

Universidade do Minho
Escola de Engenharia
Departamento de Informática

Mestrado em Engenharia Informática

Henrique José Veiga Silva

Uma interface de gestão distribuída da ferramenta MORE

Dissertação de Mestrado

Supervisionado por:

Orientador: Prof. Dra. Solange Rito Lima

Co-orientador: João Marco Cardoso da Silva

Braga, December 25, 2014

Agradecimentos

A realização desta dissertação marca o fim de uma importante etapa da minha vida e só foi possível graças à colaboração e ao contributo, de forma directa ou indirecta, de várias pessoas. Sendo assim, gostaria de agradecer a todos aqueles que contribuíram de forma decisiva para a concretização deste trabalho.

Em especial, à minha orientadora, Solange Rito Lima, do Departamento de Informática da Universidade do Minho, pelo apoio, disponibilidade, dedicação, sugestões e conhecimentos transmitidos ao longo deste trabalho.

Ao meu co-orientador, João Marco Silva, em doutoramento no Programa Doutoral em Informática, na Universidade do Minho, pelos vários artigos cedidos e também por sempre me auxiliar nas incontáveis dúvidas que surgiram durante a realização deste trabalho.

Aos meus pais que sempre me incentivaram a continuar para alcançar os meus objectivos, aos pais da minha namorada pelo apoio, compreensão e carinho que foram muito importantes durante este processo, e restante família em geral.

Aos meus amigos, pelos momentos bem passados, e principalmente ao Nuno Pinto Miranda por me ajudar em alguns problemas e dúvidas que surgiram no decorrer do trabalho.

Finalmente, um agradecimento especial, à minha namorada pela constante presença, carinho, atenção, paciência e apoio prestado durante esta etapa académica tão importante.

Resumo

SILVA, H. J. V. **Uma interface de gestão distribuída da ferramenta MORE**. 2014. 120 f. Dissertação de Mestrado - Departamento de Informática, Universidade do Minho, Braga, 2014.

Com o constante crescimento do volume de tráfego, a amostragem do mesmo é um passo crucial para obter medições de redes escaláveis, abrangendo múltiplos desafios. Uma grande variedade de cenários previsíveis de amostragem exige uma visão modular das componentes de amostragem e suas características, fundamentadas numa arquitetura consistente. Articular o ambiente de medição, o modelo de informação necessário e a estratégia de amostragem adequada é o principal problema para alcançar uma solução de amostragem eficiente. Para ajudar a resolver este problema, foi proposta uma flexível *framework* de Monitorização Otimizada de Redes e Serviços de Comunicações (MORE), baseada numa arquitetura em camadas e numa taxonomia de técnicas de amostragem existentes, distinguindo as suas características internas, capaz de combinar as características identificadas [Silva *et al.* (2013a)].

No entanto, a gestão das configurações da respetiva *framework*, quando em larga escala, pode ser um processo demorado e complexo. Neste sentido, este projeto tem por objetivo o desenvolvimento de uma aplicação que permita efetuar a gestão de configurações de uma topologia distribuída de monitorização baseada na *framework* MORE. Esta aplicação deve ter por base uma solução padronizada de comunicação entre a entidade de gestão e os pontos de medição.

Palavras-chave: Amostragem de tráfego, Técnicas de amostragem, Medição de redes, IPFIX.

Abstract

Silva, H. J. V. **A Distributed Management Interface for MORE Tool**. 2014. 120 f. Master Thesis - Departamento de Informática, Universidade do Minho, Braga, 2014.

With the steady growth of network traffic volumes, sampling is a crucial technique for achieving scalable network measurements, which involves multiple challenges. The large variety of predictable sampling scenarios requires a modular view of sampling components and characteristics, sustained by a consistent architecture. Articulating the measurement environment, the required information model and the most suitable sampling strategy is the main problem to reach an efficient sampling solution. To help solving this problem, a flexible framework for optimizing the monitoring of networks and services (Monitorização Optimizada de Redes e Serviços de Comunicações - MORE) has been proposed [Silva *et al.* (2013a)].

This framework is based on a layered architecture and on a taxonomy of existing sampling techniques, being able to combine the identified sampling characteristics. However, managing and configuring MORE in large scale networks can be a long and complex process. In this context, the main objective of the present work is to develop an application for managing the configuration of a distributed monitoring topology based on MORE. This application relies on normalized solutions for ruling the interaction between the management entity and the measurement points.

Keywords: Traffic sampling, Sampling techniques, Network measurements, IPFIX.

Conteúdo

Agradecimentos	iii
Resumo	v
Abstract	vii
Lista de Abreviaturas	xi
Lista de Figuras	xiii
Listings	xvi
1 Introdução	1
1.1 Motivação e objetivos	2
1.2 Organização da dissertação	3
2 Estado da arte	5
2.1 Amostragem de tráfego	5
2.1.1 Amostragem Sistemática	5
2.1.2 Amostragem Aleatória	6
2.1.3 Amostragem Adaptativa	6
2.1.4 Amostragem Multi-adaptativa	7
2.2 Taxonomia de amostragem de tráfego	7
2.2.1 Framework de Amostragem	8
2.3 Coleta de pacotes ou fluxos	8
2.4 Comunicação com pontos de medição	9
2.4.1 IPFIX	9
2.4.2 Yet Another Flowmeter (YAF)	9
2.4.3 System for Internet Level Knowledge (SiLK)	9
2.5 Outras soluções de monitorização	10
2.5.1 Netflow	10
2.5.2 Sampled netflow	10
2.5.3 Nagios	10
2.5.4 Cacti	11
2.5.5 Zabbix	11
2.5.6 MRTG	11

2.5.7	Netflow Sensor	11
2.6	Sumário	12
3	Arquitectura da solução	13
3.1	Objetivos de conceção	13
3.2	Interação com os pontos de medição	14
3.2.1	Primeira arquitetura analisada	14
3.2.2	Solução adotada	15
3.3	Armazenamento de dados	16
3.4	<i>Interface</i>	16
3.4.1	Privilégios de administrador	16
3.4.2	Privilégios de gestor	17
3.5	Sumário	19
4	Funcionalidades da aplicação	21
4.1	Controlo de acessos	21
4.2	Menu principal	21
4.2.1	Adicionar ou remover utilizador	22
4.2.2	Adicionar ou remover ponto de medição	23
4.2.3	Acções sobre pontos de medição	23
4.2.4	Reconfigurar serviços	25
4.2.5	Registo de <i>logs</i>	27
4.3	Sumário	27
5	Testes e resultados	29
5.1	Gestão de utilizadores	29
5.2	Gestão de pontos de medição	29
5.2.1	Adicionar ou remover pontos de medição	29
5.2.2	Acções sobre pontos de medição	30
5.3	Reconfigurar serviços	31
5.4	Registo de <i>logs</i>	31
5.5	Sumário	31
6	Conclusões e trabalho futuro	33
6.1	Sugestões para Pesquisas Futuras	34
A	Guia de instalação	35
A.1	Servidor Web	35
A.2	Ponto de medição	42
	Bibliografia	47

Lista de Abreviaturas

CSS	<i>Cascading Style Sheets</i>
DDoS	<i>Distributed Denial-of-Service</i>
GPL	<i>General Public License</i>
HTTP	<i>Hypertext Transfer Protocol</i>
HTTPS	<i>Secure Hypertext Transfer Protocol</i>
IETF	<i>Internet Engineering Task Force</i>
IP	<i>Internet Protocol</i>
IPFIX	<i>Internet Protocol Flow Information Export</i>
LAMP	<i>Linux-Apache-MySQL-PHP</i>
MORE	<i>Monitorização Optimizada de Redes e Serviços de Comunicações</i>
MRTG	<i>Multi Router Traffic Grapher</i>
PCAP	<i>Packet Capture</i>
PDO	<i>PHP Data Objects</i>
QoS	<i>Controlo de qualidade</i>
RAM	<i>Random-access memory</i>
SCTP	<i>Stream Control Transmission Protocol</i>
SHA	<i>Secure Hash Algorithm</i>
SiLK	<i>System for Internet Level Knowledge</i>
SLA	<i>Service-level Agreement</i>
SNMP	<i>Simple Network Management Protocol</i>
SPAN	<i>Switched Port Analyzer</i>
TCP	<i>Transmission Control Protocol</i>
UDP	<i>User Datagram Protocol</i>
XML	<i>Extensible Markup Language</i>
YAF	<i>Yet Another Flowmeter</i>

Lista de Figuras

1.1	Tarefas de rede [Silva <i>et al.</i> (2013a)]	1
2.1	Taxonomia de amostragem [Silva <i>et al.</i> (2014b)]	7
3.1	Primeira arquitetura analisada	14
3.2	Arquitetura da solução adotada	15
3.3	Fluxograma da aplicação Web para um utilizador do tipo "administrador"	17
3.4	Fluxograma da aplicação Web para um utilizador do tipo "gestor"	18
4.1	Página que permite efetuar login na aplicação	21
4.2	Página que permite acesso a todas as funções da aplicação	22
4.3	Página que permite inserir ou remover utilizadores	23
4.4	Página de inserção e remoção de pontos de medição (bloco de inserção)	24
4.5	Página de inserção e remoção de pontos de medição (bloco de remoção)	24
4.6	Página de seleção de pontos de medição que serão alterados	26
4.7	Página de seleção da técnica a aplicar sobre vários pontos de medição	26
4.8	Página de inserção de argumentos mediante a técnica selecionada	26
4.9	Esquema de exportação de <i>logs</i>	27
5.1	Página de iniciar, parar e visualizar o estado de pontos de medição do utilizador "admin"	30
5.2	Página de iniciar, parar e visualizar o estado de pontos de medição do utilizador "antonio"	30
5.3	Processos JAVA na máquina com nome "S0"	31
5.4	Processos JAVA na máquina com nome "S1"	31
5.5	Estado do processo do <i>Collector</i> antes e depois de atualizar o serviço na aplicação	32
5.6	Conteúdo existente no ficheiro de exportação de <i>logs</i>	32

Listings

A.1	Instalação do <i>lamp server</i>	35
A.2	Instalação do <i>libssh</i>	36
A.3	Comandos para instalar o módulo "php5-dev"	36
A.4	Linha a acrescentar no ficheiro "php.ini"	36
A.5	Comandos para instalar o SSH2	37
A.6	Comando para reiniciar o <i>apache</i>	37
A.7	Comandos para copiar o código fonte para o <i>apache</i>	37
A.8	Ligação à base de dados	38
A.9	Comandos SQL para criar a base de dados e sair	38
A.10	Comando para importar a base de dados	38
A.11	Comando de instalação do cliente SSH	38
A.12	<i>Download</i> da ferramenta "libfixbuf"	38
A.13	Comandos para instalar o libfixbuf	39
A.14	<i>Download</i> das ferramentas netsa-python e SiLK	39
A.15	Instalação o compilador Python e da biblioteca Python do NetSA	39
A.16	Comandos para instalação do SiLK	40
A.17	A criar a diretoria onde são armazenados os fluxos que chegam dos pontos de medição	40
A.18	Configuração da variável "LD_LIBRARY_PATH"	40
A.19	Conteúdo a colocar no ficheiro "/etc/ld.so.conf.d/silk.conf"	40
A.20	Comandos para atualização do <i>path</i> das bibliotecas	40
A.21	Comandos para criar o sistema de ficheiros de configuração do SiLK	41
A.22	Ficheiro de configuração "rwflowpack.conf"	41
A.23	Configuração e arranque do <i>rwflowpack</i>	41
A.24	Configuração do arranque <i>on boot</i> do <i>rwflowpack</i>	42
A.25	A atualizar os <i>symbolic links</i>	42
A.26	Comando de instalação do servidor SSH	42
A.27	Comando necessário para configurar <i>root password</i>	42
A.28	Comando de instalação dos compiladores	43
A.29	Comando de instalação dos <i>build-essential</i>	43
A.30	<i>Download</i> da ferramenta "libfixbuf"	43
A.31	Comandos para instalar o libfixbuf	43
A.32	<i>Download</i> da ferramenta "YAF"	44
A.33	A abrir o ficheiro ".bashrc" como administrador	44
A.34	Instrução a colocar no ficheiro ".bashrc"	44
A.35	Comandos para instalar o libfixbuf	44

A.36 Comando para actualização do <i>path</i> das bibliotecas	44
A.37 Comando para instalação do "openjdk-7-jdk"	45
A.38 Comandos para instalação do "jpcap"	45
A.39 A criar o sistema de diretorias da <i>framework</i>	45

Capítulo 1

Introdução

O constante crescimento no volume de dados produzido por diferentes serviços de comunicações transformou a amostragem de tráfego numa estratégia fundamental para a redução dos custos associados à monitorização e engenharia das redes de computadores, sobretudo nas redes de alto débito. O uso de amostragem consiste em estimar parâmetros de interesse sobre uma rede recorrendo apenas a um subconjunto dos pacotes que por ela transitam.

Existe, atualmente, um grande número de técnicas de amostragem de tráfego que procuram, geralmente, obter a melhor acurácia possível na estimação de um parâmetro de interesse ou um conjunto reduzido deles. No entanto, a natureza heterogénea do tráfego e os requisitos de monitorização dos vários serviços existentes nestas redes exige uma solução flexível de uso das técnicas de amostragem, procurando aproveitar suas melhores características de acordo com o parâmetro observado. Atualmente, a amostragem de tráfego sustenta uma larga gama de tarefas de rede (ver Figura 1.1). Por exemplo, a sua utilidade tem sido explorada na:

- Engenharia de tráfego para suportar a classificação e caracterização de tráfego [Tammamaro *et al.* (2012)];
- Segurança de redes para deteção de intrusões e de anomalias, *botnet* e identificação de *Distributed Denial-of-Service* (DDoS) [Androulidakis *et al.* (2009)];
- Controlo de Qualidade de Serviço (QoS) e *Service-level Agreement* (SLA) para estimar cada parâmetro como atraso, variação no atraso e perda de pacotes [Sommer *et al.* (2005)], [Hu *et al.* (2008)].

Face a essa multiplicidade de contextos, um aspeto chave para a condução dos esforços de amostragem de rede deve ser centrado no estabelecimento de uma adequada relação entre o âmbito da amostragem, o modelo de informação necessário e a estratégia de amostragem a adotar.

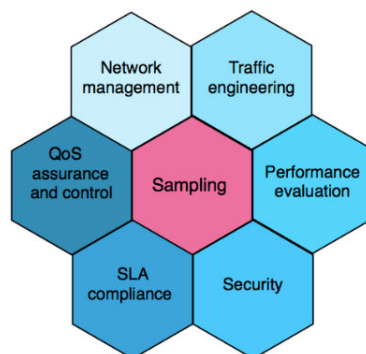


Figura 1.1: Tarefas de rede [Silva *et al.* (2013a)]

Apesar da amostragem de tráfego ser uma estratégia obrigatória para as redes de hoje e as da próxima geração, existem grandes desafios dificilmente abordados nas técnicas atuais e clássicas, nomeadamente:

- as técnicas de amostragem visam estimar corretamente o desempenho de parâmetros de rede, contudo, o seu principal alvo não é sobre como otimizar o *overhead* associado ao volume de dados envolvido no processo de recolha das amostras;
- a complexidade e os requisitos computacionais dos algoritmos de amostragem são geralmente pouco analisados, e ainda não está claro qual é o algoritmo que é mais adequado para lidar com cada tipo de tráfego [Tammamaro *et al.* (2012)];
- as técnicas de amostragem correntes tocam apenas uma única atividade de monitorização de rede.

Portanto, uma solução modular de medição baseada em amostragem de tráfego capaz de suportar de forma eficiente uma ampla gama de tarefas de rede, é ainda uma questão em aberto.

Considerando a importância e os desafios na medição de tráfego em redes multiserviço, foi proposta a *framework* MORE [Silva *et al.* (2013a)], disponibilizando esta uma solução para otimizar as tarefas de monitorização de rede, em redes de larga escala e de alta velocidade, recorrendo a técnicas de amostragem de tráfego abrangentes, desenvolvidas para melhorar a acurácia e reduzir os custos computacionais. Baseado no trabalho efetuado até ao momento, foi proposta uma arquitetura de medição de três camadas, abordando componentes chave que sustentam uma versátil estratégia de medição baseada em amostragem. Foi proposta uma taxonomia das técnicas de amostragem, que sustenta o desenvolvimento da *framework* de amostragem de tráfego MORE, e uma nova técnica multi-adaptativa que supera as técnicas de amostragem clássicas, em acurácia e no volume de dados envolvidos no processo de medição. Portanto, a MORE é capaz de integrar, de forma modular e distribuída, as diferentes propriedades comuns às técnicas de amostragem e, assim, permitir o uso das várias técnicas atualmente disponíveis, além de propostas futuras, de forma simples e escalável.

1.1 Motivação e objetivos

Estimar o tráfego total a partir de uma amostra será um grande passo para uma monitorização de redes eficiente, quer das redes atuais, quer das redes de próxima geração. O desenvolvimento da *framework* MORE surge neste contexto para trazer flexibilidade à tarefa de monitorização baseada em amostragem de tráfego.

Contudo, uma vez que a *framework* MORE [Silva *et al.* (2013a)] apenas se encontra acessível a partir de uma consola, o seu processo de configuração pode ser penoso, pois exige um período de adaptação por parte do utilizador. Este necessita de entender o que cada comando faz e o que tem disponível. Este problema aumenta ainda mais quando, está envolvido um conjunto alargado de pontos de medição, ou seja, cada vez que o gestor de rede necessitar de efetuar alguma alteração da configuração dos pontos de medição terá de se deslocar a todos eles, ou configurar remotamente um a um.

