliar o estado das ligações. Esta é também a filosofia empregue pela solução atual, onde as sondas periféricas fazem medições com as sondas de Lisboa e Porto para aferir os parâmetros de QoS das ligações.

perfSONAR: Monitorização e software *aberto*

O perfSONAR é uma *framework* de monitorização desenhada para ambientes de monitorização multi-domínio, que permite a descoberta, cálculo, armazenamento e distribuição de métricas de medição de QoS. O seu desenvolvimento surge de um esforço internacional de cooperação entre várias entidades, nomeadamente a GÉANT [29], ESnet [30], Internet2 [31] e RNP [32]. O objetivo do perfSONAR é diminuir as restrições administrativas existentes no acesso aos dados das medições entre vários domínios, facilitar a resolução de problemas de desempenho fim-a-fim em caminhos de rede que atravessam vários domínios, facilitar a gestão de redes e permitir às aplicações adaptar o seu comportamento ao estado da rede. O desenvolvimento do perfSONAR é da responsabilidade de um consórcio de organizações que procura desenvolver um *middleware* que seja interoperável entre várias redes e permita uma análise do desempenho intra e inter redes, e é constituído por:

- Um protocolo, que assume diferentes tipos de serviços e define uma sintaxe e semântica padrão através dos quais eles comunicam, e permite diferentes implementações de um serviço. Este protocolo é baseado em mensagens SOAP XML e foi desenvolvido pelo OGF NM-WG [33].
- Várias ferramentas (implementações em *software* dos vários serviços), desenvolvidas por diversos parceiros, que tentam implementar uma *framework* interoperável de monitorização de desempenho.

A arquitetura do perfSONAR é apresentada na Figura 2.

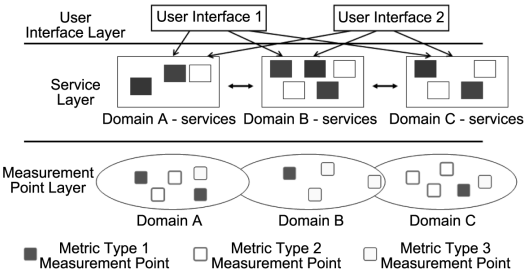

Figura 2: Arquitetura de 3 camadas do perfSONAR [22]

Os serviços do perfSONAR podem correr em múltiplos domínios, usando mensagens SOAP (transportadas em HTTP), tanto para descrever os dados das medições, como para troca de informação entre serviços.

1) Componentes: O perfSONAR, de natureza *open-source* e com uma arquitetura modular, permite aos administradores da rede implementar, combinar e personalizar várias ferramentas de acordo com as suas necessidades individuais. Para facilitar a integração das ferramentas, o perfSONAR oferece serviços individuais [34], responsáveis por funcionalidades específicas, como a recolha de dados e a visualização dos mesmos.

2) Pontos de medição: Os pontos de medição (ou MP) recolhem e publicam a informação obtida através das ferramentas de medição. As medições são geralmente efetuadas localmente, a pedido de um cliente, ou automaticamente, com intervalos programados. É também no ponto de medição que é feita a conversão entre o formato de dados que a ferramenta de medição disponibiliza e o formato de dados do perfSONAR.

3) Arquivos de medição: Os arquivos de medição (ou MA) armazenam o resultado das medições numa base de dados (SQL, por exemplo) ou num sistema de ficheiros. Estes dados

podem ser lidos pelos clientes, processados ou visualizados. Os arquivos de medição também suportam a agregação de dados.

4) Serviços de pesquisa: Cada domínio que implementa o perfSONAR deve possuir um serviço de pesquisa (ou LS). Todos os serviços disponibilizados dentro de um domínio devem ser registados no serviço de pesquisa, permitindo que os serviços de pesquisa de diferentes domínios possam comunicar entre si e partilhar informação. Assim, um utilizador apenas precisa de saber o URL do serviço de pesquisa para saber que tipo de serviços é disponibilizado por um dado domínio.

5) Ferramentas de visualização: As ferramentas de visualização permitem recolher os dados armazenados nos arquivos de medição e tratar a informação recolhida para a apresentar ao utilizador. O carácter *open-source* destas ferramentas permite adaptar a apresentação dos dados à necessidade de grupos específicos de utilizadores.

