Vera Gonçalves <veralamonteiro@gmail.com>

Visualização de Sistemas e Redes

Escola Superior de Tecnologia e de Gestão

Agosto 2012

Visualização de Sistemas e Redes

Dissertação apresentada ao Instituto Politécnico de Bragança para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Sistemas de Informação, sob a supervisão do Prof. Doutor Rui Pedro Lopes e do Eng. Tiago Pedrosa.

Vera Gonçalves <veralamonteiro@gmail.com>

Agosto 2012

Prefácio

Nas últimas décadas, o constante avanço das Tecnologias de Informação levaram à geração, manipulação e armazenamento de grandes quantidades de dados, ainda mais potenciados pela diminuição dos custos de dispositivos de armazenamento, pelo acesso facilitado à Intenet e pela evolução de sistemas de informação e ferramentas de gestão de dados. Com esta explosão, surge a necessidade de apostar em novas formas para encontrar formas eficazes e inteligentes de obter e consultar informação útil. Ao nível da gestão de sistemas e de redes, a informação de gestão segue a mesma tendência, tornando indispensável o desenvolvimento de novas formas de manipulação de informação para que a tarefa dos os administradores de sistemas seja o mais eficaz e eficiente possível.

A visualização de informação é aplicada em situações em que a quantidade de dados seja de tal forma grande que torne difícil a tomada de decisões. As diferentes técnicas de visualização procuram ajudar a interpretar os dados, auxiliando a obter conhecimento que, de outra forma seria difícil de conseguir.

Neste contexto, este trabalho procura incidir sobre os paradigmas e as abordagens de visualização aplicadas à temática da gestão de sistemas e de redes. As situações que mais podem beneficiar desta aplicação são várias e diversas, facilitando, em última instância, a interpretação dos dados que resultam da instrumentação do funcionamento dos sistemas e das redes.

Agradecimentos

É com muita satisfação que expresso aqui o mais profundo agradecimento a todos aqueles que tornaram a realização deste trabalho possível.

Agradeço ao Prof. Dr. Rui Pedro Lopes e ao Prof. Tiago Pedrosa pela competência com que orientaram a minha tese e o tempo que generosamente me dedicaram transmitindo-me os melhores e mais úteis ensinamentos, com paciência, lucidez e confiança. Pelo acesso que me facilitaram a uma pesquisa mais alargada e enriquecedora e pelas sua críticas sempre tão atempada, como construtivas, bem-haja estou-lhes muito, muito grata.

Sou muito grata a todos os meus familiares pelo incentivo recebido ao longo destes 2 anos, pela paciência, incentivo e carinho.

Conteúdo

Prefácio **iii**

Agradecimentos **v**

1 Introdução **3**

1.1 Enquadramento 4

1.2 Objectivos 4

1.3 Estrutura do Documento 5

2 Visualização **7**

2.1 O processo de visualização 9

2.1.1 Dados 9

2.1.2 Processamento 10

2.1.3 Mapeamento visual 11

2.1.4 Criação de vistas 11

2.2 Paradigmas 12

2.2.1 Quanto aos dados 12

2.2.2 Quanto à visualização 17

2.3 Resumo 22

3 Sistemas e Redes **23**

3.1 Sistemas 23

3.2 Redes 24

3.3 Segurança 28

3.4 Monitorização e registos 33

3.4.1 Monitorização 33

3.4.2 Registos (*logs*) 36

4 Visualização de Sistemas e Redes **41**
4.1 Procedimento de visualização 41
4.1.1 Dados 42
4.1.2 Visualização 43
4.2 Cenários de aplicação 43
4.2.1 Topologia de Rede 44
4.2.2 Efeito da temperatura ambiente na disponibilidade 46
4.2.3 Capacidade de processamento de routers 48
4.2.4 Controlo de acontecimentos na rede 50
4.2.5 Identificação de desiquilíbrios de rede 53
4.2.6 Identificação da localização geográfica dos utilizadores 54
4.2.7 Identificação de análises à rede 56
4.2.8 Monitorização da carga de rede nos equipamentos 57
4.2.9 Segurança para DNS 58
4.3 Conclusão 61

5 Conclusões **63**

Bibliografia **65**

Lista de Tabelas

4.1 Exemplo de cenários de aplicação 44
4.2 Exemplo de cenários de aplicação 62

Lista de Figuras

2.1 Processo de Visualização. 9
2.2 Rede de frases. 13
2.3 Geo Vista Studio. 14
2.4 WilmaScope. 15
2.5 Snap Together Visualization. 16
2.6 Grafo Simples. 18
2.7 Grafo Completo. 18
2.8 Grafo de 3 níveis. 19
2.9 Pseudografo ou Multigrafo. 19
2.10 Grafo de Árvores. 19
2.11 Treemap. 20
2.12 Séries Temporais. 21
2.13 Exemplo de Scatterplots. 22