Desta forma, este projeto tem como principal objetivo, o desenvolvimento de uma solução escalável para a configuração distribuída da *framework* MORE e aquisição de informações de monitorização fornecidas por ela. Esta solução terá como requisito o uso de mecanismos padronizados para troca de informações sobre tráfego de rede, por exemplo, o *Internet Protocol Flow Information Export* (IPFIX) [Quittek *et al.* (2004)], e a sua adaptação para uso na configuração dos pontos de medição distribuídos pela rede. Tal padronização proporcionará homogeneidade suficiente para que diferentes áreas e ferramentas de gestão de rede possam usar a amostragem de tráfego de forma unificada e transparente. Em mais detalhe, pretende-se:

- desenvolver uma solução escalável de gestão da *framework* MORE, permitindo a interação com os diversos pontos de medição distribuídos pela rede;

- desenvolver uma interface de gestão de configurações, aquisição, tratamento e visualização de resultados.

Para a concretização destes objetivos é necessário:

- identificar e explorar protocolos e ferramentas padronizadas para troca de informações sobre tráfego de rede;
- desenvolver uma aplicação Web capaz de efetuar a configuração distribuída da *framework* MORE e permitir a visualização dos resultados obtidos pela respetiva *framework*. Uma aplicação Web possibilita o acesso em qualquer lugar, sem necessidade de instalar qualquer ferramenta desde que o utilizador tenha um *browser* instalado e acesso à Internet.

Com uma interface de gestão distribuída da ferramenta MORE, a dificuldade identificada acima de configuração de múltiplos pontos de medição será facilmente ultrapassada, ou seja, será disponibilizada uma interface gráfica atraente que permitirá:

- visualizar quais as configurações atuais dos pontos de medição;
- adicionar ou remover pontos de medição;
- configurar um ou mais pontos de medição de uma só vez;
- exportar informação:
 - exportar todas as capturas registadas;
 - exportar informação previamente selecionada;
- avaliar graficamente os resultados relativos a todos os pontos de medição.

1.2 Organização da dissertação

Com o objetivo de proporcionar um melhor entendimento dos diferentes temas abordados, a dissertação foi estruturada em seis capítulos. Os primeiros dois capítulos referem-se à introdução do trabalho, onde são abordados conceitos e revisões dos temas fundamentais relacionados com o assunto. Os restantes quatro capítulos referem-se ao trabalho desenvolvido e às conclusões obtidas. Em mais detalhe:

No Capítulo 1 é feita a introdução ao tema, um enquadramento conceptual assim como a identificação da motivação e dos objetivos a cumprir.

No Capítulo 2 são abordados os fundamentos teóricos associados a amostragem de tráfego, exportação de dados e outras soluções de monitorização de redes.

No Capítulo 3 é apresentada a solução desenvolvida, todo o procedimento prático, nomeadamente, a interação com os pontos de medição.

No Capítulo 4 são apresentadas todas as funcionalidades disponibilizadas pela aplicação Web.

No Capítulo 5 são descritos alguns dos testes efetuados e os resultados obtidos.

Finalmente no capítulo 6 é apresentada a conclusão e propostas para trabalho futuro, com o objetivo de melhorar ou adicionar funcionalidades.

Capítulo 2

Estado da arte

Este capítulo descreve técnicas usadas para medição de tráfego baseadas em amostragem, assim como a taxonomia que as caracteriza. Posteriormente são comparadas as vantagens e desvantagens de coletas orientadas ao pacote e ao fluxo. Seguidamente são apresentadas as ferramentas e protocolos que sustentam a comunicação entre os pontos de medição. Finalmente é efetuada uma descrição de outras soluções de monitorização.

2.1 Amostragem de tráfego

Com o crescimento do número de dispositivos pessoais e de serviços disponíveis na Web, nomeadamente os serviços de *streaming*, verifica-se um conseqüente aumento significativo do volume de dados que circula numa rede de alto débito. Neste contexto surgem as técnicas de medição passiva baseadas em amostragem, pois revelam-se capazes de reduzir os custos de medição associados a partir de uma amostra, com grande acurácia, o que as torna cada vez mais utilizadas.

A amostragem consiste em estimar parâmetros de interesse sobre uma rede recorrendo apenas a um subconjunto dos pacotes que por ela transitam [Thompson (2002)], ou seja, dependendo da técnica em uso, o ponto de medição considera ou não o pacote. Desta forma, com o uso destas técnicas pretende-se obter níveis de acurácia semelhantes aos que se obteria se fosse usado todo o tráfego que circula na rede, enquanto se reduz os custos computacionais intrínsecos à sua monitorização [Thompson e Sever (1996)].

As técnicas de amostragem, geralmente, caracterizam-se de acordo com o método de seleção de pacotes usado. As técnicas de amostragem mais usadas são as técnicas convencionais, nomeadamente, sistemáticas ou aleatórias [Zseby *et al.* (2009)]. Estas apresentam baixa complexidade, baseando-se em regras fixas para determinar quando o pacote é selecionado ou ignorado. Durante o processo de seleção dos pacotes os aspetos a ter em conta são o tempo de chegada (relativo ou absoluto), a posição, o conteúdo do pacote, ou uma composição destes critérios.

Nas subsecções que se sequegem apresentam-se com mais detalhe as diferentes abordagens de amostragem.

2.1.1 Amostragem Sistemática

A amostragem sistemática é uma amostragem gerida por uma função determinística e divide-se em dois tipos de amostragem, nomeadamente, a amostragem *time-based* e a *count-based*. O princípio de funcionamento de ambas as amostragens é muito semelhante, mas focado em pontos distintos, ou seja, ambas têm um comportamento cíclico, começando a decrementar o valor definido no mecanismo de seleção, mas a *count-based* captura o último pacote cuja contagem decrescente do número de pacotes deu zero e a *time-based* captura o último pacote cuja contagem decrescente do tempo de chegada deu zero.

Dado que este tipo de amostragem tem um comportamento cíclico, fixo e independente do conteúdo, é uma técnica simples de desenvolver e implementar. No entanto, devido às suas caracte-

terísticas, pode ser uma técnica tendenciosa ou viciada e de acordo com [Duffield (2004)] potencia a manipulação. Como para efeito de contagem de fluxos o número de amostras está relacionado com a acurácia pretendida [Choi e Bhattacharyya (2005)] e os recursos computacionais são diretamente proporcionais à frequência de amostragem [Silva *et al.* (2014a)], conclui-se que a acurácia desta técnica tem relação direta sobre os recursos necessários para o ponto de medição operar.

2.1.2 Amostragem Aleatória

A amostragem aleatória caracteriza-se por selecionar os pontos de partida dos intervalos de amostragem de acordo com um processo aleatório, de modo a convergir para uma taxa de amostragem necessária, assegurando que a distribuição de frequência de amostragem é limitada pelos valores máximos e mínimos. Esta amostragem tipicamente recorre a um pseudo gerador aleatório ou a uma função probabilística.

A abordagem aleatória *n-out-of-N*, apresentado em [Tanmaro *et al.* (2012)], tenta evitar a previsibilidade escolhendo valores exponencialmente ou geometricamente distribuídos [Duffield (2004)]. Segundo [Zseby *et al.* (2009)] na abordagem *n-out-of-N* são selecionados n pacotes de um intervalo N de pacotes, gerando números na gama $[1, N]$ e depois seleciona todos os pacotes que têm a posição do pacote correspondente. O *Sampled NetFlow*, [Choi e Bhattacharyya (2005)] desenvolvido pela Cisco Systems, é um exemplo de uma ferramenta que se baseia nesta técnica.

Relativamente à abordagem probabilística, a decisão sobre a frequência de amostragem segue uma função densidade de probabilidade pré-definida. A função probabilística divide-se em dois tipos:

- probabilística **uniforme** - todos os pacotes têm a mesma probabilidade de ser selecionados;
- probabilística **não uniforme** - os pacotes têm diferentes probabilidades de seleção. Para mais detalhes sobre este tipo de amostragem ver [Zseby *et al.* (2009)].

De acordo com [Zseby *et al.* (2009)], quando o tráfego apresenta um comportamento uniforme, verificam-se melhores níveis de acurácia na representação do tráfego total, quando usada uma função aleatória em vez de uma técnica sistemática. Contudo, como demonstrado em [Chabchoub *et al.* (2007)], para frequências de amostragem equivalentes, as técnicas aleatórias exigem recursos computacionais ligeiramente superiores do que os das técnicas de amostragem sistemática. Esta carga computacional pode variar de acordo com a complexidade da função probabilística escolhida. De acordo com [Tanmaro *et al.* (2012)] e [Chabchoub *et al.* (2007)], não existem evidentes vantagens que levem a escolher uma técnica de amostragem aleatória ou sistemática para classificação ou caracterização de tráfego.

2.1.3 Amostragem Adaptativa

A técnica adaptativa, considera os níveis de atividade da rede e com base na mesma, ajusta o seu processo de captura. Tal flexibilidade procura identificar os subconjuntos de pacotes mais importantes de uma *stream* de tráfego de acordo com as necessidades de medição e poupar recursos durante os períodos críticos. Esta técnica efetua constantes estatísticas sobre parâmetros específicos, nomeadamente, variação do atraso [Dogman *et al.* (2010)], atraso ou perda de pacotes [Serral-Gracia *et al.* (2008)]. De acordo com [Hernandez A. *et al.* (2001)], a flexibilidade inerente às técnicas adaptativas possibilita, em alguns casos, a redução do número de capturas efetuadas para a metade, mantendo níveis de acurácia semelhantes, o que se traduz numa clara redução dos custos computacionais.

As técnicas de amostragem de pacotes adaptativas podem-se basear na lógica *fuzzy* [Jiang *et al.* (2002)], numa predição linear [Lu e He (2010)] [Wei *et al.* (2010)], ou noutras estratégias adaptativas específicas e em mecanismos que consideram o comportamento do tráfego, o conteúdo do pacote ou o estado da rede para configurar as mudanças na amostragem de referência. A lógica de *fuzzy* traduz-se na aplicação de frequências de amostragem baseadas no histórico do comportamento da rede. Esta abordagem tende a exigir mais recursos como bases de dados, com

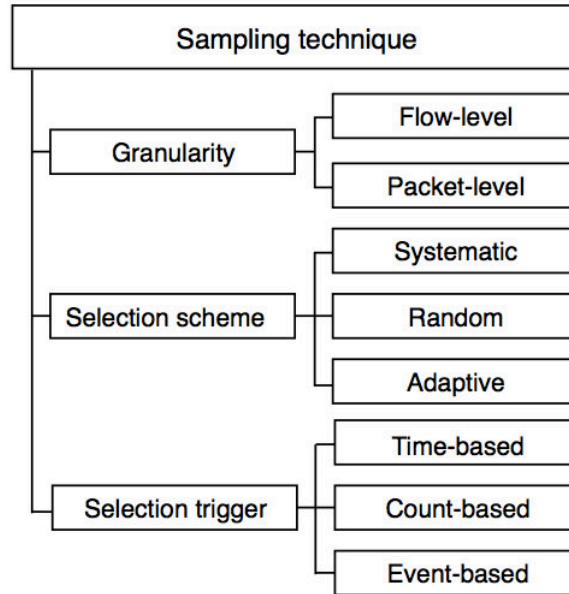


Figura 2.1: Taxonomia de amostragem [Silva et al. (2014b)]

grande volume de dados, pois exige que seja armazenada informação de longa data, já que o sistema atua de acordo com ocorrências passadas.

Com o método de predição linear, é calculado um parâmetro de referência a partir de amostras de tráfego, de modo a prever a frequência que melhor se adequa para capturar as sequências de pacotes seguintes.

Embora o uso de uma estratégia adaptativa possa pressupor o consumo de altos recursos, algumas técnicas podem constituir um custo computacional inferior, nomeadamente, o uso de processador e consumo de memória, mesmo quando comparado com as técnicas clássicas [Silva et al. (2014a)].

2.1.4 Amostragem Multi-adaptativa

Em conformidade com o que foi dito na Subsecção 2.1.3, o método de predição linear calcula um parâmetro de referência a partir das amostras de tráfego, de modo a prever a frequência que melhor se adequa para capturar os pacotes seguintes, ou seja, apenas os intervalos entre amostras variam, mantendo os tamanhos das amostras. A técnica multi-adaptativa descrita em [Silva et al. (2013b)] tem um funcionamento muito semelhante ao descrito, no entanto, esta manipula os intervalos entre amostras e os tamanhos das amostras, o que permite reduzir a carga computacional associada à medição, mantendo a acurácia na caracterização do tráfego.

2.2 Taxonomia de amostragem de tráfego

Com base no estudo das características das diferentes técnicas de amostragem existentes, em [Silva et al. (2013a)] foi proposta uma taxonomia de amostragem de tráfego. A taxonomia definida divide as técnicas de amostragem em três componentes bem definidos de acordo com a *granularidade*, *selection scheme* e *selection trigger* em uso. Cada uma destas componentes está dividida num subconjunto abordagens.

Na prática, esta taxonomia é um modelo devidamente estruturado, a partir do qual torna possível desenhar as componentes existentes nas técnicas de amostragem tradicionais e futuras. A Figura 2.1 apresenta uma visão geral sobre a taxonomia descrita. Em mais detalhes:

- *granularidade* - identifica a atonicidade do elemento em análise no processo de amostragem: Segundo a abordagem *Flow-level* apenas são selecionados os pacotes pertencentes a um con-

junto de fluxos; Na abordagem *Packet-level* os pacotes são selecionados como entradas únicas independentes.

- *selection scheme* - identifica a função de seleção que determina quais pacotes são colecionados. De acordo com a Figura 2.1, esta componente encontra-se dividida em três subconjuntos que são as técnicas de amostragem *Systematic*, *Random*, *Adaptative*. Estas técnicas são descritas com maior detalhe na Secção 2.1. A definição da técnica *Adaptative* é apresentada nas Subsecções 2.1.3 e 2.1.4. Tratam-se de técnicas distintas, mas a técnica multi-adaptativa é uma derivação da adaptativa.
- *selection trigger* - determina os limites espaciais e temporais de amostragem; Em conformidade com a Figura 2.1 esta componente divide-se em três subconjuntos distintos: *Time-based*, *Count-based* e *Event-based*, ou seja, os limites podem ser baseados em espaçamento temporal, baseada em contagem ou num evento respetivamente.

2.2.1 Framework de Amostragem

A *framework* de amostragem MORE trata-se de um desenvolvimento prático da taxonomia de amostragem apresentada na Secção 2.2.

De modo a tornar a *framework* compatível com o maior número de equipamentos possível, esta usou como interface entre o sistema de medição e a interface de rede em que será aplicada o *libpcap*. A implementação da taxonomia e as suas interações foram desenvolvidos em Java recorrendo à biblioteca Jpcap para comunicar com o *libpcap*.

A *framework* permite combinar os componentes de amostragem definidos na taxonomia. Esta combinação pode ser aplicada em cenários de medição *online* e *offline*, por exemplo, ocorrendo em tempo real ou com base em coletas previamente coletadas.

2.3 Coleta de pacotes ou fluxos

Para a analisar os dados que circulam num comutador de pacotes deve-se ter em conta as seguintes granularidades: orientada ao pacote ou ao fluxo. Ambas têm vantagens. Se se pretender, por exemplo, detetar uma falha de rede é conveniente usar uma coleta orientada ao pacote, pois estas coletas contêm informações detalhadas do que aconteceu na rede. Contudo, numa rede de alto débito, devido ao grande volume de tráfego, pode não ser viável manter esta abordagem, quer seja pelo processamento quer pelo armazenamento de coletas envolvido. É por isso comum o uso de coletas por fluxo, por exemplo, o NetFlow.

Existem várias definições do termo "fluxo" usadas pela comunidade na Internet. De acordo com [Lee e Brownlee (2007)], tipicamente define-se fluxo como uma sequência unidirecional de pacotes, em que todos os pacotes partilham as seguintes informações: endereço de IP de origem, endereço de IP de destino, porta de origem, porta de destino e protocolo.

Dentro do contexto de IPFIX, em conformidade com [Claise *et al.* (2013)] usa-se a seguinte definição:

Um fluxo é definido como um conjunto de pacotes ou *frames* que passam por um ponto de observação numa rede durante um intervalo de tempo. Todos os pacotes que pertencem a um fluxo particular têm um conjunto de propriedades comuns. Cada propriedade é definida como o resultado da aplicação de uma função para os valores de:

1. um ou mais campos do pacote da camada de rede (por exemplo, endereço de IP de destino), campos da camada transporte (por exemplo, número de porta de destino), ou campos da camada de aplicação (por exemplo, campos de cabeçalho RTP [Schulzrinne *et al.* (2003)]);
2. umas ou mais características do próprio pacote (por exemplo, o número de *labels* MPLS);
3. um ou mais dos campos derivados do tratamento de pacotes (por exemplo, o endereço do *next-hop* IP, a interface de saída).

Um pacote é definido como pertencente a um fluxo, desde que satisfaça todas as propriedades definidas pelo fluxo.

2.4 Comunicação com pontos de medição

O objetivo deste projeto é desenvolver uma aplicação de configuração e gestão distribuída da *framework MORE* que permitisse a integração com outros serviços de forma padronizada. Como tal, para o desenvolvimento desta aplicação recorreu-se a ferramentas *standard* e *freeware*. As subsecções que se seguem, apresentam as principais soluções usadas para o desenvolvimento da aplicação.

2.4.1 IPFIX

Desde há muito tempo atrás que existe necessidade de um *standard* de exportação para fluxos de informação baseados em IP. Com esta necessidade, o IETF desenvolveu o *Internet Protocol Flow Information Export*, conhecido por IPFIX. O IPFIX é um protocolo capaz de exportar informações de fluxos de rede de *routers*, pontos de medição, entre outros equipamentos [Mark *et al.* (2008)].