6) Implementações: Existem atualmente duas implementações principais que se comprometeram a interoperar: **perfSONAR MDM** [35], desenvolvida pelo GÉANT, e **perfSONAR-PS** [36], desenvolvida pela Internet2 e ESnet. Ambas utilizam o protocolo aberto desenvolvido pelo OGF para trocar dados, são baseadas em serviços *web* e partilham os mesmos propósitos: flexibilidade, escalabilidade, abertura e descentralização. Diferem no processo de desenvolvimento, no ciclo de vida dos produtos, na interação com os seus utilizadores e no modelo de implementação e distribuição.

O perfSONAR MDM foi desenvolvido como um sistema de monitorização multi-domínio, pensado para servir os parceiros do GÉANT e destinado a fornecer um serviço federado, centralmente monitorizado e coordenado, com suporte total do GÉANT. O perfSONAR-PS tem um modelo de suporte distribuído, com o objetivo de proliferar o número de nós disponíveis na comunidade.

Ambas as implementações disponibilizam um GUI de configuração e visualização dos resultados das medições baseado numa interface *web*. O perfSONAR MDM disponibiliza ainda um cliente baseado em Java para aceder aos dados das medições, designado por perfSONAR-UI [37].

De entre os utilizadores conhecidos destas duas implementações[1], destacam-se os membros das redes LHC OPN [38] e LHC ONE [39].

7) Ferramentas de medição: As implementações do perfSONAR incluem uma gama de ferramentas que permitem fazer medições e monitorização ao nível da camada IP e acima, desde ferramentas de medição do atraso e da largura de banda disponível, até ferramentas que permitem fazer medições passivas.

8) Medição da taxa de transferência: A medição da taxa de transferência é feita recorrendo ao BWCTL [40] no caso do perfSONAR-PS e ao BWCTL MP, um ponto de medição baseado no BWCTL, no caso do perfSONAR MDM. O BWCTL é de um encapsulador de ferramentas já conhecidas, como o *iperf*, *thrulay* ou o *nuttcp*.

[1]Existe ainda uma terceira implementação, denominada de perfSONAR-NC, desenvolvida pela UNINETT,

9) Medição do atraso: As medições do atraso baseiam-se no HADES [34] para o perfSONAR MDM e no OWAMP para o perfSONAR-PS. O HADES permite obter informações sobre o *one-way delay* [41], *delay jitter* [42], *packet loss rate* [43] e rotas alternativas. O OWAMP é uma ferramenta desenvolvida pela Internet2, que implementa o protocolo do mesmo nome. Esta ferramenta é usada para determinar o *one-way delay* entre dois pontos da rede.

10) Medições passivas: Ferramentas do perfSONAR baseadas na monitorização passiva da rede foram recentemente desenvolvidas (PacketLoss) ou encontram-se em desenvolvimento (ABW2) [34]. Alguns parâmetros e características da rede apenas podem ser obtidos recorrendo a medições passivas, que não influenciam o tráfego da rede.

IV. Solução Aberta: Testes e Resultados

Atendendo ao estado de desenvolvimento da implementação, à alargada comunidade de utilizadores, ao *feedback* existente por parte da FCCN e à estrutura de suporte, optou-se pelo perfSONAR-PS, versão 3.2.2, correndo sobre CentOS 5.8. Esta é portanto uma solução aberta a nível de *software* (aplicacional e SO) e foi implementada e testada com sucesso para a implementação de sondas de QoS em *hardware* da *appliance* QoSMetrics, em *hardware* genérico e também máquinas virtuais.

A. Métricas de qualidade de serviço

Para avaliar a qualidade de serviço das ligações da RCTS, a FCCN tem utilizado a solução proprietária da QoSMetrics para calcular algumas métricas, incluídas também num relatório mensal disponibilizado às instituições ligadas à RCTS. As métricas coletadas e que se pretendem continuar a obter são as seguintes:

- *one-way delay* mínimo, médio e máximo
- *jitter* mínimo, médio e máximo
- pacotes enviados, recebidos, perdidos e duplicados

Estas métricas são sempre calculadas de forma independente para os dois sentidos da ligação. Com a adoção desta nova solução baseada no perfSONAR, pretendeu determinar-se estas métricas (tanto em IPv4, como em IPv6) e manter a compatibilidade com o sistema automático de produção de relatórios que a FCCN tem em operação [27], baseado em Cacti [44] e HP Openview Service Desk (entretanto descontinuado pela HP).