3.1 Topologia em Bus 25
3.2 Topologia em Anel 25
3.3 Topologia em Estrela 26
3.4 Topologia em Estrela Estendida 27
3.5 Topologia em Hieráquica 27
3.6 Topologia em Malha 28
3.7 Criptografia Simétrica 31
3.8 Criptografia Assimétrica 31
3.9 Firewall 32
3.10 Gestão de SNMP 36

4.1 Exemplo de um ficheiro de log 42
4.2 Representação gráfica de um domínio. 47
4.3 Temperatura vs Falhas. 49
4.4 Carga média do CPU de um router 50

4.5 Grafo direcionado de sessões SSH intranet como registo por Argus . . 52
4.6 Rede criada por um clusterMaker 's. 54
4.7 Localização actual 55
4.8 InetVis 57
4.9 Gráfico Diário 59
4.10 Gráfico Semanal 59
4.11 Gráfico Mensal 59
4.12 Gráfico Anual 60
4.13 Snort 61

Lista de Acrónimos

TI Tecnologias de Informação

SI Sistema de Informação

WWW World Wide Web

2D Bidimensional

3D Tridimensional

SO Sistema Operativo

SNMP Simple Network Management Protocol

NMAP Network Mapper

SMTP Simple Mail Transfer Protocol

POP3 Post Office Protocol

HTTP Hypertext Transfer Protocol

NNTP Network News Transfer Protocol

ICMP Internet Control Message Protocol

AFT Adress Forwarding Table

DCIM Data Center Infrastruture Managent

QoS Quality of service

Capítulo 1

Introdução

Actualmente, com o desenvolvimento da tecnologia e com a constante miniaturização de plataformas computacionais, tem-se sentido um aumento extraordinário na quantidade de informação que é gerada, armazenada e consultada. Neste cenário, e apesar dos mecanismos existentes para a pesquisa de informação (como, por exemplo, o Google, Beagle, Spotlight, entre outros) torna-se, cada vez mais complexo, extrair conhecimento da informação, devido à dificuldade que os humanos apresentam em processar e correlacionar quantidades massivas de dados.

Na área da gestão de sistemas e de redes este problema também se coloca. De facto, o aumento da diversidade e do número de plataformas computacionais, associado ao aumento generalizado de formas de conectividade e ao aumento de dependência que as organizações têm sobre sistemas em rede torna crítico obter, processar e interpretar a informação de gestão, de forma a manter o sistema em perfeito e regular funcionamento.

Uma das formas possíveis de atacar este problema passa pela utilização de técnicas de visualização que possam apresentar, numa forma mais intuitiva para o utilizador, conhecimento extraído da quantidade de dados que é constantemente gerada. As questões que se colocam de imediato são:

- Como conseguir uma análise correcta dos dados?
- Pode-se usar visualização para descrever a evolução de um determinado valor?
- Como visualizar valores dependentes do tempo?
- Como introduzir ferramentas de visualização no *workflow* de um administrador de sistemas?

Em suma, o grande objectivo deste trabalho é averiguar que ferramentas existem

e que possam ser aplicadas na área da gestão de sistemas e de redes de forma a permitir ao administrador uma mais rápida e adequada tomada de decisão

1.1 Enquadramento

A visualização de problemas originários na gestão de sistemas e de redes com recurso aos meios computacionais levanta diversas questões que é necessário abordar. Inicialmente, é necessário estudar em que consiste a visualização abordando os paradigmas e as respectivas abordagens mais comuns para a visualização de dados. Muitas abordagens usadas recorrem à visualização de árvores ou grafos, representando nós e ligações com formatos diferentes. No entanto, existem diversas abordagens que são mais específicas e trazem mais valias para um determinado tipo de visualização. Nomeadamente na visualização de dados por parte do administrador de sistemas, que necessita de manter o seu sistema seguro, ou de um outro utilizador que faz a monitorização de informação específica, tal como a geração de formas abstratas com recurso a mapas, mapas de calor, vizinhança, entre outras.

Por outro lado, a informação gerada pelos sistemas como resultado da instrumentação ou do regular funcionamento dele é imensa e, como tal, é necessário efectuar um levantamento das potenciais fontes de informação a que tipicamente o administrador recorre no exercício da sua actividade.

1.2 Objectivos

O objectivo desta dissertação é estudar e seleccionar abstracções gráficas para a visualização dos problemas relacionados com a operação de sistemas e redes de comunicação de dados. Para o administrador de sistemas torna-se difícil fazer uma correcta interpretação dos dados sem a ajuda das técnicas de visualização. Para isso, tem de adoptar ferramentas consoante as suas necessidade.

Nos dias de hoje o acesso a inúmera e diversificada informação é algo do quotidiano. É necessário conseguir filtrar o que interessa e usar essa mesma informação de forma produtiva, isto é, usar informação de forma a garantir um sistema seguro, confiável e confidencial na transmissão de dados.

Nesse sentido, é necessário fazer várias pesquisas e perceber que tipo de técnicas e abordagens existem ao nível da visualização e da gestão de sistemas.