Um dispositivo IPFIX pode ser visto simplesmente como uma entrada que implementa o protocolo IPFIX. No dispositivo IPFIX, funcionalmente o protocolo reside na exportação de processos. O processo de exportação recebe registos de fluxos a partir de um *Metering Process* e envia-os para o *Collector*.

Em alto nível, um dispositivo executa as seguintes tarefas:

1. codifica a informação de controlo em *Templates*;
2. codifica os pacotes analisados nos pontos de medição em registos de fluxos;
3. transforma os *Templates* selecionados e os registos de fluxos em mensagens IPFIX;
4. envia as mensagens IPFIX para o *Collector*.

O protocolo IPFIX comunica informações de um *IPFIX Exporter* para um *IPFIX Collector*. Essas informações incluem não só registos de fluxos, mas também informações sobre o *Metering Process*. Estas informações, tipicamente chamadas *Control Information*, incluem detalhes dos campos de dados nos registos de fluxos. Podem também incluir as estatísticas do *Metering Process*, tais como, o número de pacotes perdidos [Sadavisan *et al.* (2009)].

2.4.2 Yet Another Flowmeter (YAF)

De acordo com [CERT (2006)], o *Yet Another Flowmeter* (YAF) é usado como um sensor de captura de informações de fluxos numa rede, que permite exportar as mesmas em formato IPFIX. Lê os dados dos pacotes a partir de um ficheiro PCAP ou captura em tempo real e exporta os dados via IPFIX sobre SCTP, TCP ou UDP.

Ainda de acordo com [CERT (2006)], o YAF é projetado para ser implantado em *white-box sensors* conectados a segmentos de rede local ou a *Switched Port Analyzer* (SPAN) simétricas em pontos de *routing*, suporta nativamente fluxos bidirecionais. A exportação *BiFlow* é feita através do método de exportação especificado em [Trammell e Boschi (2008)].

2.4.3 System for Internet Level Knowledge (SiLK)

De acordo com [CERT (2012)], o *System for Internet-Level Knowledge* (SiLK) trata-se de um conjunto de ferramentas de análise de tráfego desenvolvidas pelo grupo Network Situational Awareness do CERT para facilitar a análise de segurança em redes de larga escala. O SiLK facilita a recolha, armazenamento e análise de fluxos de rede, permitindo assim ao gestor de rede rapidamente consultar o histórico de grande volume de dados. Idealmente o SiLK é usado para analisar tráfego no *backbone*, fronteira de uma empresa, grande empresa distribuída ou ISP médio.

A instalação do SiLK envolve duas categorias de aplicações: o *packing system* e a *analysis suite*. O *packing system* coleciona pacotes IPFIX, NetFlow v9, ou NetFlow v5 e converte num formato mais eficiente a nível de espaço, efetuando o registo de pacotes num serviço binário específico de ficheiros *flat*. A *analysis suite* consiste em ferramentas que lêem esses ficheiros *flat* e realizam várias operações de consulta. A ferramenta de análise interopera com *pipes*, permitindo ao utilizador desenvolver consultas relativamente sofisticadas a partir de entradas simples.

2.5 Outras soluções de monitorização

Nas Subsecções que se seguem é apresentada uma descrição das ferramentas de monitorização mais usadas atualmente. Estas foram analisadas no sentido de fornecerem conhecimentos adicionais de aspetos relevantes a contemplar na aplicação a desenvolver.

2.5.1 Netflow

O NetFlow é um módulo de uma ferramenta de monitorização de tráfego desenvolvido pela Cisco Systems. Por defeito, o NetFlow processa todos os pacotes que entram no ponto de medição, mantendo incrementalmente estatísticas de dados sobre cada fluxo na memória *cache*. Este processo requer altos recursos de processamento, principalmente em redes de alto débito [Cisco Systems (2004)].

2.5.2 Sampled netflow

O Sampled NetFlow é uma ferramenta monitorização de redes de larga escala baseada em amostragem, desenvolvida pela Cisco Systems.

O Sampled NetFlow permite colecionar estatísticas NetFlow de um subconjunto de entradas de tráfego de uma interface, selecionando apenas um em "N" pacotes da sequência, onde "N" é um parâmetro configurável.

Esta amostragem de pacotes vai diminuir substancialmente o processamento necessário para a contagem dos pacotes NetFlow, permitindo assim que a maioria dos pacotes sejam trocados mais rapidamente, ou seja, o ponto de medição não vai necessitar do processamento adicional provocado pelo NetFlow [Cisco Systems (2003)].

2.5.3 Nagios

Nagios é um dos mais populares sistemas de monitorização Linux, com suporte *Simple Network Management Protocol* (SNMP), baseado em Web. Distribuído sob a licença *General Public License* (GPL), Nagios permite monitorizar a disponibilidade e o tempo de resposta dos serviços de rede, o uso de recursos do sistema, como a carga do processador, alocação de memória *Random-access memory* (RAM), entre outros. Usar esta ferramenta implica instalar uma instância Nagios principal (servidor) que coleciona informações a partir de qualquer sistema local ou clientes Nagios (agentes). Isso significa que é possível instalar o *software* cliente em sistemas Linux, Windows ou dispositivos Cisco, e ver o estado dos seus serviços e/ou processos de forma centralizada, pois trata-se de uma ferramenta de monitorização baseada em Web.

Em caso de qualquer falha ou anomalia detetada pelo servidor, o administrador receberá um alerta. O Nagios suporta uma grande variedade de alertas, incluindo e-mail, sms, mensagens de chat e notificações de chamadas de telefone.

Existe um grande número de *plugins* oficiais e de terceiros que podem estender as funcionalidades do Nagios. Este monitoriza estados, contudo não apresenta quaisquer gráficos, por exemplo, o uso da interface de rede [Galstad *et al.* (1996)].

2.5.4 Cacti

O Cacti é um sistema de monitorização baseado na web, escrito em PHP Hypertext Preprocessor e distribuído sob a licença GPL. Esta ferramenta foi projetado principalmente para os gráficos. Este usa a ferramenta RRDTool para construir os gráficos, o que permite facilmente anexar ao Cacti gráficos que estejam em ficheiros `*.rrd`, juntando tudo nesta aplicação.

Assim, é possível, por exemplo, ver gráficos de carga de processador, uso de memória RAM, estatísticas em tempo de ida e volta, e uso da largura de banda, colecionada a partir de vários *hosts*. O Cacti suporta SNMP, o que torna possível monitorizar praticamente qualquer dispositivo de rede: Unix, clientes Windows, dispositivos Cisco, equipamentos Juniper, telefones VoIP, routers, switches [Cacti Group (2004)].

Por defeito, o Cacti não suporta alertas, por isso, para usufruir deste serviço é necessário instalar um *plugin* de terceiros.

2.5.5 Zabbix

O Zabbix é uma classe empresarial de sistema de monitorização Linux, com uma grande lista de recursos disponíveis. Este está distribuído sob a licença GPL, com suporte SNMP e escrito em PHP. Uma vez que este serviço é disponibilizado via Web, é possível monitorizar os agentes a partir de qualquer local.

Com esta ferramenta é possível monitorizar vários parâmetros de rede, nomeadamente, a integridade e disponibilidade de servidores. De modo a conseguir um permanente acompanhamento do que está a acontecer na rede, este *software* disponibiliza um mecanismo flexível de configuração de alertas. Com base na informação guardada na base de dados, o Zabbix faz relatórios e disponibiliza uma ótima interface de visualização sobre os respetivos dados. Esta ferramenta mostra vários tipos de gráficos, incluindo as estatísticas de rede e a carga do processador [SIA (2001)].

2.5.6 MRTG

O Multi Router Traffic Graffer, conhecido por MRTG, distribuído sob a licença GPL, consiste num *script* Perl que usa SNMP para ler os contadores de tráfego do seus *routers* e um código escrito em C que regista os dados de tráfego e cria os gráficos que representam o tráfego na conexão de rede que está a ser monitorizada. Como estes gráficos estão incluídos em páginas Web, podem ser vistos a partir de qualquer local.

Além de uma detalhada visão diária, o MRTG também cria representações visuais dos últimos sete dias, das últimas 5 semanas e dos últimos 12 meses. Isto é possível porque o MRTG mantém um registo de todos os dados que recebeu do *router*. Estes registos são automaticamente consolidados de forma que o crescimento dos dados não é significativo ao longo do tempo, mas ainda assim contém todos os dados relevantes para todo o tráfego analisado ao longo dos últimos dois anos [OETIKER+PARTNER (2006)].

2.5.7 Netflow Sensor

O Netflow Sensor, conhecido por NfSen, é um Netflow *collector*, *open source*, baseado em Web e é distribuído sob a licença BSD. Este *software* é um *front-end* para a ferramenta nfdump NetFlow e permite efetuar as seguintes operações:

- apresentar dados NetFlow: Fluxos, pacotes e bytes através do RRD (Round Robin Database);
- navegar com facilidade através de dados NetFlow;
- processar os dados Netflow dentro de um intervalo de tempo especificado;
- configurar alertas com base em diversas condições;
- criar os próprios *plugins* para processar dados NetFlow num intervalo especificado.

Diferentes tarefas precisam de diferentes interfaces para os seus dados NetFlow. O NfSen mantém todas as funções do nfdump, quando se usa a linha de comandos e permite uma implementação dos dados NetFlow [[Teleinformatikdienste fuer Lehre und Forschung \(2004\)](#)].

2.6 Sumário

Neste capítulo foram abordadas as principais técnicas de amostragem existentes - sistemática, aleatória, adaptativa e multi-adaptativa. De seguida foi descrita a taxonomia de amostragem de tráfego e a sua representação prática através de uma *framework* de amostragem. Posteriormente foram apresentados os fundamentos teóricos e as ferramentas que proporcionaram a comunicação entre os pontos de medição e o *Collector*.

Por fim, foram apresentadas as principais ferramentas atualmente usadas para monitorização de redes.

No capítulo seguinte será apresentada a arquitetura da solução de interação com a *framework* MORE, ou seja, são enumerados os objetivos de conceção e é apresentado o modo como os pontos de medição interagem.

Capítulo 3

Arquitectura da solução

Neste capítulo são apresentadas as soluções analisadas para a conceção da arquitetura da aplicação para interação com a *framework* MORE, desde a fase de configuração dos pontos de medição até à visualização dos dados coletados.

3.1 Objetivos de conceção

Para elaborar este projeto pretende-se usar a *framework* proposta no artigo [Silva *et al.* (2013a)]. Esta *framework* materializa a taxonomia de técnicas de amostragem apresentada na Secção 2.2, foi implementada em Java, com recurso à classe *Jpcap* e permite uma combinação flexível de várias características envolvidas num processo de amostragem.

Para elaborar a interface de gestão distribuída da *framework* MORE, teve-se em atenção os seguintes objetivos de conceção:

- normalização;
 - formatos normalizados que facilitem a integração com ferramentas existentes;
 - coordenação da comunicação entre entidades;
- otimização do volume de dados envolvidos;
- versatilidade;
- escalabilidade.

Tendo em conta que se trata de uma ferramenta de amostragem, o volume de dados capturados pela *framework* será muito inferior ao volume de dados total que passa no ponto de medição. Contudo, tratando-se de uma rede de alto débito, entende-se que o volume de dados é ainda elevadíssimo. Assim, tendo em conta que esta ferramenta gera coletas orientadas ao pacote, a gestão centralizada de pontos de medição distribuídos torna-se ainda muito pesada. Como tal, decidiu-se acrescentar um mecanismo que a partir das coletas geradas pela *framework*, gere coletas orientadas ao fluxo, reduzindo assim drasticamente o volume de dados a armazenar no *Collector*. Com esta alteração de tipologia, uma vez que se deixa de ter toda a informação dos pacotes selecionados e passa-se a ter apenas informação global dos mesmos, perde-se em detalhe, mas ainda assim existem uma larga gama de anomalias que podem ser detetadas (por exemplo, *DoS*, *worms*, ou *scans*), garantindo-se escalabilidade e uma otimização do volume de dados envolvidos.

A comunicação entre o *Collector* e cada ponto de medição é feita com o *daemon* SiLK e o YAF. O *daemon* pode ser instalado na mesma máquina que o Servidor Web ou em máquinas distintas, e o YAF deve ser instalado em cada ponto de medição. Assim, é possível colocar no YAF, como *input*, a amostragem gerada pela *framework* MORE. Este gera coletas orientadas ao fluxo em formato IPFIX (para mais detalhes sobre IPFIX, ver a Subsecção 2.4.1) e envia-os para o *daemon* SiLK.

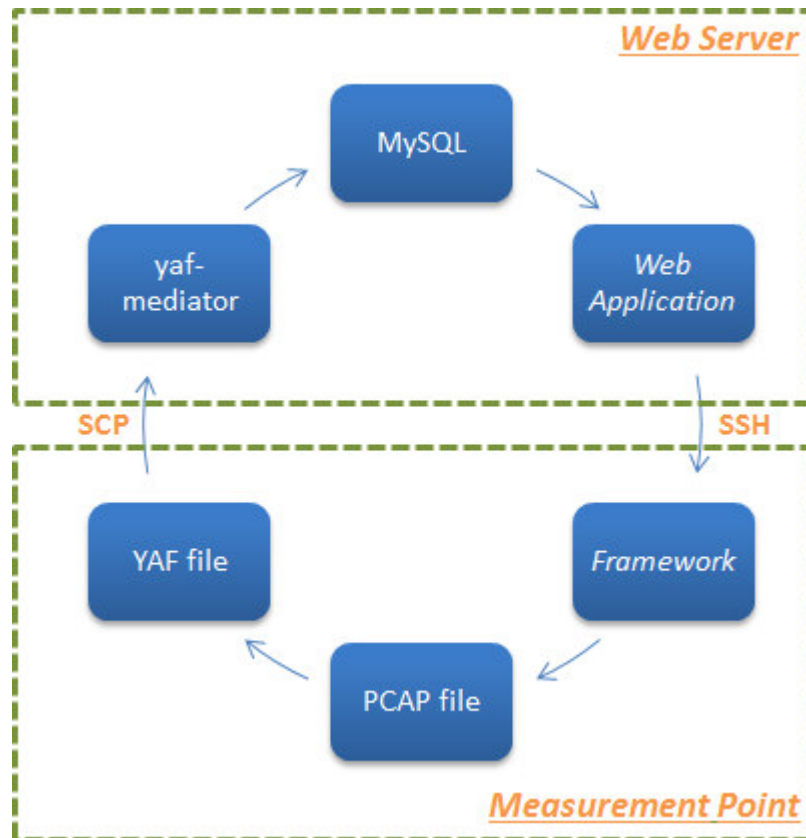


Figura 3.1: Primeira arquitetura analisada

3.2 Interação com os pontos de medição

Até chegar à solução final passou-se por várias possíveis soluções que, por algum aspeto, não foram consideradas as mais adequadas. Nesta Secção será abordada a interação com os pontos de medição, bem como as principais arquiteturas que conduziram até ao produto final.

3.2.1 Primeira arquitetura analisada

Como base desta arquitetura um Servidor Web configura os pontos de medição e posteriormente recebe os dados coletados.

Em conformidade com a figura 3.1, é possível verificar que a configuração é feita por *secure shell*, que envia todos os parâmetros de configuração necessários para o ponto de medição. Rececionados os parâmetros de configuração, a *framework* inicia o processo de amostragem, guardando a informação em formato PCAP na diretoria pretendida. Ainda durante o processo de estudo da solução, chegou-se à conclusão de que o formato PCAP consegue disponibilizar muita informação, mas o volume de dados aumenta muito rapidamente e ainda mais rápido quando a *framework* se encontra numa rede de *core*. Assim, de forma a reduzir o respetivo volume de dados, decidiu-se usar o YAF [CERT (2006)] (ver Subsecção 2.4.2).

Esta ferramenta processa dados de um arquivo PCAP ou de uma interface em captura em tempo real, em fluxos bidirecionais e exporta os dados para o formato IPFIX. Esta exportação pode ser feita para uma diretoria local ou para um *daemon*.

Neste caso, os fluxos são exportados para uma diretoria local e são enviados para o Servidor Web por *secure copy*. Rececionado o ficheiro com os fluxos, em formato IPFIX, é executado o *super_mediator*.

O *super_mediator* é uma ferramenta capaz de receber informação em formato IPFIX e exportar toda a informação para um ficheiro TXT ou carregar a informação numa base de dados MySQL.

Assim, pode-se inserir a informação numa base de dados a partir do *super_mediator*, possibilitando efetuar filtros sobre a informação e desenhar gráficos, facilitando a leitura dos resultados obtidos.

Contudo, esta arquitetura não segue as recomendações do IPFIX para a interoperabilidade com soluções existentes com suporte a este padrão. Além disso, como a comunicação entre a entidade de gestão e o ponto de medição é ativa, a aplicação tem de tomar a decisão de quando capturar as estatísticas de tráfego, que pode ser um processo pesado para a aplicação de gestão.

3.2.2 Solução adotada

A figura 3.2 apresenta a arquitetura da solução adotada. À semelhança do que aconteceu na arquitetura apresentada na Subsecção 3.2.1, a configuração da *framework* MORE é feita por *secure shell*. Recepcionados os parâmetros de configuração, a *framework* inicia o processo de amostragem, guardando a informação em formato PCAP. Com uma periodicidade configurada por defeito para 15 segundos, a *framework* invoca o YAF e este envia os fluxos resultantes em formato IPFIX para o *Collector* SiLK (ver Subsecção 2.4.3) que está idealmente está a correr no Servidor Web. Também é disponibilizada uma opção, em que o utilizador pode definir um *Collector* alojado numa máquina diferente do servidor Web. Contudo, a escolha desta opção tem a limitação de o utilizador não conseguir, através da aplicação Web, confirmar se o *Collector* está a receber informação dos pontos de medição (ver Subsecção 4.2.3).