B. Estabilidade da referência temporal

Para determinar as métricas referidas em IV-A, o OWAMP, usado pelo perfSONAR-PS, requer que os pontos entre os quais são feitas as medições estejam corretamente sincronizados com o NTP. Isto é necessário para assegurar que os relógios das máquinas envolvidas nas medições estejam sincronizados, de forma a reduzir o erro dos carimbos temporais atribuídos aos pacotes de medição. Para configurar o NTP, os autores do OWAMP sugerem que sejam configurados como referências temporais pelo menos quatro servidores. Contrariamente, em [45], os autores sugerem que, para manter

a estabilidade do NTP, se deve configurar apenas um servidor (próximo e de baixo *stratum*) como referência temporal. Quando se utiliza apenas um servidor como referência, a variação destes parâmetros torna-se menor, mas em caso de falha desta única referência, o comportamento do NTP irá variar drasticamente [45] [46], produzindo carimbos temporais com flutuações acentuadas que irão interferir nas medições do *one-way delay*.

C. Testes e Resultados

A FCCN possui quatro servidores NTP *stratum 1*, tendo sido utilizados os do Porto e Lisboa, devido à distância dos mesmos às sondas de teste, localizadas na UMinho e em Lisboa. Inicialmente, para facilitar a operacionalização, a sonda da UMinho foi colocada na rede interna da Universidade.

As medições iniciais, configuradas com os valores por omissão (enviar 10 pacotes por segundo, de 20 bytes cada), mostraram valores mínimos próximos daqueles obtidos com a solução proprietária da QoSMetrics, mas valores máximos com picos de dezenas e centenas de milisegundos. Mesmo os valores mínimos continham algumas flutuações, e ocasionalmente alguns picos. Estes resultados eram semelhantes nas duas direções, e estão demonstrados nas Figuras 3 e 4.

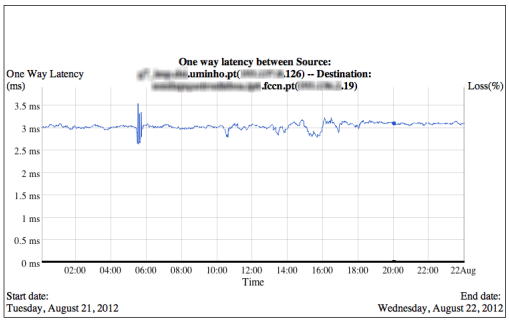

Figura 3: Sonda QoS; valores obtidos para OWD (Braga-Lisboa) mínimo

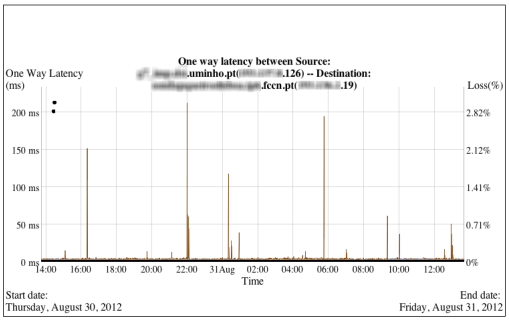

Figura 4: Sonda QoS; valores obtidos para OWD (Braga-Lisboa) máximo

Após análise cuidada e comparação direta com os resultados das sondas em produção, foi possível detetar algumas das flutuações nos resultados das medições diretamente correlacionadas com alterações da temperatura ambiente da sonda. As alterações da temperatura ambiente influenciam as propriedades do cristal do relógio presente na máquina, criando variações na frequência de oscilação, com influência direta na sincronização NTP e, consequentemente, nas medições do OWAMP, tal como assinalado na Figura 5.

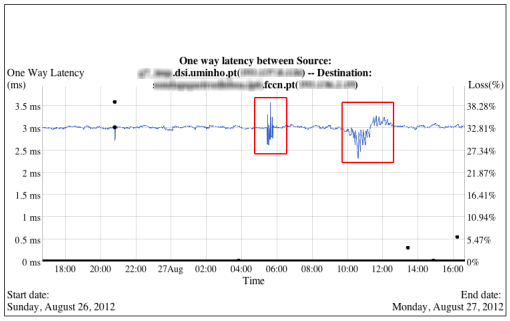

Figura 5: Influência da temperatura na determinação de OWD mínimo entre a sonda da UMinho e Lisboa

Alterações no ambiente de testes

No sentido de melhorar a precisão dos resultados, tentando minimizar a influência de fatores externos, realizaram-se as seguintes alterações:

- sonda na UMinho: instalou-se um disco SSD e mais 512 Mbytes de RAM;
- a sonda foi realocada no *datacenter* da UMinho e ligada diretamente à rede do *backbone* RCTS.