1.3 Estrutura do Documento

Este documento está estruturado em cinco capítulos. O actual faz uma introdução e enquadramento ao problema que se vai tratar. O capítulo 2 aborda os paradigmas e as abordagens mais comuns para a visualização de dados. O capítulo 3, descreve as tecnologias e arquitecturas mais comuns de sistemas e redes, abordando os sistemas operativos, aplicações, topologia, segurança e informação de gestão.

O capítulo 4 associa os paradigmas e as abordagens mais adequados para a visualização de informação que resulta do funcionamento de sistemas e de redes. Por último, o capítulo 5 encerra o documento com algumas conclusões.

Capítulo 2

Visualização

A visão é, provavelmente, o sentido humano mais valorizado. De facto, até na linguagem corrente a importância da visualização se encontra patente. Frases como "uma imagem vale mil palavras" ou "vi com os meus própios olhos" são lugares comuns que nos acompanham diariamente e que salientam a característica exclusiva da visão na compreensão do que nos rodeia.

A enorme quantidade de informação que é absorvida pelos olhos é processada pelo cérebro humano que, naturalmente, evoluiu de forma a dedicar uma maior parte a este sentido do que a qualquer dos outros.

É dado adquirido que uma imagem pode ser usada para comunicar uma grande quantidade de informação. Desde uma simples fotografia, que transmite de forma bastante próxima do real um momento, uma paisagem, um tema, e que não se consegue reproduzir tão fielmente por palavras ou por gestos, à representação gráfica de dados numéricos, a imagem providencia um meio de comunicação com uma largura de banda elevada para a transmissão de informação.

Neste sentido, visualização pode ser definida como o processo de geração de imagens baseadas em dados. Consequentemente, visualização procura uma correspondência entre os dados e uma possível representação visual [Marty, 2008]. Assim, a capacidade cognitiva inerente do utilizador é mais eficazmente aproveitada para, em termos visuais, obter, explorar e interpretar informação [Teyseyre and Campo, 2009].

A visualização é um campo de pesquisa muito vasto, que se divide em dois subcampos principais: científico e o da informação. O campo científico representa objectos ou conceitos associados a fenómenos do mundo físico, tais como, a química, metereologia ou o corpo humano. Já o campo da informação representa dados não espaciais que envolvem conceitos e relações, tais como, dados financeiros, bibliografias, fontes gráficas ou software. Os dados não são mais que uma recolha de informação organizada proveniente dos mais diversos locais, tais como, resultados

de uma pesquisa, experiências, observação de outras informações dentro de um SI ou através de um conjunto de instalações. Os dados podem ser números, palavras ou imagens.

A visualização encontra aplicação em várias áreas, como na matemática, física, lógica e informática, em que o objectivo consiste em facilitar ao utilizador uma representação visual clara da informação. A questão principal reside em identificar como melhor extrair informação útil, de forma rápida, de um grande volume de dados [Fry, 2008]. [Knight, 2000] destaca três desafios na visualização:

- o volume crescente de dados,
- ferramentas e modelos de análise de dados cada vez mais complexos e dados cada vez mais abstratos,
- utilizadores menos especializados.

Segundo [Hessen, 2000], a essência do conhecimento está intimamente relacionada com a correlação entre dois objectos: o **utilizador** e o **objecto**, considerando que a função do primeiro é aprender o segundo e que a função do segundo é ser aprendido pelo primeiro. Em contexto de visualização a aquisição de conhecimento pode ser interpretada de forma diferente pelos diversos utilizadores, dependendo da percepção que cada um tem perante determinado objecto. A representação da informação não é mais que um processo cognitivo que estabelece uma relação entre o utilizador e os dados. Ao tratar os dados, o utilizador aplica os seus conhecimentos e experiências, provocando um desencadear de ideias. Desta forma um outro utilizador ao possuir o mesmo tipo de informação também faz a sua própria interpretação, podendo chegar a resultados diferentes.

Na opinião de [Targino, 1995] a Web pemite amenizar a ansiedade de informação dos utilizadores nas pesquisas, contribuindo para que encontrem o que procuram em ambientes informacionais, com o objectivo final de obter informação correcta.

Os paradigmas e abordagens aplicadas à visualização acompanham a evolução dos meios de cálculo, tendo iniciado com a evolução do computador na II Grande Guerra nos EUA. A Marinha, a Universidade de Harvard e a IBM desenvolveram o Mark I, em 1944. Por sua vez, o exército também desenvolvia o seu computador, com a finalidade de o usar para calcular as trajectórias dos projecteis com mais precisão, ficando concluido só em 1946. Em 1947, um grupo de Stanford inventou o transistor, o que revolucionou a indústria da computação devido, essencialmente, a ser mais rápido, mais duradouro e de menor consumo que as tradicionais válvulas electrónicas [Hamann, 2011]. A visualização, tal como muitas outras tecnologias, tiram proveito desta evolução, tendo surgido novos paradigmas e abordagens para

representar estruturas de dados e informação complexa, tais como o comércio, segurança, redes e pesquisas de mercado [Andery, 2010].