A comunicação entre a entidade de gestão e o ponto de medição é passiva, seguindo as recomendações de interação e exportação de dados IPFIX.

Esta solução cumpre os objetivos de conceção delineados na medida em que se conseguiu versatilidade através do uso de ferramentas que respeitam protocolos padrão, escalabilidade e otimização do volume de dados envolvidos através do uso de coletas orientadas ao fluxo em vez de ao pacote, e normalização a dois níveis: formatos normalizados que facilitam a integração com ferramentas existentes; na coordenação da comunicação entre as entidades.

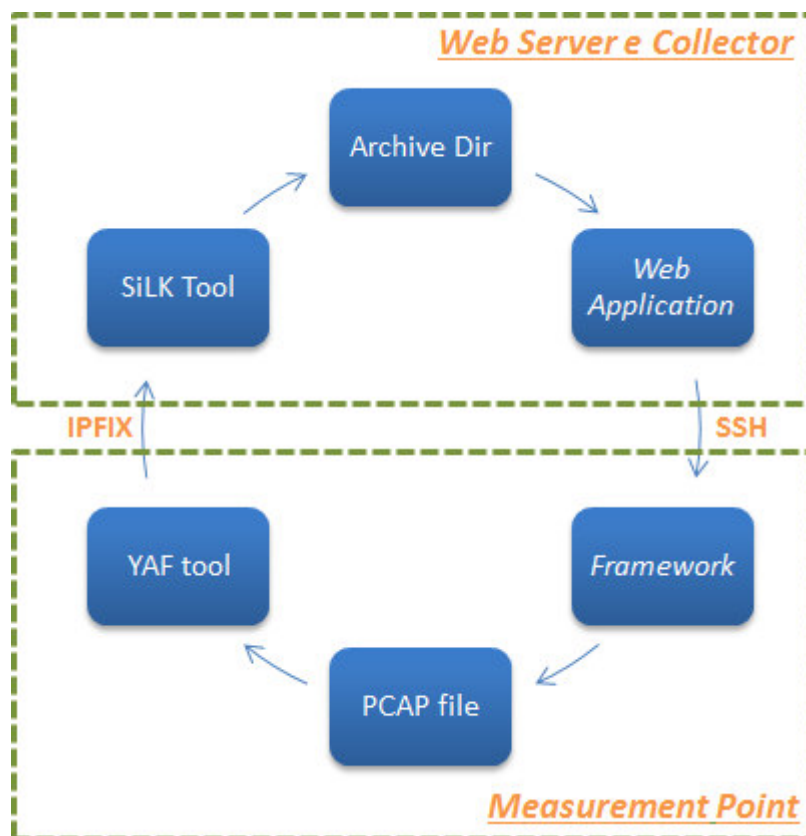


Figura 3.2: Arquitetura da solução adotada

3.3 Armazenamento de dados

Toda a informação de gestão de utilizadores, *logs* e pontos de medição da aplicação Web é guardada numa base de dados. Como a base de dados ainda não se encontra encriptada decidiu-se fazer a inserção do *hash* da password, em vez da password. Com esta solução, em caso da base de dados ser atacada, o atacante pode conseguir aceder aos dados dos utilizadores, mas não descobre a password. O resumo fez-se em *Secure Hash Algorithm* (SHA). Todos os dados são inseridos através da classe PDO do PHP, o que já resolve problemas associados a *SQL injection*.

O *Collector* guarda os respetivos fluxos num conjunto de diretorias devidamente estruturadas, ou seja, o tráfego é separado por tipo e cada tipo tem diretorias com o ano de captura. Cada ano tem diretorias com o mês de captura e abrindo esta última é possível verificar que contém os ficheiros com a informação proveniente dos pontos de medição. O nome destes ficheiros também não é atribuído ao acaso, pois existe necessidade de distinguir a origem do tráfego. O nome destes ficheiros é atribuído da seguinte forma: "«tipo de tráfego»-«nome do ponto de medição de origem»_«data da captura».«hora da captura»".

3.4 Interface

Esta secção apresenta uma visão geral sobre a aplicação Web, nomeadamente, as funções permitidas e os privilégios ou limitações dos diferentes tipos de utilizador.

A interface Web tem como base o pacote Linux-Apache-MySQL-PHP (LAMP) e a *framework bootstrap*. O LAMP disponibiliza um ambiente de desenvolvimento Web para sistemas operativos Linux. O LAMP permite criar aplicações Web com Apache2, PHP e MySQL. O *bootstrap* é um *front-end framework* que disponibiliza templates com design baseado em HTML, *Cascading Style Sheets* (CSS) e extensões JavaScript.

3.4.1 Privilégios de administrador

A figura 3.3 apresenta o fluxograma que ilustra as opções disponibilizadas ao utilizador do tipo "administrador". De acordo com a figura 3.3, após o sucesso no *login*, a aplicação encaminha para a área "home" que está dividida em cinco grandes módulos. Cada um destes módulos pode ou não encontrar-se subdividido:

1. Inserir ou remover utilizadores - neste módulo o administrador pode inserir e remover utilizadores;
2. Inserir ou remover pontos de medição e visualizar do estado de todos os pontos de medição - neste módulo o administrador pode inserir e remover pontos de medição. O administrador consegue inserir os seu próprios pontos de medição e inserir pontos de medição noutras utilizadores. O administrador consegue remover pontos de medição que estejam associados ao seu nome de utilizador e pontos de medição que estejam associados a outros nomes de utilizador. No bloco de remoção, o administrador tem a possibilidade de ver o estado de todos os pontos de medição;
3. Iniciar ou terminar a captura e visualização do estado dos pontos de medição - neste módulo o administrador pode iniciar ou parar o processo de amostragem de todos os pontos de medição. Estes pontos de medição podem estar associados ao seu nome de utilizador ou não. Nesta área o administrador também tem a possibilidade de ver o estado de todos os pontos de medição;
4. Inicia, pára, reinicia e reconfigura o *Collector* - neste módulo o administrador tem a possibilidade de iniciar, parar, reiniciar ou reconfigurar o *Collector*. Reconfigurar consiste em atualizar o ficheiro `"/etc/silk/sensor.conf"` de acordo com a informação que consta na base de dados e em seguida reiniciar o *Collector*;

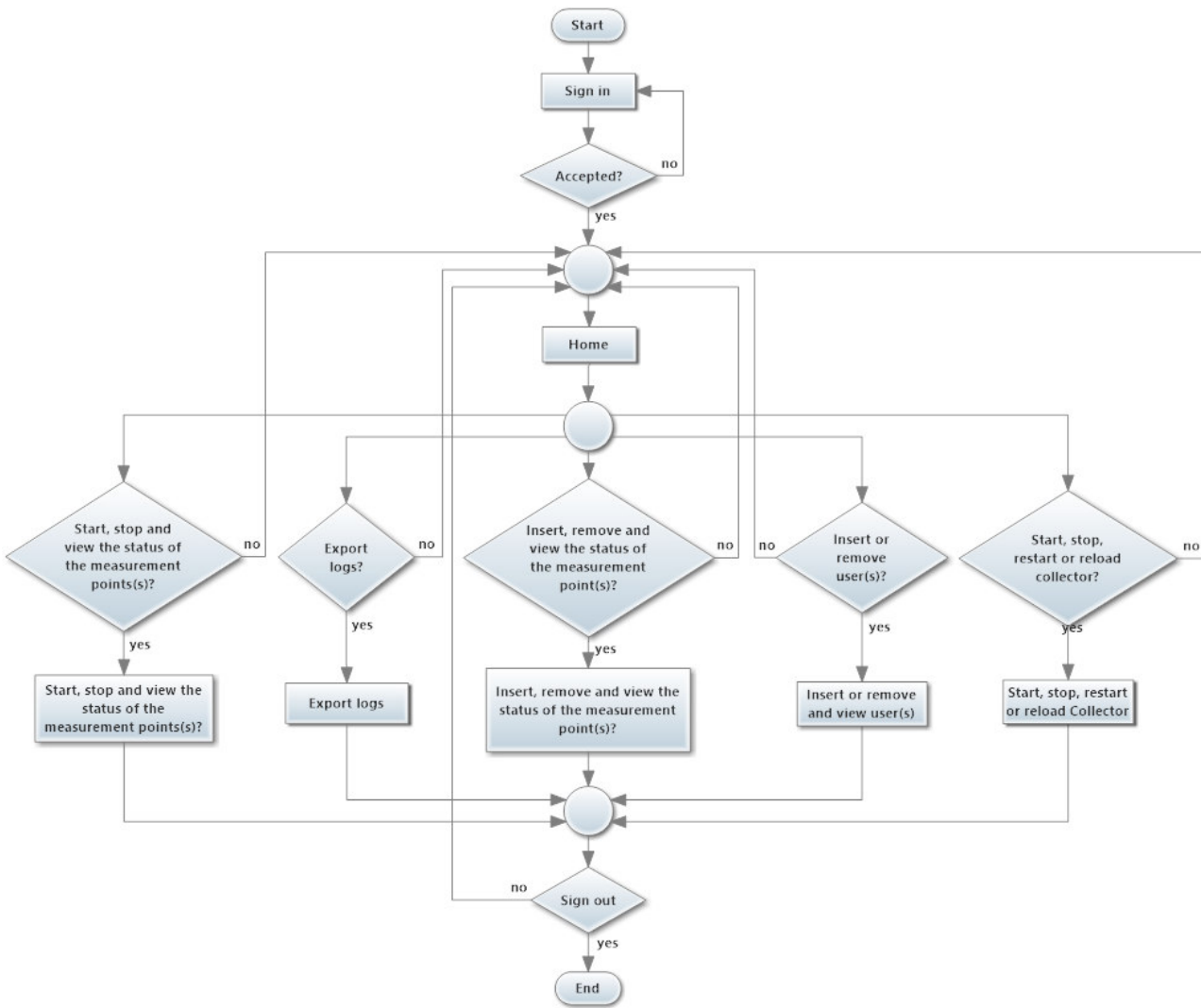


Figura 3.3: Fluxograma da aplicação Web para um utilizador do tipo "administrador"

5. Exportar o registo de *logs* - este módulo permite exportar o registo de todas as operações efetuadas na aplicação e está apenas disponível para um utilizador do tipo "administrador".

3.4.2 Privilégios de gestor

A figura 3.4 apresenta o fluxograma que ilustra as opções disponibilizadas o utilizador do tipo "gestor". De acordo com a figura 3.4, após o sucesso no *login*, a aplicação encaminha para a área "home" que está dividida em dois grandes módulos. Cada um destes módulos pode ou não encontrar-se subdividido:

1. Inserir ou remover pontos de medição e visualização do estado de todos os pontos de medição - neste módulo o gestor pode inserir, remover e visualizar apenas os seu próprios pontos de medição. O gestor deve ter consciência que após a inserção de um novo ponto de medição é necessário reconfigurar o serviço do *Collector* e não tem privilégios para isso. Assim, é enviada uma notificação para o administrador para reconfigurar o *Collector* e só depois de validada, o gestor poderá usar o seu novo ponto de medição;
2. Iniciar ou terminar a captura e visualizar do estado dos pontos de medição - neste módulo o gestor pode iniciar ou parar o processo de amostragem dos seus pontos de medição. Nesta área o gestor também tem a possibilidade de ver o estado dos seus próprios pontos de medição;

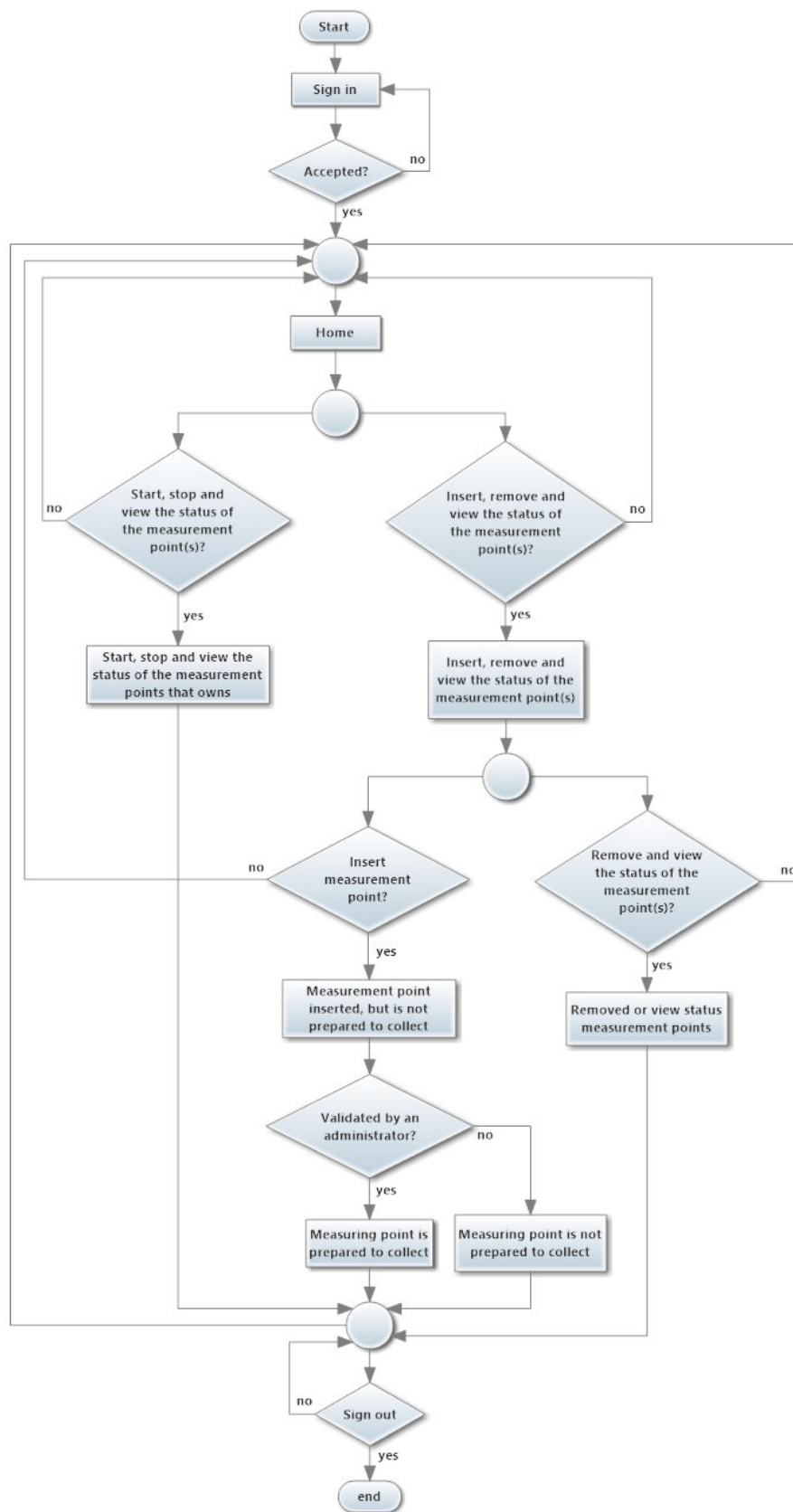


Figura 3.4: Fluxograma da aplicação Web para um utilizador do tipo "gestor"

3.5 Sumário

Neste capítulo foram apresentados os requisitos e as principais abordagens que conduziram para o desenvolvimento da solução adotada. Foi também apresentado o método usado para a comunicação entre o Servidor Web e os pontos de medição, e o método usado para a comunicação entre os pontos de medição e o *Collector*. De seguida foi explicado como o *Collector* armazena a informação capturada pelos pontos de medição. Seguidamente são apresentadas todas funcionalidades da aplicação Web e privilégios dos utilizadores.

No capítulo seguinte serão apresentadas com mais detalhe a interface desenvolvida e as funções que a aplicação disponibiliza.

Capítulo 4

Funcionalidades da aplicação

Neste capítulo são apresentadas todas as funções que a aplicação Web permite a um utilizador do tipo "administrador". Dado que a versão da aplicação para utilizadores do tipo "gestor" é uma versão limitada da versão de administração, não é aqui documentada. Contudo, são descritas as limitações para um utilizador do tipo "gestor".

Um guia detalhado de instalação da aplicação e de todos os módulos que a sustentam é apresentado no anexo A.

4.1 Controlo de acessos

A figura 4.1 apresenta a página que permite o acesso à interface Web de gestão distribuída da ferramenta MORE. Após o *login* bem sucedido, a aplicação encaminha para a área *home*, descrita na Secção 4.2, local que pode encaminhar para todos os serviços disponibilizados pela aplicação.

4.2 Menu principal

A figura 4.2 apresenta a página que pode encaminhar para todos os serviços disponibilizados pela aplicação.