Com este ambiente de teste, foi possível efetuar testes na rede RCTS entre Braga e Lisboa, mimetizando, em parte, a configuração da solução proprietária atualmente existente e em utilização pela FCCN para medição dos parâmetros de QoS na rede da RCTS. Os testes de *one-way delay* foram agora configurados de forma a que fosse enviado um pacote de 1500 bytes a cada segundo.

Com esta alteração no ambiente de teste, as medições passaram a apresentar valores mínimos de *one-way delay* estáveis, como se pode ver na Figura 6. No entanto, os valores máximos continuaram a apresentar picos de várias dezenas de milisegundos, representados na Figura 7. Foi analisada a carga nas sondas, na rede, e a configuração do NTP, mas não foi encontrada nenhuma justificação para os picos apresentados.

Após discussão desta ocorrência com o suporte do perfSONAR-PS, a explicação mais provável para os picos apresentados nas medições é de que estes são decorrentes dos atrasos presentes na pilha do *software*. Em alguns casos, as variações dos atrasos no processo de receber um pacote, gerar uma interrupção, e colocar um carimbo temporal, poderão introduzir falsos atrasos nos pacotes de medição. A carga do CPU das sondas ronda os 10% (estando *idle* 90% do tempo). Foi dada prioridade aos processos do `ntpd` e do `owampd`, através do comando `nice`, mas tal não produziu diferenças notáveis nos valores obtidos com as medições, como seria de esperar dada a carga das máquinas. O tráfego total gerado pelas sondas ronda os 100 kbps, permitindo concluir que o

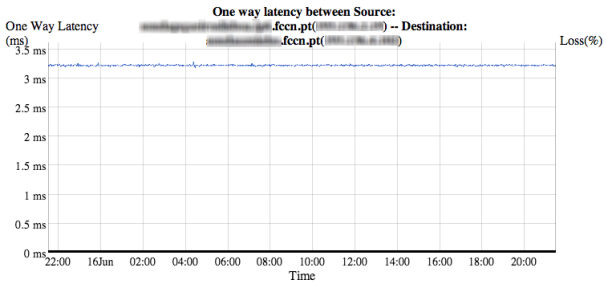

Figura 6: Amostra do *one-way delay* mínimo entre a sonda de Lisboa e Braga

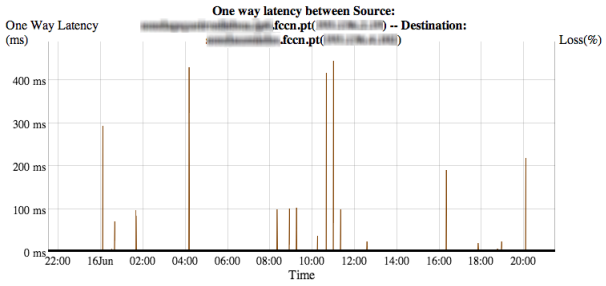

Figura 7: Amostra do *one-way delay* máximo entre a sonda de Lisboa e Braga

processo de monitorização não tem impacto significativo na operação normal da rede.

O NTP das sondas reporta erros de relógio estimados num intervalo de 2 a 35 microsegundos, aproximadamente. No entanto, o próprio NTP estima o *one-way delay* entre o cliente e o servidor NTP como sendo metade do RTT. Comparando com a solução da QoSMetrics, as medições do perfSONAR-PS apresentam variações entre os 0.2 e os 0.4 milisegundos.

Não se considerou a reconfiguração de *hardware* com placas adicionais, nem a inclusão de outros dispositivos de sincronização externa (como p.ex. GPS) e testaram-se apenas soluções e protocolos abertos (não se considerou nunca a hipótese de usar protocolos de sincronização proprietários, como p.ex. o QTP). Existe portanto espaço para aprofundar a investigação ao nível dos mecanismos de sincronização de relógio que utilizem *hardware* genérico.

Importa referir que os picos apresentados são resultado de apenas um ou dois pacotes com um atraso visivelmente superior à média, num total de 60 pacotes em cada sessão de um minuto. Usando métodos estatísticos, estes *outliers* podem ser facilmente descartados.