2.1 O processo de visualização

A visualização dos dados recorre à exploração visual do ser humano, às suas capacidades de percepção. O objectivo final é conseguir visualizar os dados de forma simples e conseguir rapidamente ver as representações visuais de dados para detectar problemas e consequente ajudar na tomada de decisões.

O modelo genérico subjacente à visualização (ou, por outras palavras, à construção de vistas que representam graficamente dados) compreende diversas fases, iniciando com a aquisição dos dados em bruto (Figura 2.1):

- Pré-processamento e transformação de dados
- Mapeamento visual
- Criação de vistas

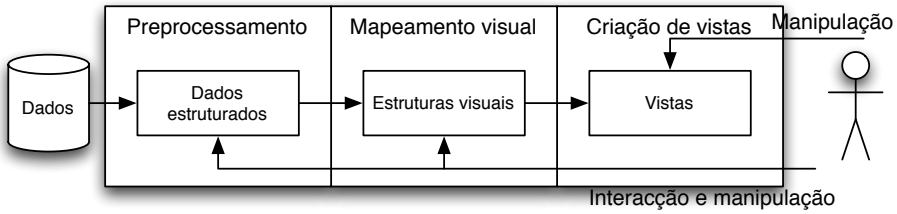


Figura 2.1: Processo de Visualização.

2.1.1 Dados

A proveniência dos dados apresenta diferenças ao nível de sintaxe e de semântica. Eventualmente, também poderá ser necessário correlacionar dados de fontes distintas conforme os dados a serem visualizados.

Os dados podem ser gerados ou obtidos de diversas formas, com origem igualmente diversa. Estes podem ser adquiridos de sondas, instrumentação de monitorização, ou medidos por um processo de quantificação. A natureza destes é diversa, pois podem representar qualquer variável que se encontra na natureza ou mesmo resultante de conceitos sociais ou humanos.

Previamente a qualquer processamento, os dados encontram-se numa forma tipicamente austera, difícil de compreender de forma directa. Adicionalmente, podem

apresentar-se não estruturados e com um nível elevado de verbosidade, o que aumenta a dificuldade em os interpretar. Neste sentido, é necessário proceder a algum condicionamento ou processamento, de forma a extrair alguma estrutura.

Em termos de natureza, os dados podem ser quantitativos (inteiros ou números reais), ordinais (que apresentam uma noção de ordem) ou que representam categorias (sem ordem). Adicionalmente, os dados a representar podem descrever um determinado número de dimensões:

- unidimensional – variação de uma variável em função de outra (independente),
- bidimensional – variação de duas variáveis inter-dependentes em função de outra (independente),
- tridimensional – variação de três variáveis inter-dependentes em função de outra (independente),
- multidimensional – variação de quatro ou mais variáveis inter-dependentes, em função de outra (independente).

2.1.2 Processamento

O formato original dos dados não permite, tipicamente, uma extracção efectiva de conhecimento. De facto, a natureza e a quatidade destes requer um processamento prévio, no sentido de lhe conceder alguma estrutura lógica. Os dados são um conjunto de registos complexos e difíceis de visualizar e tirar conclusões. Na sua forma mais simples, a estrutura pode passar por um formato tabular, mais adequado a ser processado pelo software. A fase de processamento pode ser relativamente simples, como retirar campos ou modificar separadores, ou complexa, com algoritmos de inteligência artificial para extracção de características, por exemplo [sec, 2011].

Após o processamento, a informação estará, à partida, num formato mais adequado e/ou a sua quantidade controlada de forma a possibilitar a sua visualização segundo a técnica escolhida. Este processo envolve:

- transformação dos dados atráves de métodos (filtragem de ruído, mudandança de resolução, transformação de escala, etc.);
- adaptação para uma forma apropriada de representação gráfica (côr, atributos, etc.);
- sequência de imagens.

Como resultado do processamento dos dados, são construídas estruturas de dados. Estas podem ser:

- lineares – vectores, tabelas, arrays, ...
- espaciais ou geográficas – correspondendo a objectos físicos, tais como mapas, plantas, ...
- temporais – que alteram em função do tempo,
- hierárquicas – organizados de forma a representar uma hierarquia (genealogia, fluxogramas, ...),
- rede – representando relações entre entidades (topologia, ...).

2.1.3 Mapeamento visual

O principal problema a resolver nesta fase assenta na definição de estruturas visuais que serão usadas para fazer a correspondência entre os dados e a sua localização na imagem. Nem todos os dados terão correspondência, no entanto, outros, terão facilmente enquadramento e significado visual [Card et al., 1999].

Como exemplo, há dados, com relevância geográfica, que podem ser representados numa posição específica num mapa. Por outro lado, dados hierárquicos também ocuparão uma posição lógica, tal como o mapa de uma rede de computadores.