A barra de tarefas, uma vez que está fixada no topo da página, permite acesso a qualquer função disponibilizada pela aplicação, em qualquer instante. O botão:

- "home" - encaminha para o menu principal;
- "users" - encaminha para a página de gestão de utilizadores. (ver a Subsecção 4.2.1);

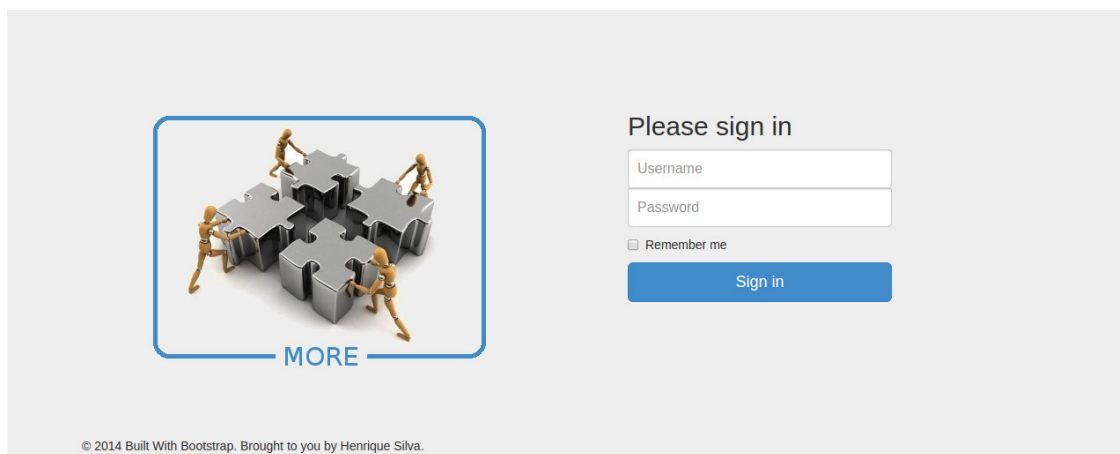


Figura 4.1: Página que permite efetuar login na aplicação

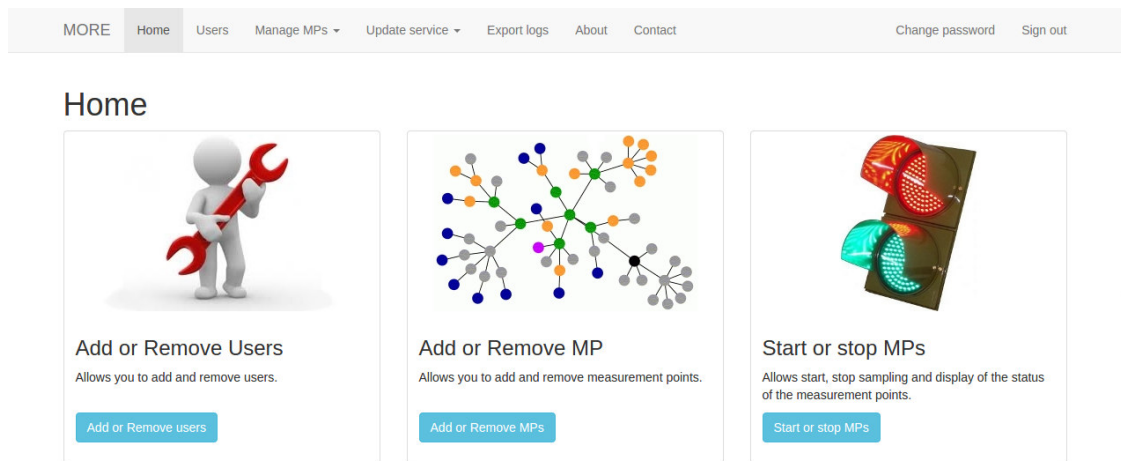


Figura 4.2: Página que permite acesso a todas as funções da aplicação

- "Manage MPs" - abre um *dropdown*, que permite ao utilizar:
 - adicionar ou remover pontos de medição (ver a Subsecção 4.2.2);
 - iniciar ou parar o processo de amostragem e visualizar o estado, ou seja, ver se o ponto de medição se encontra a amostrar ou se está parado (ver a Subsecção 4.2.3).
- "Update service" - abre um *dropdown*, que permite ao utilizar iniciar, parar, reiniciar ou reconfigurar o *Collector* (ver a Subsecção 4.2.4);
- "Export logs" - exporta o registo de todas as operações que foram efetuadas na aplicação (ver a Subsecção 4.2.5);
- "Change password" - permite ao utilizador mudar a palavra-passe;
- "Sign out" - termina a sessão do utilizador.

Ainda de acordo com a figura 4.2, abaixo da barra de tarefas estão disponíveis três módulos distintos. Cada um deles apresenta-se com uma figura ilustrativa, seguido de uma breve descrição e um botão que permite encaminhar para a página de adicionar ou remover utilizadores, adicionar ou remover pontos de medição, e iniciar ou parar amostragem e visualização do estado, respetivamente.

4.2.1 Adicionar ou remover utilizador

Para adicionar ou remover utilizadores, é necessário premir o botão "Add or remove users", no menu principal. Posteriormente é apresentada a página ilustrada na figura 4.3. O primeiro bloco permite adicionar novos utilizadores. Um utilizador é composto por um nome, um *username*, uma *password* e um tipo de utilizador. O nome de *username* tem obrigatoriamente de ser único. Caso não seja único, será enviado um alerta e o utilizador não é inserido. O tipo de utilizador serve para definir o nível de privilégios do utilizador:

- Administrador - tem privilégios para executar qualquer tipo de tarefas, sobre qualquer ponto de medição;
- Manager - um utilizador deste tipo:
 - Pode adicionar e remover pontos de medição do seu utilizador. A adição e remoção de pontos de medição de um utilizador deste tipo, carece de uma validação de utilizador do tipo administrador, ou seja, o novo ponto de medição só estará disponível para entrar em funcionamento depois da validação de um administrador.
 - Pode iniciar ou parar o processo de amostragem dos pontos de medição que é responsável.

The screenshot shows a web interface for user management. At the top is a navigation menu with items: MORE, Home, Users, Manage MPs, Update service, Export logs, About, Contact, Change password, and Sign out. Below the menu is the title 'Add or Remove user'. The interface is split into two main sections. The first section, 'Insert user', is a form with four input fields: 'Name' (placeholder: Insert Name), 'Username' (placeholder: Insert username), 'Password' (placeholder: Insert password), and 'Type' (placeholder: Select type). A green 'Save user' button is at the bottom of this form. The second section, 'Remove user', is a table with three columns: 'Name', 'Username', and 'Type'. The table contains one row with the name 'Henrique', username 'manager', and type 'manager'. A red 'Remove Users' button is located below the table.

Figura 4.3: *Página que permite inserir ou remover utilizadores*

- Não pode adicionar, nem remover utilizadores;
- Não pode iniciar, parar, reiniciar ou reconfigurar o *Collector*;
- Não pode exportar o registo de *logs*.

No segundo bloco, são apresentados todos os utilizadores que este utilizador pode remover, ou seja, existem dois utilizadores inseridos, o "admin" e o "manager", contudo apenas está listado o "manager" porque se está a visualizar a lista de utilizadores a partir do utilizador "admin". Esta medida tem o objetivo de evitar que um utilizador da aplicação se remova a si próprio.

4.2.2 Adicionar ou remover ponto de medição

Esta área permite ao utilizador inserir ou remover pontos de medição. Os dados necessários para preparar um novo ponto de medição são:

- Nome a atribuir ao ponto de medição;
- Tipo de exportação, ou seja, o formato da mensagem que será exportada pelo ponto de medição;
- Credenciais de acesso ao ponto de medição. Estas credenciais têm de ter privilégios *root*;
- Descrição do ponto de medição. Este campo é facultativo;
- Nome do grupo de trabalho onde o ponto de medição está a atuar;
- Endereço e máscara de rede onde o ponto de medição está a atuar.

Após preencher o formulário e selecionando o botão "Save MP", apresentado na figura 4.4, será introduzida na base de dados a informação do novo ponto de medição e retorna-se à página "home".

Caso se pretenda remover pontos de medição inseridos na base de dados, basta selecionar os pontos de medição a remover e selecionar no botão "Remove MPs", apresentado na figura 4.5.

4.2.3 Acções sobre pontos de medição

Nesta área é permitida a configuração de um conjunto de pontos de medição ou efetuar uma pausa de uma amostragem previamente configurada. É também possível visualizar quais os pontos de medição que não se encontram em funcionamento e os pontos de medição que se encontram a operar. Este formulário tem a seguinte informação disponível:

MORE Home Users **Manage MPs** Update service Export logs About Contact Change password Sign out

Add or Remove measurement point

Insert measurement point

Name

Type

IP

IP Collector

Credentials

Owner

Description

Group Name

Network

Figura 4.4: *Página de inserção e remoção de pontos de medição (bloco de inserção)*

MORE Home Users **Manage MPs** Update service Export logs About Contact Change password Sign out

Remove measurement points

	Name	Type	IP	IP Collector	Network	Username	Owner
<input type="checkbox"/>	S0	ipfix	192.168.10.8	188.82.17.225	192.168.10.0/24	root	admin
<input type="checkbox"/>	S1	ipfix	192.168.10.11	188.82.17.225	192.168.10.0/24	root	antonio
<input type="checkbox"/>	S2	ipfix	192.168.10.13	188.82.17.225	192.168.10.0/24	root	andre
<input type="checkbox"/>	S4	ipfix	10.0.0.12	188.82.17.225	10.0.0.0/8	root	admin
<input type="checkbox"/>	S3	silk	10.10.10.1	188.82.17.225	10.0.0.0/8	root	andre

Figura 4.5: *Página de inserção e remoção de pontos de medição (bloco de remoção)*

- "Name" - apresenta o nome do ponto de medição;
- "Type" - apresenta o formato da mensagem de exportação dos fluxos;
- "IP M.P." - apresenta o IP do ponto de medição;
- "IP Collector" - apresenta o endereço de IP do *collector* que irá receber os fluxos gerados pelo ponto de medição;
- "Network" - apresenta a rede onde o ponto de medição se encontra;
- "Status" - apresenta o estado em que se encontra o ponto de medição. Se o servidor Web for a mesma máquina que o *collector*:
 - caso não esteja a receber dados, aparece "stopped" a vermelho;
 - no caso de estar a receber dados, aparece uma mensagem a verde que pode ser diversificada. Estas diferenças têm a ver com os argumentos necessários ao funcionamento de cada técnica. No exemplo apresentado na figura 4.6, o ponto de medição está a amostrar dados na interface com nome lógico "1", capturando 5 em cada 100 pacotes e a operar com a técnica sistemática *count-based*.

Se o servidor Web não for o mesmo que o *collector*, aparecerá "no information" a cor-de-laranja.

Configurar técnica de amostragem

Para iniciar amostragem de um ou vários pontos de medição, é necessário selecionar o botão "start or stop" no menu principal (ver figura 4.2). Será apresentado o menu da figura 4.6, onde se deve selecionar qual ou quais os pontos de medição que vão iniciar amostragem. Ao pressionar o botão "Configure MPs", será apresentado o menu de escolha da técnica de amostragem pretendida, como apresentado na figura 4.7. Dependendo da técnica de amostragem, será apresentado um menu que coleta a informação necessária para iniciar amostragem nos pontos de medição pretendidos, com a técnica escolhida. A figura 4.8 apresenta um caso em que foi selecionada a técnica de amostragem sistemática *count-based*. Neste caso, por se tratar da técnica de amostragem *count-based* apenas necessita de conhecer o número de amostras e o intervalo entre amostras. Os restantes campos apresentados na figura 4.8 são obrigatórios para todas as técnicas.

Caso o campo "Exporting to collector" seja definido como "yes" o ponto de medição começa a exportar os dados para o *Collector* que se encontra definido no Servidor Web. Se for definido como "no", o ponto de medição inicia o processo de amostragem, mas armazena as capturas na diretoria local "/home/results/".

O "interface number" é o número da interface que se pretende analisar.

Parar processo de amostragem

De modo a parar os processos de amostragem de um ou vários pontos de medição, é necessário pressionar o botão "Stop or start MPs" no menu principal (ver figura 4.2). Será apresentado o menu da figura 4.6, onde se deve selecionar quais os pontos de medição que serão alterados e pressionar o botão "Stop MPs".

4.2.4 Reconfigurar serviços

O botão "update service" permite ao utilizador atualizar todas as informações dos pontos de medição no *Collector*, ou simplesmente iniciar, parar e reiniciar o *Collector*.

De cada vez que é inserido um ponto de medição novo, deve-se reconfigurar o serviço. Para reconfigurar o serviço basta premir o botão "update service", onde será lançado um *dropdown*, seguido da opção "reload service". Ao premir o botão "reload service", todas as informações dos pontos de medição que se encontram na base de dados, serão atualizadas e o *Collector* será reiniciado. Caso

More Home Users **Manage MPs** Update service Export logs About Contact Change password Sign out

Select measurement point

Select measurement points

Name	Type	IP M. P.	IP Collector	Network	Owner	Status
<input type="checkbox"/> S0	ipfix	192.168.10.8	188.82.17.225	192.168.10.0/24	admin	Exporting to Collector (Packet-level systematic count-based; interface:1; size:5; interval:100;)
<input type="checkbox"/> S1	ipfix	192.168.10.11	188.82.17.225	192.168.10.0/24	antonio	Exporting to Collector (Packet-level systematic count-based; interface:1; size:5; interval:100;)
<input type="checkbox"/> S2	ipfix	192.168.10.13	188.82.17.225	192.168.10.0/24	andre	Stopped
<input type="checkbox"/> S4	ipfix	10.0.0.12	188.82.17.225	10.0.0.0/8	admin	Error Unable to establish connection
<input type="checkbox"/> S3	silk	10.10.10.1	188.82.17.225	10.0.0.0/8	andre	Without information

[Configure MPs](#) [Stop MPs](#)

Figura 4.6: Página de seleção de pontos de medição que serão alterados

More Home Users **Manage MPs** Update service Export logs About Contact Change password Sign out

Manager settings

Select technique

Packet-level	Flow-level
<input checked="" type="radio"/> Systematic count-based	<input type="radio"/> Systematic count-based
<input type="radio"/> Systematic time-based	<input type="radio"/> Systematic time-based
<input type="radio"/> Systematic event-based	<input type="radio"/> Systematic event-based
<input type="radio"/> Random count-based	<input type="radio"/> Random count-based
<input type="radio"/> Multiadaptive sampling	<input type="radio"/> Multiadaptive sampling
<input type="radio"/> Adaptive linear prediction sampling	<input type="radio"/> Adaptive linear prediction sampling

[Next](#)

Figura 4.7: Página de seleção da técnica a aplicar sobre vários pontos de medição

More Home Users **Manage MPs** Update service Export logs About Contact Change password Sign out

Manager settings

Packet-level systematic count-based

Exporting to Collector Yes No

Interface number

Insert sample size

Insert interval between samples

[Apply](#)

Figura 4.8: Página de inserção de argumentos mediante a técnica selecionada

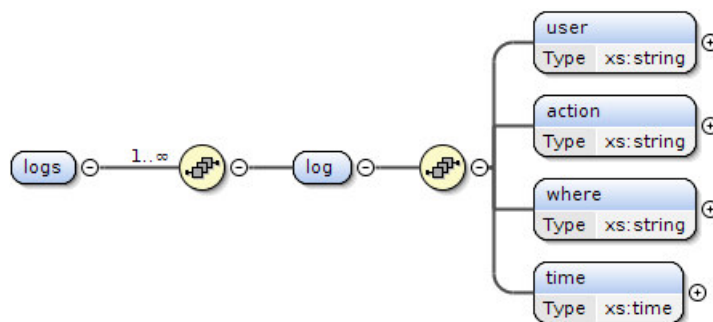


Figura 4.9: Esquema de exportação de logs

esta operação não seja efetuada, o *Collector* pode não receber a informação que o novo ponto de medição está a capturar.

As operações iniciar, parar e reiniciar, permitem iniciar, parar ou reiniciar, respetivamente.

Todas as operações apenas se encontram disponível para um utilizador do tipo "administrador". Um utilizador do tipo "gestor" apenas pode notificar o administrador. Caso este valide a notificação, a operação será realizada.

4.2.5 Registo de logs

À medida que o utilizador vai executando operações na aplicação, esta cria um registo de *logs* que tem como objetivo facilitar a deteção de anomalias. Esta informação é disponibilizada em formato XML apenas a utilizadores do tipo administrador, de acordo com o esquema apresentado na figura 4.9. Em conformidade com este esquema, "logs" é uma sequência de "log", em que um "log" é composto por:

- "user" - guarda o identificador do utilizador que efetuou a operação;
- "action" - identifica a operação que se efetuou;
- "where" - identifica o nome do último *script* executado quando a operação foi feita;
- "time" - guarda a data no formato "Y-m-d H:i:s".

Esta operação não se encontra disponível para um "gestor".

4.3 Sumário

Neste capítulo foram descritas as funcionalidades da aplicação desenvolvidas. Aqui são apresentadas todas as possíveis ações sobre os pontos de medição, gestão de utilizadores e controlo de acessos.

Foram descritas e ilustradas a página de controlo de acessos, bem como as várias tarefas possíveis, nomeadamente, adicionar e remover pontos de medição, iniciar, visualizar e parar as configurações dos pontos de medição, adicionar e remover utilizadores à aplicação, atualizar as configurações do *Collector* e exportação do registo de *logs*.

No próximo capítulo são apresentados os testes de funcionalidades realizados e resultados obtidos.

Capítulo 5

Testes e resultados

Neste capítulo são apresentados testes efetuados e os resultados obtidos que confirmam o bom funcionamento da ferramenta desenvolvida.

Todo o processo de validação e teste da aplicação foi contínuo, acompanhando a implementação dos módulos de suporte. Desta forma, foi possível verificar e validar as funcionalidades, de forma a avançar para o desenvolvimento de outros módulos. Foi feita uma validação final, onde foram testadas todas as funcionalidades descritas no capítulo anterior.