V. Conclusão

Apresentou-se uma arquitetura para a implementação de uma rede de sondas de QoS baseada apenas em soluções *Open Source*, utilizando unicamente *hardware* genérico, sem dispositivos (internos ou externos) adicionais. Esta arquitetura foi implementada e testada sobre a infraestrutura de *backbone* da RCTS e os resultados obtidos foram muito satisfatórios.

O perfSONAR mostrou ser uma *framework* consolidada, com duas grandes implementações distintas, mas que, essencialmente, têm o mesmo objetivo. No caso específico do perfSONAR-PS, o suporte oferecido pela comunidade (incluindo programadores e utilizadores) é rápido e eficiente, sendo possível obter respostas a dúvidas e sugestões de instalação para casos específicos.

Embora não apresentado neste trabalho, verificou-se ainda que a possibilidade de aceder à base de dados onde são guardados os dados de medição do OWAMP permite uma fácil integração com outras ferramentas (especificamente, foi realizada a integração com a ferramenta de criação de relatórios RCTS).

A dependência do OWAMP em relação ao NTP também influencia os resultados das medições, sendo desejável que no futuro seja possível suportar outros métodos de sincronização temporal. No caso da utilização do NTP, é necessário proceder à sua correta configuração (não utilizar a que vem instalada por definição). Idealmente, o NTP deve ser sempre utilizado em conjunto com uma referência externa de relógio precisa e fiável, como um recetor GPS. O preço destes recetores não é, hoje em dia, muito alto, mas, ainda assim, as questões logísticas implícitas na instalação destes continuam a ser um grande entrave à sua utilização.

Seria interessante poder realizar testes com outras configurações de *hardware* e *software*, para poder aferir com exatidão a causa dos picos nos valores máximos do atraso, e comparar os valores obtidos com diferentes configurações. Infelizmente, por restrições logísticas e temporais, tal não foi possível.

VI. Agradecimentos

Este trabalho foi parcialmente financiado por Fundos FEDER, através do Programa Operacional Fatores de Competitividade - COMPETE - e por Fundos Nacionais através da FCT - Fundação para a Ciência e Tecnologia, através do Projecto: FCOMP-01-0124-FEDER-022674.

Agradece-se ainda a cooperação da FCCN / FCT na facilitação do acesso à RCTS e aos meios computacionais necessários para a prossecução deste trabalho. Um agradecimento muito especial a João Nuno Ferreira, Carlos Friaças, Emanuel Massano e Ana Pinto, todos da FCCN, por toda a ótima colaboração prestada.

Referências

[1] The Story of the PING program. [Online]. Available: http://ftp.arl.mil/~mike/ping.html (accessed July 2012)

[2] Traceroute for Linux. [Online]. Available: http://traceroute.sourceforge.net (accessed July 2012)

[3] IP Performance Metrics. [Online]. Available: http://datatracker.ietf.org/wg/ippm/charter/ (accessed July 2012)

[4] A One-way Active Measurement Protocol (OWAMP). [Online]. Available: http://tools.ietf.org/html/rfc4656 (accessed July 2012)

[5] An implementation of the One-Way Active Measurement Protocol. [Online]. Available: http://www.internet2.edu/performance/owamp/index.html (accessed July 2012)

[6] QoSMet - Quality of Service Metrology. [Online]. Available: http://fabien.michaut.free.fr/qosmet/ (accessed July 2012)

[7] H. Veiga, T. Pinho, J. Oliveira, R. Valadas, P. Salvador, A. Nogueira, *"Active traffic monitoring for heterogeneous environments"*, 4th International Conference on Networking, ICN'05, April 17-21, 2005 - Reunion Island.