Adicionalmente, esta fase também incide sobre a definição de elementos gráficos, tais como símbolos ou formas específicas, e suas propriedades (côr, dimensão, etc.).

Por último, dependendo da interação com o utilizador, esta fase determina se a visualização é:

- estática – impressão em papel, por exemplo
- transformável – permitindo alteração de valores ou alteração de parâmetros de mapeamento
- manipulável – o utilizador pode modificar patâmetros que regulam a geração das vistas (zoom, rotação, ...)

2.1.4 Criação de vistas

As vistas são o resultado final do processo de visualização. Estas procuram representar a correspondência entre a estrutura de dados e as estruturas visuais, gerando uma representação visual tipicamente com o auxílio do computador. A representação visual procura dar resposta eficiente a questões colocadas no início do processo.

A dificuldade principal nesta fase é lidar com a quantidade de dados a representar face ao espaço visual disponível. Tipicamente, há grandes quantidades de dados

que se procura representar. No entanto, apenas uma pequena área no ecran ou no papel para tirar proveito. Nestas situações há um conjunto de ferramentas que ajudam, como sendo a possibilidade de *zooming*, *panning*, *scrolling*, entre outros [Mazza, 2009].

2.2 Paradigmas

A evolução e aparecimento de novos paradigmas da visualização tem acompanhado o avanço tecnológico, em particular das tecnologias de informação e comunicação. Com a disponibilização de mais capacidade gráfica, de processamento e de armazenamento, novas formas de geração de imagens são exploradas, tirando partido da capacidade adicional. Adicionalmente, uma maior quantidade de dados consegue ser processada por unidade de tempo, tornando viável o alargamento de áreas que podem tirar proveito dos paradigmas da visualização

Em termos simplistas, a visualização de informação obedece a um *workflow* bem definido: aquisição, processamento, representação, depuração e interacção. Enquanto que a aquisição de informação compreende uma tarefa relativamente simples, o mesmo já não se pode dizer quanto aos passos subsequentes. No entanto, o resultado final deve ser preciso, intuitivo e atraente para o utilizador [Cerqueira, 2010].

Independentemente da técnica de visualização e dos algoritmos de processamento, os tipos de dados básicos subjacentes à visualização podem ser de vários tipos e de várias dimensões. Neste sentido, tipicamente, os dados abrangem desde o unidimensional ao multi-dimensional, sendo casos específicos os dados com informação temporal ou hierárquica. Já a visualização, enquadra-se, tipicamente, em ambientes 2D, 2.5D ou 3D.

2.2.1 Quanto aos dados

O paradigma de visualização depende, em grande parte, do tipo de dados a visualizar. Tal como referido anteriormente, estes correspondem, tipicamente, a uma ou mais variáveis dependentes que variam em função de outra(s), independente(s).

Unidimensionais

A visualização unidimensional compreende dados em que uma variável varia relativamente a um ou mais atributos independentes. Como exemplo, incluem-se documentos textuais, código fonte de programas, listas sequenciais (nomes em ordem alfabética, por exemplo), etc. Cada item do conjunto de dados é uma linha de texto

que contém uma lista de características. No entanto, pode ser necessário haver uma linha adicional contendo outros atributos de interesse (independentes), como a data da última actualização. Neste tipo de visualização é importante que a interface utilizada represente essas características, tais como, o tipo de letra, côr, tamanho, entre outros. Este tipo de visualização oferece uma resposta mais efectiva às necessidades do utilizador para além de uma vista global e compacta entre si através de pequenas alterações, como por exemplo, a côr, tipo de fonte, tamanho, orientação, posição e escala, que torna a visualização mais rica e fácil de comparar, filtar ou pesquisar os dados na lista. Utiliza como conceito básico de listas sequenciais baseadas na leitura/pesquisa de textos.

Um exemplo de visualização unidimensional incide sobre a construção de redes de frases, uma técnica de visualização de texto [Van Ham et al., 2009]. Neste caso, é gerada uma imagem com determinada ordenação de texto não estruturado. O resultado final é um grafo em que as palavras são os nós e as arestas indicam palavras relacionadas de acordo com uma determinada especificação. Estas relações podem ser especificadas a nível semântico ou sintáxico, sendo que diferentes relações produzem diferentes perspectivas sobre o mesmo texto. No seu conjunto, estas perspectivas providenciam um apanhado de conceitos e relações chave num documento ou num conjunto de documentos (Figura 2.2).