5.1 Gestão de utilizadores

Depois de preencher o bloco "Insert user" e premir o botão "Save user", apresentados na figura 4.3, verificou-se que foi criada uma linha na base de dados com os respetivos dados inseridos no formulário e no bloco "Remove user" passou-se a apresentar as informações do novo utilizador.

O bloco "Remove user", ilustrado pela figura 4.3, apresenta todos os utilizadores que o utilizador que tem sessão iniciada pode remover. Tal como esperado, como o utilizador que tem sessão iniciada é o utilizador "admin", este não é apresentado na lista de possíveis utilizadores a remover. Esta medida tem o objetivo evitar que um utilizador se remova a si próprio. Ao selecionar vários utilizadores e ao premir o botão "Remove users", verificou-se que todos os utilizadores selecionados e apenas estes foram removidos da aplicação.

Ao iniciar a sessão com um utilizador com privilégios de gestor, verificou-se que esta opção não está disponível.

5.2 Gestão de pontos de medição

Nesta secção são apresentados alguns dos testes efetuados às páginas de gestão de pontos de medição.

5.2.1 Adicionar ou remover pontos de medição

Após inserir os dados do ponto de medição no formulário apresentado na figura 4.4 e premir o botão "Save MP", verificou-se que a respetiva informações foi inserida na base de dados. Verificou-se também que quando um administrador inseriu um ponto de medição, este foi alertado para a necessidade de efetuar *reload* do *Collector*. No caso de o utilizador ser um gestor, o utilizador e os administradores recebiam um alerta a informar a necessidade efetuar *reload* do *Collector*. Contudo, apenas os administradores têm privilégios para efetuar qualquer tarefa sobre o *Collector*, o que fazia com que o gestor apenas pudesse usar o novo ponto de medição depois de o administrador validar a notificação. Para além das diferenças descritas, verificou-se também que o administrador pode atribuir pontos de medição, enquanto o gestor apenas pode inserir os seus próprios pontos de medição.

Name	Type	IP M. P.	IP Collector	Network	Owner	Status
<input type="checkbox"/> S0	ipfix	192.168.10.8	188.82.17.225	192.168.10.0/24	admin	Exporting to Collector (Packet-level systematic count-based; interface:1; size:5; interval:100;)
<input type="checkbox"/> S1	ipfix	192.168.10.11	188.82.17.225	192.168.10.0/24	antonio	Exporting to Collector (Packet-level systematic count-based; interface:1; size:5; interval:100;)
<input type="checkbox"/> S2	ipfix	192.168.10.13	188.82.17.225	192.168.10.0/24	andre	Stopped
<input type="checkbox"/> S4	ipfix	10.0.0.12	188.82.17.225	10.0.0.0/8	admin	Error Unable to establish connection
<input type="checkbox"/> S3	silk	10.10.10.1	188.82.17.225	10.0.0.0/8	andre	Without information

Figura 5.1: Página de iniciar, parar e visualizar o estado de pontos de medição do utilizador "admin"

Name	Type	IP M. P.	IP Collector	Network	Owner	Status
<input type="checkbox"/> S1	ipfix	192.168.10.11	188.82.17.225	192.168.10.0/24	antonio	Exporting to Collector (Packet-level systematic count-based; interface:1; size:5; interval:100;)

Figura 5.2: Página de iniciar, parar e visualizar o estado de pontos de medição do utilizador "antonio"

O módulo de remoção ilustrado na figura 4.5 permite a visualização e remoção de pontos de medição. Tal como esperado, depois de selecionar os pontos de medição que se pretende remover e premir o botão "Remover MPs", o administrador conseguiu remover qualquer ponto de medição, de qualquer utilizador. O gestor uma vez que apenas consegue gerir os seus pontos de medição, apenas conseguiu remover os seus próprios pontos de medição.

5.2.2 Acções sobre pontos de medição

Depois de iniciada a amostragem no ponto de medição com IP "192.168.10.13", seleccionou-se a máquina "S2", premiu-se o botão "Stop MPs" e o campo de estado passou a "Stopped" de acordo com a figura 5.1.

Ao iniciar amostragem nas máquinas com os IPs "192.168.10.8", "192.168.10.11" e "10.0.0.12", com a técnica de amostragem sistemática *count-based*, com 5 amostras num intervalo de 100, da forma descrita na Secção 4.2.3, obteve-se os resultados apresentados nas figuras 5.1, 5.2, 5.3 e 5.4. Os pontos de medição seleccionados têm donos diferentes, pelo que apenas foi possível iniciar amostragem em todos eles por ter iniciado a sessão com um utilizador do tipo administrador, neste caso o utilizador "admin". Tal como esperado, uma vez que o utilizador "antonio" apenas tem inserido um ponto de medição e é um utilizador do tipo gestor, na figura 5.2 é apresentado apenas um ponto de medição, o ponto de medição com IP "192.168.10.11". Posteriormente, de acordo com as figuras 5.3 e 5.4, analisou-se a lista de processos JAVA nas distintas máquinas, o que levou a concluir o sucesso na execução do comando para as máquinas com IP "192.168.10.8" e "192.168.10.11". Naturalmente, uma vez que a máquina com IP "10.0.0.12" não existia, em conformidade com a figura 5.2, não foi possível estabelecer ligação comunicação com a máquina.

```

root@framework:~
root@framework:~# ps -ef | grep 'java -jar /home/SamplingFramework.jar'
root    3358  3201  1 19:33 pts/1    00:00:00 java -jar /home/SamplingFramework.jar 1 188.82.17.225 1 5 100 true
root    3381  2100  0 19:33 pts/0    00:00:00 grep --color=auto java -jar /home/SamplingFramework.jar
root@framework:~#

```

Figura 5.3: Processos JAVA na máquina com nome "S0"

```

root@webServer:~
root@webServer:~# ps -ef | grep 'java -jar /home/SamplingFramework.jar'
root    3457  3400  0 19:35 pts/4    00:00:00 java -jar /home/SamplingFramework.jar 1 188.82.17.225 1 5 100 true
root    3694  3632  0 19:36 pts/5    00:00:00 grep --color=auto java -jar /home/SamplingFramework.jar
root@webServer:~#

```

Figura 5.4: Processos JAVA na máquina com nome "S1"

5.3 Reconfigurar serviços

Após iniciar sessão com um utilizador com privilégios de administrador, ao pressionar o botão "Update service", seguido da opção iniciar, parar ou reiniciar, verificou-se que o *Collector* foi iniciado, parado ou reiniciado, de acordo com a opção escolhida. O administrador também tem a opção *reload*. Esta opção permite ao utilizador atualizar o ficheiro de configuração do *Collector* e reiniciar o serviço. A figura 5.5 apresenta um *print out* do estado do serviço do *Collector* antes e depois de efetuar a operação *reload*. Pode-se concluir que a atualização foi efetuada com sucesso, porque o serviço tem um novo PID, logo foi reiniciado. Além disso, após inserir um novo ponto de medição, o ficheiro de configuração `"/etc/silk/sensor.conf"`, foi atualizado com base nas novas informações da base de dados.

Após iniciar sessão com um utilizador com privilégios de gestor, verificou-se que esta opção não se encontrava disponível.

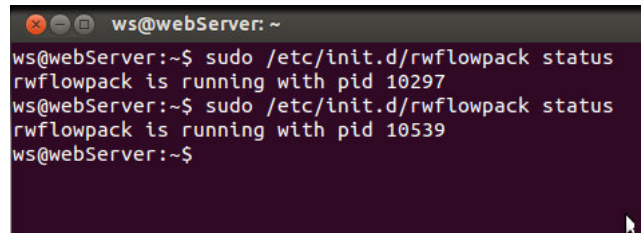
5.4 Registo de logs

Após iniciar a sessão com um utilizador com privilégios de administrador, ao pressionar o botão "Export logs", foi gerado e descarregado um ficheiro em formato XML semelhante ao apresentado na figura 5.6. Após iniciar a sessão com um utilizador com privilégios de gestor, verificou-se que esta opção não se encontrava disponível.

5.5 Sumário

Neste capítulo são apresentados exemplos dos testes efetuados e os resultados obtidos. É possível verificar a inserção, remoção e ou exportação dos dados e a reconfiguração do *Collector* no caso de este se encontrar na mesma máquina que o servidor Web. Os resultados mostram que o método utilizado foi eficaz e as ações descritas acima foram efetuadas com sucesso.

No capítulo seguinte serão apresentadas as conclusões finais e sugestões para trabalho futuro.



```
ws@webServer: ~  
ws@webServer:~$ sudo /etc/init.d/rwflowpack status  
rwflowpack is running with pid 10297  
ws@webServer:~$ sudo /etc/init.d/rwflowpack status  
rwflowpack is running with pid 10539  
ws@webServer:~$
```

Figura 5.5: Estado do processo do Collector antes e depois de atualizar o serviço na aplicação

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
--<logs>  
--<log>  
  <user>admin</user>  
  <action>insert</action>  
  <where>mp</where>  
  <time>2014-10-09 18:00:00</time>  
</log>  
--<log>  
  <user>admin</user>  
  <action>delete</action>  
  <where>mp</where>  
  <time>2014-10-09 18:00:00</time>  
</log>  
</logs>
```

Figura 5.6: Conteúdo existente no ficheiro de exportação de logs

Capítulo 6

Conclusões e trabalho futuro

Desde há muito que existe uma preocupação em monitorizar as redes. Com esta necessidade muitas técnicas de amostragem foram surgindo, procurando a melhor acurácia possível na estimação de parâmetros de interesse ou num conjunto reduzido deles. Contudo, a heterogeneidade da natureza do tráfego e requisitos de monitorização dos vários serviços existentes nestas redes exige uma solução flexível de uso das técnicas de amostragem, procurando aproveitar as suas melhores características de acordo com o parâmetro em análise.

Com esta necessidade foi desenvolvida a *framework* de amostragem MORE baseada numa taxonomia de técnicas de amostragem, devidamente estruturada, a partir da qual é possível desenhar as componentes existentes nas técnicas de amostragem tradicionais e futuras. No entanto, a *framework* MORE nos moldes em que se encontrava tinha grandes dificuldades de exportação da informação seleccionada, pois esta mediante a técnica de amostragem, colecionava a informação numa diretoria local em formato PCAP. Mesmo estando perante dados amostrados, uma vez que este formato tem todo o conteúdo do pacote, o volume de dados pode ser muito alto, sobretudo quando se está a falar de redes de alto débito. Assim, havia duas soluções: ou se acedia ao ponto de medição remotamente e se exportava toda a informação para um local que possibilitasse a análise de resultados ou o administrador tinha de se deslocar a cada ponto de medição para fazer uma cópia dos dados capturados. Além disso, a configuração da *framework* de amostragem MORE teria de ser via *terminal*, o que fazia desta abordagem, uma abordagem complexa sobretudo para novos utilizadores e/ou dispendiosa.

Perante esta situação desenvolveu-se uma solução escalável para configuração distribuída da *framework* de amostragem MORE e aquisição de informações de monitorização fornecidas por ela. Assim, desenvolveu-se uma solução baseada em ferramentas existentes, normalizadas, e que facilita a integração com ferramentas externas, desde que estas respeitem os protocolos padrão usados no desenvolvimento desta solução, o que faz desta solução, uma solução muito versátil. Como se pretendia desenvolver uma solução escalável, houve necessidade de efetuar uma cuidada análise à arquitetura do sistema a adotar, de modo a proporcionar uma otimização do volume de dados envolvido, o que encaminhou para uma abordagem de coletas orientadas ao fluxo, em vez de coletas orientadas ao pacote. A abordagem orientada ao fluxo, fez reduzir substancialmente o volume de dados gerados. Esta redução de volume de dados leva a concluir que este é o caminho para conseguir uma solução escalável para a configuração distribuída da *framework* de amostragem MORE, com gestão centralizada.

Uma vez que esta aplicação é uma aplicação baseada em Web, permite efetuar a gestão dos pontos de medição a partir de qualquer local, a partir de um computador ou telefone inteligente, desde que estes disponibilizem um *browser* e acesso à Web.

De acordo com o descrito, considera-se que todos os objetivos foram atingidos com sucesso, com exceção da representação gráfica dos resultados rececionados dos pontos de medição, o que se remete para trabalho futuro, com a instalação do FlowViewer, conforme indicado na Secção 6.1

6.1 Sugestões para Pesquisas Futuras

Este trabalho teve como finalidade desenvolver uma interface de gestão distribuída da ferramenta MORE, ou seja, foi criada uma aplicação Web para gerir um conjunto de pontos de medição distribuídos na rede e coletar os dados colecionados pelos pontos de medição.

De forma a melhorar o trabalho realizado propõe-se como trabalho futuro:

- aperfeiçoar a experiência de uso da aplicação Web, incluindo novos módulos em Javascript, de forma a tornar a aplicação mais intuitiva;
- incluir a ferramenta FlowViewer para visualização gráfica dos resultados. O FlowViewer é uma ferramenta *open source*, desenvolvida pela NASA e pela ESDIS e usa como base o SiLK. Esta ferramenta permite *tracking* e visualização gráfica de resultados baseada em Web [NASA e ESDIS (2012)];
- encriptar a base de dados, melhorando assim a segurança dos dados inseridos e fazendo com que os dados inseridos não sejam perceptíveis para utilizadores não autorizados;
- criar um método de autenticação do ponto de medição automatizado, evitando assim a prévia inserção de cada ponto de medição na aplicação Web. As dificuldades de colocar em funcionamento esta solução prendem-se ao facto de após se efetuarem todas as configurações necessárias para colocar em funcionamento o novo ponto de medição, é necessário reiniciar o *Collector*, o que pode provocar a perda de alguma informação dos pontos de medição que se encontram em funcionamento. Contudo, a perda de informação em cada ponto de medição pode não ser relevante sobretudo se estivermos a falar de um ponto de medição que esteja incluído numa rede de alto débito. Ou seja, como a informação é exportada de 15 em 15 segundos e o processo de reiniciar o *Collector* tem um tempo muito inferior, quer dizer que no máximo se pode perder 15 segundos dos fluxos gerados por cada ponto de medição, o que constitui um volume de dados muito baixo quando comparado com o total capturado. A solução para este problema pode passar por: após o ponto de medição, o *Collector* e o servidor Web trocarem todas as informações necessárias para o bom funcionamento do processo de amostragem e exportação de fluxos, enviar uma notificação para o administrador para que este valide a inserção do novo ponto de medição. Caso o administrador valide, a informação seria atualizada no *Collector* e no servidor Web, e o ponto de medição estaria pronto a colecionar e exportar dados. Caso contrário, se o administrador não validar, esta informação passaria para quarentena e o administrador não voltaria a ser notificado sobre este ponto de medição, até que os dados fossem eliminados da quarentena;
- de acordo com o explicado na Subsecção 2.4.3, este produto usa o SiLK como *Collector*. Este *Collector* permite usar os pontos de medição em várias abordagens distintas. Neste momento, o servidor apenas contempla a abordagem em que é criada uma *probe* separado para cada *sensor*, usando a mesma porta, variando o campo do *host* a aceitar como forma de distinção. Neste momento a base de dados está estruturada de forma a contemplar as diferentes abordagens permitidas pelo SiLK, no entanto, é necessário criar formulários que sustentem a inserção de dados para as diferentes abordagens e adaptar o módulo que gera o ficheiro de configuração do SiLK, mais concretamente o módulo (`update.php`) que gera o ficheiro `"/etc/silk/sensor.conf"`;
- atualmente a aplicação encontra-se em funcionamento através do protocolo de comunicação HTTP. De forma a garantir confidencialidade da informação trocada nos dados que são comunicados entre o servidor Web e o utilizador da aplicação, seria interessante adicionar ao servidor atual capacidades de segurança SSL/TLS alterando o serviço para HTTPS;
- os sistemas Unix necessitam de privilégios *root* para aceder a informações da placa de rede. Como tal, para que a *framework* conseguisse iniciar o processo de amostragem com base no tráfego que passa em determinada porta, foi necessário inserir na aplicação Web uma password com privilégios *root*, o que deveria ser alterado.

Apêndice A

Guia de instalação

Neste capítulo pretende-se divulgar todos os passos necessários à instalação desde produto, tornando possível usar, corrigir possíveis *bugs* e continuar o desenvolvimento novas funcionalidades.

O guia de instalação encontra-se dividido em duas secções. A Secção A.1 apresenta detalhadamente todos os passos para a configuração do Servidor Web e a Secção A.2 apresenta detalhadamente todos os passos para configurar um ponto de medição. É de referir que o sistema operativo usado para preparar o Servidor Web e o ponto de medição foi o Ubuntu 12.04 a 64 bits. Contudo, é esperado que com pequenos ajustes não haja grande dificuldade de instalar estas ferramentas em qualquer sistema operativo baseado em Linux.

A.1 Servidor Web

O Servidor Web disponibiliza uma *interface Web* atraente e de manuseamento simples. Esta máquina permite também desencadear ações sobre um ponto de medição ou um grupo de pontos de medição, desde que estes se encontrem devidamente registados no Servidor Web. Após se configurar uma ação, o *daemon* que se encontra à "escuta" no Servidor Web ficar a aguardar que o ponto de medição ou os pontos de medição devolvam os resultados. Estes resultados são o subconjunto de tráfego coletado pela *framework* que se encontra em funcionamento em cada ponto de medição, mediante a configuração inicialmente definida a partir do Servidor Web, como já foi dito anteriormente.

Com o comando apresentado no *Listing A.1* é possível instalar o *lamp server*. O *lamp server* consiste num conjunto de ferramentas *open source* agrupadas que permitem criar um Servidor Web. O *lamp server* é constituído pelas seguintes ferramentas:

- *Apache Server*;
- Base de dados *MySQL*;
- Compilador *PHP*.