[8] Zhang Shu and Katsushi Kobayashi, *"HOTS: An OWAMP-Compliant Hardware Packet Timestamper"*, 2005

[9] iperf. [Online]. Available: http://sourceforge.net/projects/iperf/ (accessed July 2012)

[10] nuttcp. [Online]. Available: http://www.lcp.nrl.navy.mil/nuttcp/stable/nuttcp.html (accessed July 2012)

[11] thrulay. [Online]. Available: http://sourceforge.net/projects/thrulay/ (accessed July 2012)

[12] S. Ubik, P. Zejdl, *"Passive monitoring of 10Gb/s lines with pc hardware"*, 2008

[13] John Cleary, Stephen Donnelly, Ian Graham, Anthony McGregor, Murray Pearson, *"Design Principles for Accurate Passive Measurement"*, 2000

[14] ENDACE DAG CARDS. [Online]. Available: http://www.endace.com/endace-dag-high-speed-packet-capture-cards.html (accessed July 2012)

[15] Se-Hee Han, Myung-Sup Kim, Hong-Taek Ju, James Won-Ki Hong, *"The Architecture of NG-MON: a Passive Network Monitoring System for High-Speed IP Networks"*, 2002

[16] Cisco IOS Flexible NetFlow Technology White Paper. [Online]. Available: http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6555/ps6601/ps6965/prod_white_paper0900aecd804be1cc.html (accessed July 2012)

[17] Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information. [Online]. Available: http://tools.ietf.org/html/rfc5101 (accessed July 2012)

[18] Nick Duffield, *"Sampling for Passive Internet Measurement: A Review"*, 2004

[19] Surveyor. [Online]. Available: http://www.isoc.org/inet99/proceedings/4h/4h_2.htm#surveyor (accessed July 2012)

[20] RIPE Test Traffic Measurement Service. [Online]. Available: http://www.ripe.net/data-tools/stats/ttm/test-traffic-measurement-service (accessed July 2012)

[21] RIPE Atlas [Online]. Available: https://atlas.ripe.net (accessed August 2012)

[22] perfSonar. [Online]. Available: http://www.perfsonar.net (accessed July 2012)

[23] PingER. [Online]. Available: http://www-iepm.slac.stanford.edu/pinger/ (accessed July 2012)

[24] Etomic. [Online]. Available: http://www.etomic.org/ (accessed July 2012)

[25] Archipelago Measurement Infrastructure. [Online]. Available: http://www.caida.org/projects/ark/ (accessed July 2012)

[26] CAIDA. [Online]. Available: http://www.caida.org/ (accessed July 2012)

[27] Emanuel Massano, *"SONAR - Supervisão da RCTS"*, 2008

[28] Anne-Cécile Orgerie, P. Gonçalves, M. Imbert, J. Ridoux, D. Veitch, *"Survey of network metrology platforms"*, 2012

[29] GÉANT. [Online]. Available: http://www.geant.net/ (accessed June 2013)

[30] ESnet. [Online]. Available: http://www.es.net/ (accessed June 2013)

[31] Internet2. [Online]. Available: http://www.internet2.edu/ (accessed June 2013)

[32] RNP. [Online]. Available: http://www.rnp.br/ (accessed June 2013)

[33] Open Grid Forum's Network Measurement Working Group. [Online]. Available: http://nmwg.internet2.edu (accessed June 2013)

[34] D. Vicinanza, *"Intercontinental Multi-Domain Monitoring for LHC with perfSONAR"*, 2012

[35] perfSONAR Multi Domain Monitoring. [Online]. Available: https://forge.geant.net/forge/display/perfsonar/Home (accessed June 2013)

[36] perfSONAR Perl Services. [Online]. Available: http://psps.perfsonar.net/ (accessed June 2013)

[37] perfSONAR User Interface. [Online]. Available: http://www.perfsonar.net/perfsonarUI.html (accessed June 2013)

[38] LHC Optical Private Network. [Online]. Available: http://lhcopn.web.cern.ch/lhcopn/ (accessed June 2013)

[39] LHC Open Network Environment. [Online]. Available: http://lhcone.net/ (accessed June 2013)

[40] BWCTL. [Online]. Available: http://www.internet2.edu/performance/bwctl/ (accessed June 2013)

[41] One-way Delay. [Online]. Available: http://tools.ietf.org/html/rfc2679 (accessed June 2013)

[42] Jitter Delay. [Online]. Available: http://tools.ietf.org/html/rfc3393 (accessed June 2013)

[43] Packet loss. [Online]. Available: http://tools.ietf.org/html/rfc2680 (accessed June 2013)

[44] Cacti. [Online]. Available: http://www.cacti.net/ (accessed June 2013)

[45] Wolfgang John, Sven Tafvelin, Tomas Olovsson, *"Passive internet measurement: Overview and guidelines based on experiences"*, 2010

[46] OWAMP Details. [Online]. Available: http://www.internet2.edu/performance/owamp/details.html#NTP (accessed June 2013)