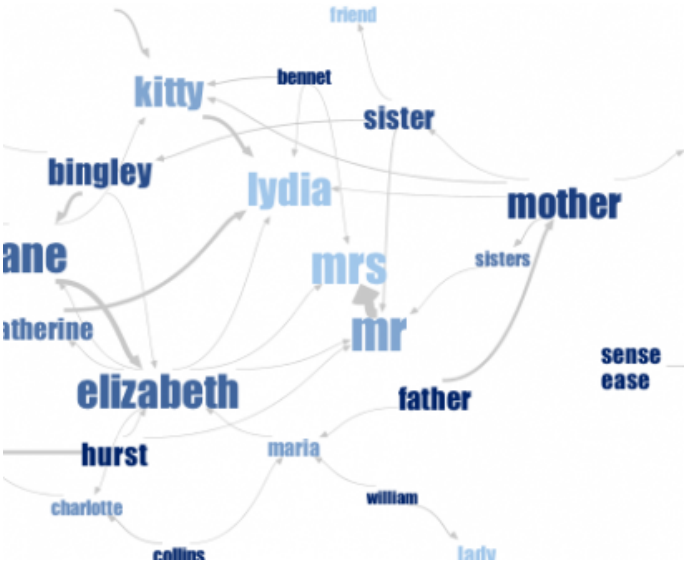


Figura 2.2: Rede de frases.

Outros exemplos incluem gráficos de barras, tabelas (duas colunas) ou outros.

Bidimensionais

Os dados do tipo bidimensional compreendem duas variáveis dependentes que variam em função de um ou mais valores independentes. Incluem-se, naturalmente, dados que, pela sua natureza, têm mais de dois atributos. Um exemplo é a relação entre as importações e exportações de países, que pode ser representado por intermédio de uma tabela com três colunas. Para ser mais representativo, este mesmo exemplo pode ser apresentado como um gráfico de dois eixos, sendo um eixo as importações e outro as exportações, servindo como coordenadas para posicionar o país no gráfico.

Outros exemplos incluem mapas geográficos, plantas de prédios, esquemas de hardware. Aplicações que tiram proveito de paradigmas bidimensionais incluem, por exemplo, programas de edição de imagem, programas de composição de publicações, entre outros. Por exemplo, o método Geo Vista Studio permite fazer o estudo da visualização dos dados geográficos, a construção rápida de aplicações e processamento de dados, tendo como vantagem ser desenvolvido em código aberto (Figura 2.3).

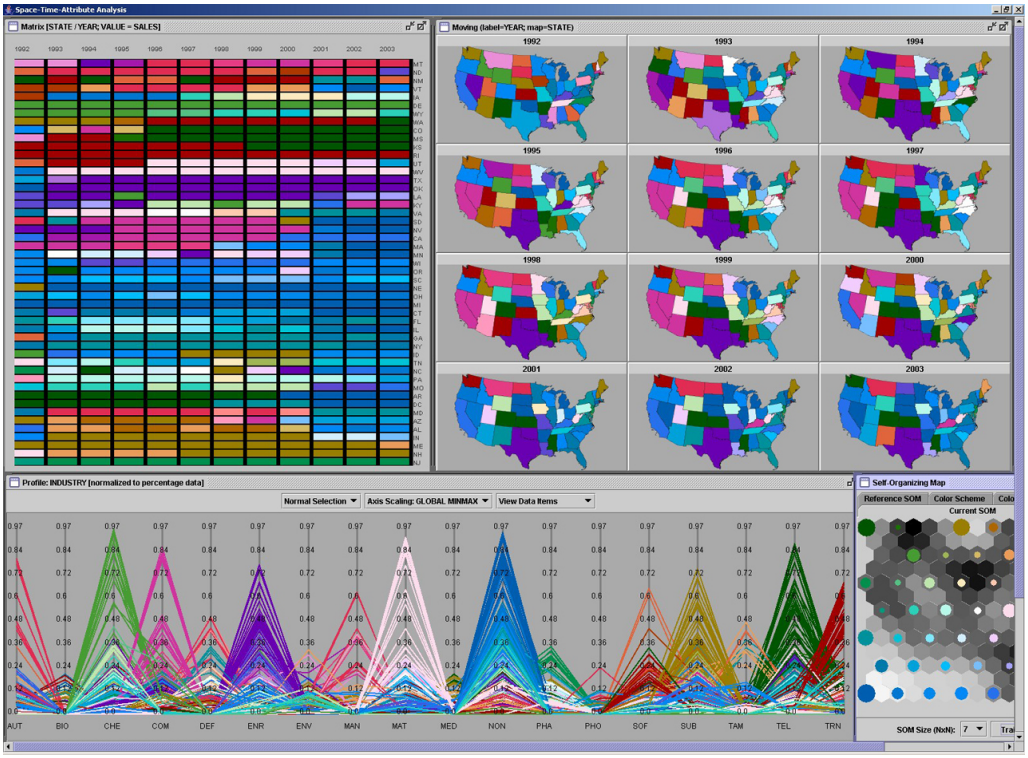


Figura 2.3: Geo Vista Studio.

Tridimensionais

A visualização tridimensional inclui dados em que três variáveis dependentes variam em função de um ou mais valores independentes. Nestes casos, a representação não é possível ser feita recorrendo a duas dimensões, sendo necessário aplicar métodos de visualização mais complexas (3D, por exemplo). Para ilustrar dados tridimensionais, tome-se o exemplo anterior, em que era feita a correlação entre as importações e as exportações em função do país e adicione-se o PIB (produto interno bruto) de cada país. Neste caso, os eixos de coordenadas serão, respectivamente, as importações, exportações e o PIB, fixando a posição de cada país num espaço 3D. O resultado final poderá ser um *scatterplot* (secção 2.2.2).