```
1 $ sudo apt-get install lamp-server^
```

Listing A.1: *Instalação do lamp server*

Para poder usar o SSH2, é necessário instalar algumas dependências. O SSH2 é um método do PHP que permite executar ações por *secure shell* (ver [Achour *et al.* (1997)]).

Para instalar o SSH2 usaram-se os comandos do *Listing A.2* que se passam a explicar:

1. O "libssh-dev", pois é uma dependência do "libssh";
2. Criou-se uma pasta de nome "libssh";

3. Alterou-se o *owner* da directoria "libssh" para "USER";
4. Abriu-se a directoria de nome "libssh";
5. Fez-se *download* do "libssh2-1.4.3";
6. Descompactou-se o ficheiro comprimido de nome "libssh2-1.4.3.tar.gz";
7. Abriu-se a directoria que é gerada ao descompactar o respetivo ficheiro;
8. Efetuou-se o *build* da ferramenta "libssh2";
9. Com o comando "sudo make install", instalou-se o "libssh";

```

1 $ sudo apt-get install libssl-dev
2 $ sudo mkdir /usr/local/src/libssh
3 $ sudo chown USER.USER /usr/local/src/libssh
4 $ sudo cd /usr/local/src/libssh
5 $ wget http://libssh2.org/download/libssh2-1.4.3.tar.gz
6 $ tar vxzf libssh2-1.4.3.tar.gz
7 $ cd libssh2-1.4.3
8 $ ./configure && make
9 $ sudo make install

```

Listing A.2: *Instalação do libssh*

O "php-ssh2" também necessita de uma dependência, o "php5-dev", e é possível instalá-lo, com os comandos descritos no *Listing A.3*.

É extremamente importante adicionar o *path* para ssh2.so. Para tal, basta acrescentar a linha apresentada no *Listing A.4*, ao ficheiro "php.ini".

```

1 $ sudo apt-get install php5-dev
2 $ sudo vi /etc/php5/apache2/php.ini

```

Listing A.3: *Comandos para instalar o módulo "php5-dev"*

```

1 extension=/usr/lib/php5/20090626/ssh2.so

```

Listing A.4: *Linha a acrescentar no ficheiro "php.ini"*

Agora está tudo pronto para instalar o SSH2. Passa-se a explicar o que realizam as instruções mencionadas no *Listing A.5*:

1. Criou-se uma pasta de nome "phpssh";
2. Alterou-se o *owner* da directoria "phpssh" para "USER";
3. Abriu-se a directoria de nome "phpssh";
4. Fez-se *download* do "ssh2-0.12";
5. Descompactou-se o ficheiro comprimido de nome "ssh2-0.12.tgz";
6. Abriu-se a directoria que é gerada ao descompactar o respetivo ficheiro;
7. Antes de efetuar o *build* da ferramenta "ssh2", é necessário fazer *run* do *phpize*.

O *phpize* é um comando usado para preparar o *build environment* para uma extensão PHP [Achour *et al.* (1997)];

8. Efetuou-se o *build* da ferramenta "ssh2-0.12";
9. Com o comando "sudo make install", instalou-se o "ssh2";

```

1 $ sudo mkdir /usr/local/src/phpssh
2 $ sudo chown USER.USER /usr/local/src/phpssh
3 $ cd /usr/local/src/phpssh
4 $ wget http://pecl.php.net/get/ssh2-0.12.tgz
5 $ cd ssh2-0.12
6 $ phpize
7 $ ./configure --with-ssh2 && make
8 $ sudo make install

```

Listing A.5: *Comandos para instalar o SSH2*

Após o processo de instalação destas ferramentas é necessário reiniciar o *apache* com o comando apresentado no *Listing A.6*.

```

1 $ sudo /etc/init.d/apache2 restart

```

Listing A.6: *Comando para reiniciar o apache*

Depois de reiniciar o *apache*, é necessários colocar o código fonte na diretoria raiz do *apache*. Para tal, é necessário executar os comandos seguintes:

1. Criou-se uma pasta de nome "app";
2. Alterou-se o *owner* da diretoria "app" para "USER";
3. Abriu-se a diretoria de nome "app";
4. Fez-se *download* do código fonte;
5. Descompactou-se o ficheiro comprimido de nome "more.tar.gz";
6. Abriu-se a diretoria que é gerada ao descompactar o respetivo ficheiro;
7. Copia-se a diretoria gerada ao descompactar o ficheiro comprimido de nome "more.tar.gz" para a diretoria "/var/www/";
8. Alterou-se o *owner* da diretoria "more" para "USER";
9. Alterou-se as permissões da diretoria copiada, para que os utilizadores fora do grupo apenas possam executar o código fonte.

```

1 $ sudo mkdir /usr/local/src/app
2 $ sudo chown USER.USER /usr/local/src/app
3 $ cd /usr/local/src/app
4 $ wget https://dl.dropboxusercontent.com/u/49961504/application/more.tar.gz
5 $ tar -xvzf more.tar.gz
6 $ cp /usr/local/src/app/more /var/www/
7 $ sudo chown USER.USER /var/www/more/
8 $ sudo chmod -R 755 /var/www/more/

```

Listing A.7: *Comandos para copiar o código fonte para o apache*

Para se ligar à base de dados é possível fazê-lo de acordo com o apresentado no *Listing A.8*. Esta base de dados foi criada para gestão de acessos à aplicação, controlo de *logs*, gestão de pontos de medição e visualização de resultados.

```
1 $ mysql --user=USER --password=PASSWORD
```

Listing A.8: *Ligação à base de dados*

Depois de ligado à base de dados, é necessário criar uma base de dados com o nome "more". É mesmo importante que o nome seja "more", pois caso contrário, a aplicação *Web* não funcionará corretamente. Pode-se criar a base de dados com os comandos apresentados no *Listing A.9*.

```
1 mysql> CREATE DATABASE more;
2 mysql> quit;
```

Listing A.9: *Comandos SQL para criar a base de dados e sair*

É necessário importar a base de dados que se encontra em "/usr/local/src/app/more/more.sql". Agora, fora do SQL, basta executar o comando apresentado no *Listing A.10*.

Nota: O facto de o "-p" estar junto da "PASSWORD", não é engano.

```
1 $ mysql -u USER -pPASSWORD more < /usr/local/src/app/more/more.sql
```

Listing A.10: *Comando para importar a base de dados*

Para aceder à aplicação basta colocar a *url* <http://localhost/more/docs/examples/signin> no *browser* e serão pedidas as credenciais. Por definição, inicialmente o *username* e a *password* são "admin", as quais devem ser alteradas após o primeiro *login*.

Com o comando apresentado no *Listing A.11* é possível instalar o "openssh-client". Esta ferramenta permite ao Servidor Web ligar-se ao ponto de medição para efetuar ações sobre o último.

```
1 $ sudo apt-get install openssh-client
```

Listing A.11: *Comando de instalação do cliente SSH*

O "libfixbuf" é uma ferramenta compatível com o RFC 5101 [Claise (2008)] que faz parte do grupo de ferramentas disponíveis pela NetSA da CERT. Os comandos apresentados no *Listing A.12* fazem o seguinte:

1. Cria uma diretoria com o nome "netsa";
2. Altera o *owner* da diretoria "netsa" para "USER";
3. Abre a diretoria "netsa";
4. Faz o *download* da versão 1.5.0 do "libfixbuf".

```
1 $ sudo mkdir /usr/local/src/netsa
2 $ sudo chown USER.USER /usr/local/src/netsa
3 $ cd /usr/local/src/netsa
4 $ wget http://tools.netsa.cert.org/releases/libfixbuf-1.5.0.tar.gz
```

Listing A.12: *Download da ferramenta "libfixbuf"*

Os comandos apresentados no *Listing A.31* fazem o seguinte:

1. descompacta o ficheiro comprimido de nome "libfixbuf-1.5.0.tar.gz";
2. abre a diretoria que é gerada ao descompactar o respetivo ficheiro;
3. efetua o *build* da ferramenta "libfixbuf";
4. faz a instalação do "libfixbuf".

```
1 $ tar zxvf libfixbuf-1.5.0.tar.gz
2 $ cd libfixbuf-1.5.0
3 $ ./configure && make
4 $ sudo make install
```

Listing A.13: Comandos para instalar o libfixbuf

Os comandos apresentados no *Listing A.14* fazem as seguintes operações:

1. Cria uma diretoria com o nome "netsa";
2. Altera o *owner* da diretoria "netsa" para "USER";
3. Abre a diretoria de nome "netsa";
4. Faz *download* da biblioteca netsa-python;
5. Faz *download* da ferramenta SiLK;

```
1 $ sudo mkdir /usr/local/src/netsa
2 $ sudo chown USER.USER /usr/local/src/netsa
3 $ cd /usr/local/src/netsa
4 $ wget http://tools.netsa.cert.org/releases/netsa-python-1.4.3.tar.gz
5 $ wget http://tools.netsa.cert.org/releases/silk-2.5.0.tar.gz
```

Listing A.14: Download das ferramentas netsa-python e SiLK

A ferramenta SiLK necessita da biblioteca "netsa-python" e esta biblioteca necessita do compilador Python, os quais se instalam da forma apresentada no *Listing A.15*.

```
1 $ sudo apt-get install python-dev
2 $ cd /usr/local/src/netsa
3 $ tar zxvf netsa-python-1.4.3.tar.gz
4 $ cd netsa-python-1.4.3
5 $ sudo python setup.py install
```

Listing A.15: Instalação o compilador Python e da biblioteca Python do NetSA

Estão reunidas todas as condições para instalar o SiLK. O *Listing A.16* apresenta todos os procedimentos para a instalação desta ferramenta, os quais se passam a explicar:

1. Abre-se a diretoria de nome "netsa";
2. Descompacta-se o ficheiro comprimido de nome "silk-2.5.0.tar.gz";
3. Efetua-se o *build* da ferramenta "SiLK". Os argumentos inseridos no comando "configure" têm a seguinte função:

- O argumento "-with-python" inclui a biblioteca "netsa-python" anteriormente instalada;
- O argumento "-with-libfixbuf=/usr/local/lib/pkgconfig/" inclui a biblioteca "libfixbuf" e especifica o local onde esta se encontra instalada;
- O argumento "-sysconfdir=/etc/silk" especifica o local onde o SiLK vai buscar as configurações, nomeadamente, o ficheiro de configuração "rwflowpack.conf" mencionados nos *Listing A.22* e o "sensor.conf" gerado automaticamente a partir das configurações inseridas, quando se insere um novo ponto de medição;
- O argumento "-enable-data-rootdir=/var/flows/all_framework" especifica o local onde os fluxos que chegam dos pontos de medição são guardados;

4. Efetua-se a instalação da ferramenta.

```

1  $ cd /usr/local/src/netsa
2  $ tar zxvf silk-2.5.0.tar.gz
3  $ cd silk-2.5.0
4  $ ./configure --with-python --sysconfdir=/etc/silk --with-libfixbuf=/usr/local
    /lib/pkgconfig/ --enable-output-compression --enable-data-rootdir=/var/
    flows/all_framework
5  $ make && sudo make install

```

Listing A.16: Comandos para instalação do SiLK

O comando apresentado no *Listing A.17*, cria a diretoria que guarda os fluxos que chegam dos pontos de medição.

```

1  $ sudo mkdir /var/flows/all_framework

```

Listing A.17: A criar a diretoria onde são armazenados os fluxos que chegam dos pontos de medição

Neste instante todas as ferramentas necessárias ao funcionamento do SiLK se encontram instaladas. De modo a evitar que de cada vez que se inicie o Servidor Web seja necessário atualizar a variável "LD_LIBRARY_PATH", é necessário criar um ficheiro de nome "silk.conf" na diretoria "/etc/ld.so.conf.d/" com os comandos que se encontram descritos no *Listing A.18*. O respetivo ficheiro deve conter a informação apresentada no *Listing A.19*.

```

1  $ cd /etc/ld.so.conf.d/
2  $ sudo vi silk.conf

```

Listing A.18: Configuração da variável "LD_LIBRARY_PATH"

```

1  /usr/local/lib
2  /usr/local/lib/silk

```

Listing A.19: Conteúdo a colocar no ficheiro "/etc/ld.so.conf.d/silk.conf"

É importante fazer *run* do comando inserido no *Listing A.20* para fazer *binding*, para poder usar as bibliotecas "libfixbuf" e o "netsa-python".

```

1  $ sudo ldconfig

```

Listing A.20: Comandos para atualização do path das bibliotecas

No *Listing A.21* são apresentados os comandos necessários para criar o sistema de ficheiros de configuração do SiLK, os quais se explicam:

1. Cria-se a diretoria de nome "silk";
2. Abre-se a respetiva diretoria "silk";
3. Faz-se uma cópia do ficheiro de configuração modelo, de nome "twoway-silk.conf" para a diretoria que se acabou de criar e dá-lhe o nome "silk.conf";
4. Faz-se uma cópia do ficheiro de configuração modelo, de nome "rwflowpack.conf" para a diretoria "/etc/silk"

```

1  $ sudo mkdir /etc/silk
2  $ cd /etc/silk
3  $ sudo cp /usr/local/share/silk/twoway-silk.conf ./silk.conf
4  $ sudo cp /usr/local/share/silk/etc/rwflowpack.conf /etc/silk

```

Listing A.21: Comandos para criar o sistema de ficheiros de configuração do SiLK

O módulo *rwflowpack* é um componente do SiLK responsável por colecionar os fluxos de dados e armazená-los em ficheiros binários, em formato SiLK. O ficheiro 'rwflowpack.conf' é o ficheiro de configuração deste módulo. Após a cópia do ficheiro modelo, deve-se alterar para as configurações apresentadas no *Listing A.22*. Existem muitas outras opções possíveis, no entanto, não são relevantes para colocar este serviço em funcionamento, no contexto desta aplicação.

```

1  ENABLED=True
2  statedirectory=/usr/local/var/lib/rwflowpack
3  CREATE_DIRECTORIES=yes
4  SENSOR_CONFIG=/etc/silk/sensor.conf
5  DATA_ROOTDIR=/var/flows/all_framework
6  SITE_CONFIG=/etc/silk/silk.conf
7  LOG_TYPE=syslog
8  LOG_DIR=${statedirectory}/log
9  COMPRESSION_TYPE=best

```

Listing A.22: Ficheiro de configuração "rwflowpack.conf"

Para terminal a configuração do *rwflowpack*, falta apenas efetuar operações apresentadas no *Listing A.23*:

1. Copia-se o ficheiro de arranque modelo do *rwflowpack* para a diretoria "/etc/init.d";
2. Abriu-se o ficheiro de nome "rwflowpack" e alterar a variável "SCRIPT_CONFIG_LOCATION" de "/usr/local/etc/" para "/etc/silk";
3. Inicia-se o serviço.

```

1  $ sudo cp /usr/local/share/silk/etc/init.d/rwflowpack /etc/init.d
2  $ sudo vi /etc/init.d/rwflowpack
3  $ sudo /etc/init.d/rwflowpack start

```

Listing A.23: Configuração e arranque do *rwflowpack*

De forma a evitar que toda a vez que o servidor é iniciado, seja necessário iniciar o serviço *rwflowpack*, deve-se introduzir no *terminal* o comando apresentado no *Listing A.24*.

```
1 $ sudo update-rc.d rwflowpack defaults
```

Listing A.24: *Configuração do arranque on boot do rwflowpack*

Por fim basta executar os comandos apresentados no *Listing A.25* para atualizar os *symbolic links*.

```
1 $ sudo ln -s /etc/silk/silk.conf /var/flows/all_framework
2 $ sudo ln -s /etc/silk/sensor.conf /var/flows/all_framework
```

Listing A.25: *A atualizar os symbolic links*

A.2 Ponto de medição

A Secção A.2 apresenta detalhadamente todos os passos para configurar um ponto de medição. Para configurar um grupo de pontos de medição, é necessário repetir cada um dos passos mencionados na Secção A.2 para cada ponto de medição ou após efetuar a instalação do primeiro ponto de medição, efetuar um *backup* e *restore* nas máquinas que se pretender configurar.

Para a instalação do YAF é necessário instalar alguns pré-requisitos, para que este consiga gerar registos de fluxos IPFIX e encaminhar os mesmos para o SiLK. Para saber mais detalhes sobre YAF, IPFIX e SiLK ver a Secção 2.4. Os pré-requisitos são os seguintes:

- Ferramentas de compilação, por exemplo, gcc e make;
- O Glib para rotinas extra para a linguagem C;
- Bibliotecas de desenvolvimento para o libpcap;
- Bibliotecas de desenvolvimento Perl compatíveis com expressões regulares.

Com o comando apresentado no *Listing A.26* é possível instalar o "openssh-server". Esta ferramenta permite ao Servidor Web efetuar ações sobre qualquer ponto de medição, permitindo assim a interação com a *framework* nele colocada. Este tipo de interação pode ser iniciar, parar ou alterar o processo de amostragem. Também tem a opção de capturar todo o tráfego.

```
1 $ sudo apt-get install openssh-server
```

Listing A.26: *Comando de instalação do servidor SSH*

É essencial executar o comando apresentado no *Listing A.27*. Com este comando é possível definir a *root password*, para acesso via SSH. É necessário ter em atenção que **a password aqui definida tem obrigatoriamente de ser a mesma que se inserir na interface gráfica do Servidor Web, no menu inserir ponto de medição.**

```
1 $ sudo passwd
```

Listing A.27: *Comando necessário para configurar root password*

Com o comando apresentado no *Listing A.28* é possível instalar os compiladores necessários.