Outro exemplo ambrange a representação de objectos do mundo real, combinando-os com a visualização 2D, de forma a visualizar um relacionamento entre vários items ou características [Carvalho and Marcos, 2009]. Neste tipo de visualização, o utilizador consegue explorar o espaço tridimensional, que requer posicionamento e capacidade de orientação neste tipo de campo. Por exemplo, o software WilmaScope usa Java3D e faz animação de grafos em 3D permitindo uma navegação interactiva (Figura 2.4).

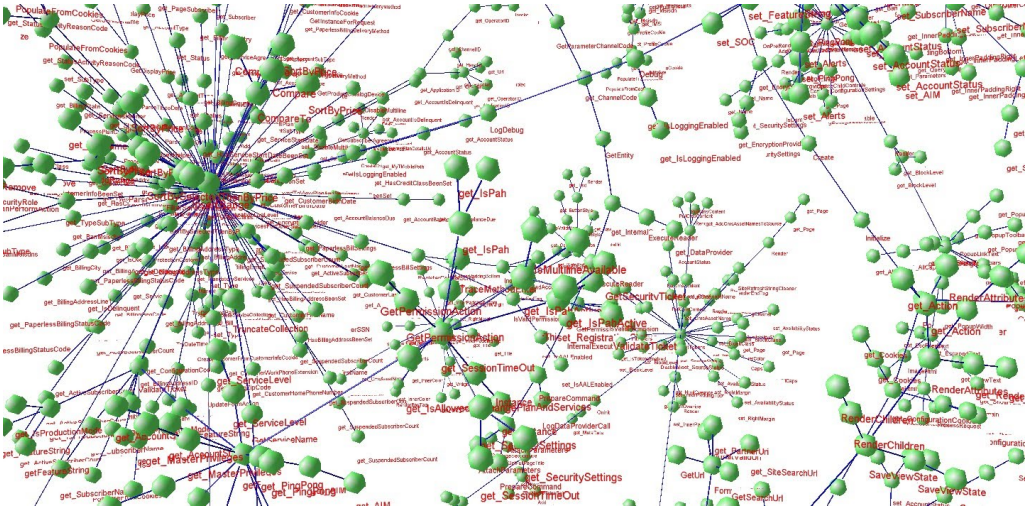


Figura 2.4: WilmaScope.

Multidimensionais

A visualização multidimensional está relacionada com um conjunto de dados relacionais e estatísticos que são manipulados como multidimensionais, uma vez que a cada objecto está associado a N atributos. As principais dificuldades desta visualização

é encontrar padrões, grupos e correlações entre pares de objectos. Os dados podem ser representados sob a forma 3D acrescido de técnicas de apresentação múltiplas de várias perspectivas [Silva, 2007]. Os objectivos básicos desta visualização são:

- Perceber um conjunto de dados de N dimensões, permitindo encontrar padrões, relações, limites, falhas, etc..
- Encontrar um determinado dado especifíco.

Por exemplo, o Snap Together Visualization permite criar e publicar na Internet mútiplas visualizações de base de dados relacionais e as suas respectivas relações (Figura 2.5). Podemos concluir, que a visualização apresenta um conjunto de soluções para representar e explorar as diversas técnicas de visualização.