```
1 $ sudo apt-get install make gcc g++
```

Listing A.28: *Comando de instalação dos compiladores*

O comando apresentado no *Listing A.29*, permite instalar 16 pacotes dos *build-essential*, nomeadamente, Glib.

```
1 $ sudo apt-get install libglib2.0-dev libpcap-dev libpcre3-dev glib2.0 python-  
dev
```

Listing A.29: *Comando de instalação dos build-essential*

O "libfixbuf" é uma ferramenta compatível com o RFC 5101 [Claise (2008)] que faz parte do grupo de ferramentas disponíveis pela NetSA da CERT. Os comandos apresentados no *Listing A.30* fazem o seguinte:

1. cria-se uma diretoria com o nome "netsa";
2. altera-se o *owner* da diretoria "netsa" para "USER";
3. abre-se a diretoria "netsa";
4. faz-se o *download* da versão 1.5.0 do "libfixbuf".

```
1 $ sudo mkdir /usr/local/src/netsa  
2 $ sudo chown USER.USER /usr/local/src/netsa  
3 $ cd /usr/local/src/netsa  
4 $ wget http://tools.netsa.cert.org/releases/libfixbuf-1.5.0.tar.gz
```

Listing A.30: *Download da ferramenta "libfixbuf"*

Os comandos apresentados no *Listing A.31* fazem o seguinte:

1. descompacta-se o ficheiro comprimido de nome "libfixbuf-1.5.0.tar.gz";
2. abre-se a diretoria que é gerada ao descompactar o respetivo ficheiro;
3. efetua-se o *build* da ferramenta "libfixbuf";
4. faz-se a instalação do "libfixbuf".

```
1 $ tar zxvf libfixbuf-1.5.0.tar.gz  
2 $ cd libfixbuf-1.5.0  
3 $ ./configure && make  
4 $ sudo make install
```

Listing A.31: *Comandos para instalar o libfixbuf*

Os comandos apresentados no *Listing A.30* à semelhança do que já foi dito anteriormente para o "libfixbuf", abrem a pasta "netsa" e descarregam o ficheiro "yaf-2.5.0.tar.gz" para a respetiva diretoria.

```
1 $ cd /usr/local/src/netsa
2 $ wget http://tools.netsa.cert.org/releases/yaf-2.5.0.tar.gz
```

Listing A.32: *Download da ferramenta "YAF"*

De modo a garantir que o *PATH* é especificado sempre que a máquina é reiniciada, é necessário adicionar o *PATH* ao ficheiro ".bashrc". Para isso, basta abrir o ".bashrc" com o comando mencionado no [Listing A.33](#) colocar a instrução mencionada no [Listing A.34](#) e guardar as alterações.

```
1 $ sudo vi ~/.bashrc
```

Listing A.33: *A abrir o ficheiro ".bashrc" como administrador*

```
1 export PKG_CONFIG_PATH=/usr/local/lib/pkgconfig
```

Listing A.34: *Instrução a colocar no ficheiro ".bashrc"*

À semelhança do que foi dito para o "libfixbuf", as instruções apresentadas no [Listing A.35](#), servem para instalar o YAF (para mais detalhes do YAF ver a Subsecção 2.4.2):

1. descompacta-se o ficheiro comprimido de nome "yaf-2.5.0.tar.gz";
2. abre-se a diretoria que é gerada ao descompactar o respetivo ficheiro;
3. efetua-se o *build* da ferramenta "yaf". É de salientar que o argumento "-enable-applabel" instrui o YAF para gravar o protocolo de aplicação utilizado para os fluxos de rede. Por exemplo, o HTTP. YAF também consegue detetar o uso de protocolos de aplicação em portas não *standard*. Esta informação é útil para usar na análise de resultados;
4. Faz-se a instalação do "yaf".

```
1 $ tar zxvf yaf-2.5.0.tar.gz
2 $ cd yaf-2.5.0
3 $ ./configure --enable-applabel && make
4 $ sudo make install
```

Listing A.35: *Comandos para instalar o libfixbuf*

É importante fazer *run* do comando inserido no [Listing A.36](#) para fazer *binding*, para poder usar a biblioteca "libfixbuf".

```
1 $ sudo ldconfig
```

Listing A.36: *Comando para actualização do path das bibliotecas*

Neste momento a ferramenta YAF já se encontra instalada, no entanto, ainda falta instalar o "openjdk-7-jdk" e o "jpcap". Para instalar o "openjdk-7-jdk", basta inserir no *terminal* o comando apresentado no [Listing A.37](#). No caso do "jpcap", é necessário ter em atenção a versão do sistema operativo que se está a usar, ou seja, após descompactar o ficheiro comprimido, deve-se escolher a versão adequada ao sistema operativo. Neste caso, por se tratar de um sistema operativo de 64 bits, instalou a versão "amd". É possível ver a metodologia de instalação no [Listing A.38](#).

```
1 $ sudo apt-get install openjdk-7-jdk
```

Listing A.37: *Comando para instalação do "openjdk-7-jdk"*

```
1 $ sudo mkdir /usr/local/src/jpcap
2 $ sudo chown USER.USER /usr/local/src/jpcap
3 $ cd /usr/local/src/jpcap
4 $ wget https://dl.dropboxusercontent.com/u/49961504/jpcap/jpcap.tar.gz
5 $ tar -xvzf jpcap.tar.gz
6 $ sudo dpkg -i jpcap-0.7.deb
```

Listing A.38: *Comandos para instalação do "jpcap"*

Neste instante, já todas as ferramentas necessárias para o funcionamento do ponto de medição se encontram instaladas, mas ainda falta criar o sistema de diretorias necessário para que a *framework* consiga operar. Todo o processo é explicado no *Listing A.39*

```
1 $ sudo mkdir /usr/local/src/framework
2 $ cd /usr/local/src/framework
3 $ wget https://dl.dropboxusercontent.com/u/49961504/framework/Sampling%20
   Framework%20v3.0.tar.gz
4 $ tar -xvxf Sampling\ Framework\ v3.0.tar.gz
5 $ sudo cp /usr/local/src/framework/Sampling\ Framework\ v3.0/Sampling\
   Framework\ v3.0/SamplingFramework3_0.jar /home/framework/SamplingFramework.
   jar
6 $ sudo chown USER.USER /home/framework/results
7 $ sudo mkdir /home/framework/results
```

Listing A.39: *A criar o sistema de diretorias da framework*

Bibliografia

- Achour et al. (1997)** Mehdi Achour, Friedhelm Betz, Antony Dovgal, Nuno Lopes, Hannes Magnusson, Georg Richter, Damien Seguy e Jakub Vrana. PHP Manual, 1997. URL <http://pt2.php.net/manual/en/>. Citado na pág. 35, 36
- Androulidakis et al. (2009)** Georgios Androulidakis, Vassilis Chatzigiannakis e Symeon Papavasiliou. Network anomaly detection and classification via opportunistic sampling. *IEEE Network*, 23(1):6–12. ISSN 0890-8044. doi: 10.1109/MNET.2009.4804318. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4804318>. Citado na pág. 1
- Cacti Group (2004)** Inc. Cacti Group. Cacti® - The Complete RRDTool-based Graphing Solution, 2004. URL <http://www.cacti.net/>. Accessed: 15/09/14. Citado na pág. 11
- CERT (2012)** NetSA CERT. System for Internet-Level Knowledge (SiLK), 2012. URL <https://tools.netsa.cert.org/silk/>. Accessed: 06/01/2014. Citado na pág. 9
- CERT (2006)** NetSA CERT. Yet Another Flowmeter (YAF), 2006. URL <https://tools.netsa.cert.org/yaf/>. Accessed: 06/01/2014. Citado na pág. 9, 14
- Chabchoub et al. (2007)** Y. Chabchoub, C. Fricker, F. Guillemin e P. Robert. *Deterministic Versus Probabilistic Packet Sampling in the Internet*. Springer Berlin Heidelberg. Citado na pág. 6
- Choi e Bhattacharyya (2005)** Baek-Young Choi e Supratik Bhattacharyya. Observations on cisco sampled netflow. *ACM SIGMETRICS Performance Evaluation Review*, 33(3):18–23. URL http://dl.acm.org/ft_gateway.cfm?id=1111579&type=pdf&CFID=577703042&CFTOKEN=62021565. Citado na pág. 6
- Cisco Systems (2003)** Inc. Cisco Systems. Sampled NetFlow, 2003. URL http://www.cisco.com/c/en/us/td/docs/ios/12_0s/feature/guide/12s_sanf.html#wp1023709. Accessed: 24/09/14. Citado na pág. 10
- Cisco Systems (2004)** Inc. Cisco Systems. NetFlow gives network managers a detailed view of application flows on the network, 2004. URL http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6555/ps6601/prod_case_study0900aecd80311fc2.pdf. Accessed: 24/09/14. Citado na pág. 10
- Claise (2008)** B. Claise. Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information, 2008. URL <http://tools.ietf.org/pdf/rfc5101.pdf>. Accessed: 23/01/14. Citado na pág. 38, 43
- Claise et al. (2013)** B. Claise, Ed. Trammell e P. Aitken. Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information, 2013. URL <http://tools.ietf.org/pdf/rfc7011.pdf>. Accessed: 20/02/14. Citado na pág. 8
- Dogman et al. (2010)** A. Dogman, R. Saatchi e S. Al-Khayatt. An adaptive statistical sampling technique for computer network traffic. Em *Communication Systems Networks and Digital Signal Processing (CSNDSP)*, páginas 479 – 483, Newcastle upon Tyne. International Symposium on. Citado na pág. 6

- Duffield (2004)** Nick Duffield. Sampling for Passive Internet Measurement : A Review. 19(3): 472–498. doi: 10.1214/088342304000000206. Citado na pág. 6
- Galstad et al. (1996)** Ethan Galstad, Mary Starr, Mike Okeefe, Thomas Guyot-Sionnest, Ludmil Miltchev, Nicholas Scott, Jake Omann, Luke Groschen, Eric Stanley, Shamas Demoret, Scott Wilkerson, Katie Montour, Sam Lansing, Andy Brist e Trevor McDonald. Nagios - The Industry Standard In IT Infrastructure Monitoring, 1996. URL <http://www.nagios.org/about>. Accessed: 15/09/14. Citado na pág. 10
- Hernandez A. et al. (2001)** Edwin Hernandez A., Matthew Chidester C. e Alan George D. Adaptive Sampling for Network Management. *Journal of Network and Systems Management*, 9 (4):409–434. doi: 10.1023/A:1012980307500. Citado na pág. 6
- Hu et al. (2008)** C. Hu, S. Wang, J. Tian, B. Liu, Y. Cheng e Y. Chen. Accurate and Efficient Traffic Monitoring Using Adaptive Non-Linear Sampling Method. Em *2008 IEEE INFOCOM - The 27th Conference on Computer Communications*, páginas 26–30. IEEE. ISBN 978-1-4244-2026-1. doi: 10.1109/INFOCOM.2008.14. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4509609>. Citado na pág. 1
- Jiang et al. (2002)** Qingshan Jiang, R. Srinivasan e D. Slonowsky. Measurement based traffic prediction using fuzzy logic. Em *Electrical and Computer Engineering*, volume 2, páginas 834–840. Electrical and Computer Engineering, 2002. IEEE CCECE 2002. Canadian Conference on. doi: 10.1109/CCECE.2002.1013050. Citado na pág. 6
- Lee e Brownlee (2007)** DongJin Lee e Nevil Brownlee. Passive measurement of one-way and two-way flow lifetimes. *ACM SIGCOMM Computer Communication Review*, 37(3):17. ISSN 01464833. doi: 10.1145/1273445.1273448. URL <http://portal.acm.org/citation.cfm?doid=1273445.1273448>. Citado na pág. 8
- Lu e He (2010)** Yiyi Lu e Chen He. Resource allocation using adaptive linear prediction in WDM/TDM EPONs. *AEU - International Journal of Electronics and Communications*, 64(2): 173–176. Citado na pág. 6
- Mark et al. (2008)** L. Mark, M. Stiernerling e P. Aitken. IP Flow Information Export (IPFIX) Implementation Guidelines, 2008. URL <http://tools.ietf.org/pdf/rfc5153.pdf>. Accessed: 27/02/14. Citado na pág. 9
- NASA e ESDIS (2012)** NASA e ESDIS. FlowViewer, 2012. URL <http://sourceforge.net/projects/flowviewer/files/?source=navbar>. Accessed: 06/06/14. Citado na pág. 34
- OETIKER+PARTNER (2006)** OETIKER+PARTNER. MRTG - Tobi Oetiker's MRTG - The Multi Router Traffic Grapher, 2006. URL <http://oss.oetiker.ch/mrtg/index.en.html>. Accessed: 25/09/14. Citado na pág. 11
- Quittek et al. (2004)** J. Quittek, T. Zseby, B. Claise e S. Zander. Requirements for IP Flow Information Export (IPFIX), 2004. URL <http://www.rfc-editor.org/rfc/pdf/rfc3917.txt.pdf>. Accessed: 17/02/14. Citado na pág. 2
- Sadavisan et al. (2009)** G. Sadavisan, N. Brownlee, B. Claise e J. Quittek. Architecture for IP Flow Information Export, 2009. URL <http://tools.ietf.org/pdf/rfc5470.pdf>. Accessed: 20/12/13. Citado na pág. 9
- Schulzrinne et al. (2003)** H. Schulzrinne, S. Casner, R. Frederick e V. Jacobson. RTP : A Transport Protocol for Real-Time Applications, 2003. URL <http://tools.ietf.org/pdf/rfc3550.pdf>. Accessed: 25/09/14. Citado na pág. 8

- Serral-Gracia et al. (2008)** R. Serral-Gracia, A. Cabellos-Aparicio e J. Domingo-Pascual. Packet Loss Estimation Using Distributed Adaptive Sampling. Em *Network Operations and Management Symposium Workshops*, páginas 124 – 131, Salvador da Bahia. NOMS Workshops 2008. IEEE. doi: 10.1109/NOMSW.2007.22. Citado na pág. 6
- SIA (2001)** ZABBIX SIA. Zabbix - The Enterprise-class Monitoring Solution for Everyone, 2001. URL <http://www.zabbix.com/>. Accessed: 25/09/14. Citado na pág. 11
- Silva et al. (2013a)** João Marco C. Silva, Paulo Carvalho e Solange Rito Lima. Enhancing Traffic Sampling scope and efficiency. Em *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, páginas 71–72. Ieee. ISBN 978-1-4799-0056-5. doi: 10.1109/INFCOMW.2013.6562848. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6562848>. Citado na pág. v, vii, xiii, 1, 2, 7, 13
- Silva et al. (2013b)** João Marco C. Silva, Paulo Carvalho e Solange Rito Lima. A multiadaptive sampling technique for cost-effective network measurements. Citado na pág. 7
- Silva et al. (2014a)** João Marco C. Silva, Paulo Carvalho e Solange Rito Lima. Computational Weight of Network Traffic Sampling Techniques. *The nineteenth IEEE Symposium on Computers and Communications*. Citado na pág. 6, 7
- Silva et al. (2014b)** João Marco C. Silva, Solange Rito Lima e Paulo Carvalho. *Monitoring and Securing Virtualized Networks and Services*, volume 8508 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, sperotto, ed. ISBN 978-3-662-43861-9. doi: 10.1007/978-3-662-43862-6. URL <http://link.springer.com/10.1007/978-3-662-43862-6>. Citado na pág. xiii, 7
- Sommers et al. (2005)** Joel Sommers, Paul Barford, Nick Duffield e Amos Ron. Improving accuracy in end-to-end packet loss measurement. Em *Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '05*, volume 35, página 157, New York, New York, USA. ACM Press. ISBN 1595930094. doi: 10.1145/1080091.1080111. URL <http://dl.acm.org/citation.cfm?id=1080091.1080111>. Citado na pág. 1
- Tammaro et al. (2012)** Davide Tammaro, Silvio Valenti, Dario Rossi e Antonio Pescapé. Exploiting packet-sampling measurements for traffic characterization and classification. *International Journal of Network Management*, 22(6):451–476. ISSN 10557148. doi: 10.1002/nem.1802. URL <http://doi.wiley.com/10.1002/nem.1802>. Citado na pág. 1, 2, 6
- Teleinformatikdienste fuer Lehre und Forschung (2004)** Teleinformatikdienste fuer Lehre und Forschung. NfSen tool, 2004. URL <http://nfsen.sourceforge.net/>. Accessed: 25/09/14. Citado na pág. 12
- Thompson (2002)** Steven. K. Thompson. Sampling. *Wiley-Interscience*. Citado na pág. 5
- Thompson e Sever (1996)** Steven. K. Thompson e G.A. Sever. Adaptive sampling. *Wiley Series in Probability in Statistics*. Citado na pág. 5
- Trammell e Boschi (2008)** B. Trammell e E. Boschi. Bidirectional Flow Export Using IP Flow Information Export (IPFIX). páginas 1–24. URL <http://tools.ietf.org/pdf/rfc5103.pdf>. Citado na pág. 9
- Wei et al. (2010)** Yongtao Wei, Wang Jinkuan e Wang Cuirong. A Traffic Prediction Based Bandwidth Management Algorithm of a Future Internet Architecture. Em *Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on*, páginas 560 – 563. IEEE. doi: 10.1109/ICINIS.2010.166. Citado na pág. 6

Zseby et al. (2009) T. Zseby, M. Molina, Nick Duffield, S. Niccolini e F. Raspall. Sampling and Filtering Techniques for IP Packet Selection Status, 2009. URL <http://tools.ietf.org/pdf/rfc5475.pdf>. Accessed: 14/03/14. Citado na pág. [5](#), [6](#)