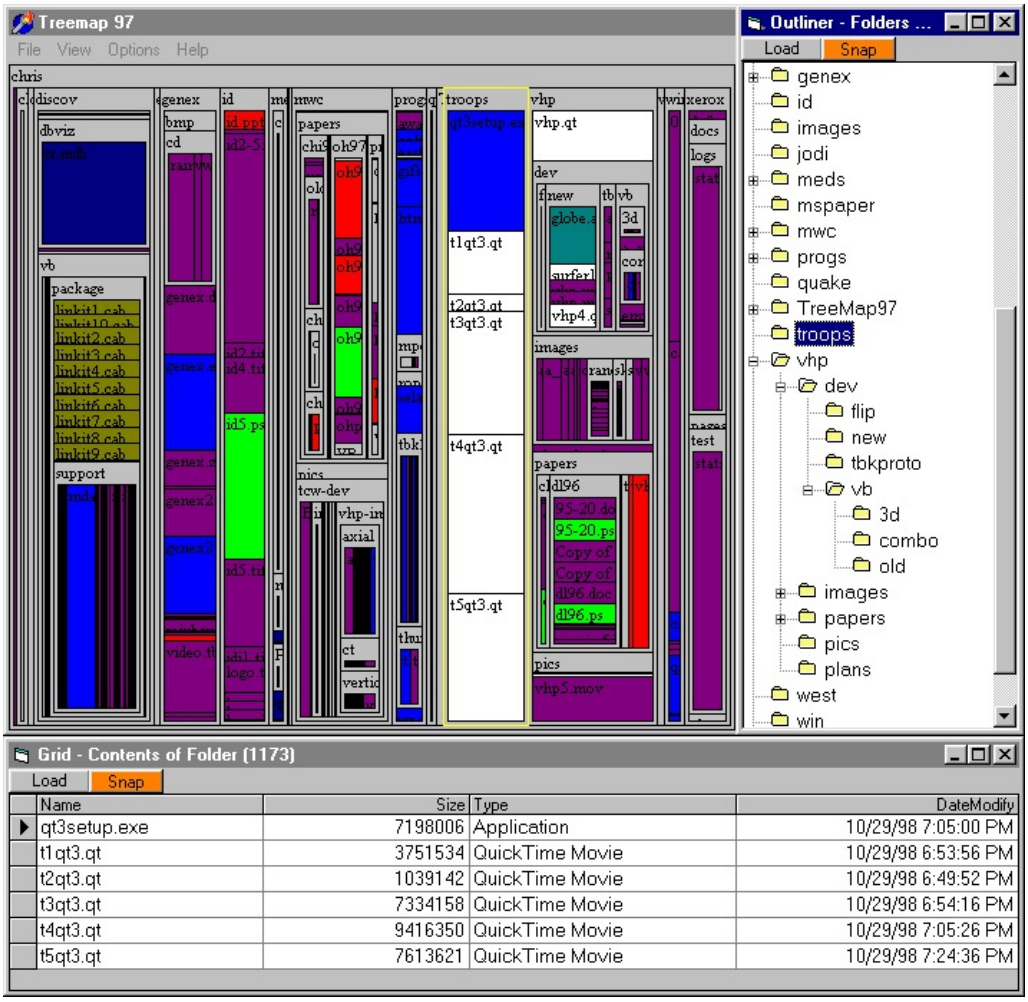


Figura 2.5: Snap Together Visualization.

2.2.2 Quanto à visualização

As técnicas de visualização têm como finalidade criar representações visuais e em cada representação os dados são mapeados num elemento visual, como por exemplo num círculo, num ponto ou esfera, num espaço com as suas respectivas relações que reflectem o tipo de relacionamento entre os dados [Paulovich, 2009]. Estes podem ser apresentados do tipo: escalares (unidimensional), vectoriais (bidimensional), tensoriais (tridimensional), multidimensional [Carvalho and Marcos, 2009].

Visualização leva à interpretação de grandes quantidades de informação. Por isso, o uso das técnicas de visualização e exploração tentam ajudar a interpretar de forma amigável e clara. Ao realizar uma determinada pesquisa a palavra chave é Informação, logo é, necessário planear como a apresentar de forma a chamar a atenção do utilizador. Os paradigmas visuais apelam às características naturais do ser humano (visão), e com o seu conhecimento para criar interfaces que podem ser facilmente aprendidas. Para mostrar a informação é necessário planear e escolher a melhor técnica de visualização a ser usada e associar um ou mais atributos reais. Para ajudar a interpretar esta ferramenta, visualização, surgem algumas técnicas de representação visual: grafos, árvores, mapas, agrupamentos (*clusters*), gráficos temporais e *scartterplots* [Teyseyre and Campo, 2009].

Subjacente a todos estes paradigmas encontra-se a forma de representação, em concreto, o número de dimensões da vista: 2D ou 3D.

2D e 3D

A criação de vistas é feita, tipicamente, sobre o ecran de um computador. Este dispositivo é intrinsecamente 2D, permitindo representar elementos visuais sobre um plano de duas coordenadas. A visualização em 3D, em que se representam objectos com base em três coordenadas, é feita por intermédio de uma transformação de coordenadas com adição de perspectiva.

A visualização em 2D é, tipicamente, mais clara e rigorosa, enquanto que a visualização 3D possui alguns problemas intrínsecos. Um dos problemas é relacionado com a própria percepção do utilizador, levando a que, em alguns casos, haja uma sobrecarga cognitiva, ou seja, um aumento do esforço mental para interpretar a imagem.

Este facto não quer dizer que a construção de vistas em 3D deva ser sempre evitada. Pode haver situações em que este tipo de abordagem funciona melhor.

Grafos

Os grafos são um modelo matemático que estuda as relações entre objectos de um determinado conjunto. A sua estrutura é muito simples, obedecendo à função $g = f(v, a)$, em que v são os vértices e correspondem aos objectos, e a são as arestas, que representam as relações entre os objectos [Nascimento, 2005]. Existem diversos tipos de grafos [Wikipédia,]:

- Grafo simples – É um grafo não direccionado em que existe no máximo uma aresta entre dois vértices mas sem arestas paralelas (Figura 2.6);
- Grafo completo – É um grafo simples em que o vértice é adjacente a todos os outros vértices (Figura 2.7);
- Grafo regular – É um grafo onde todos os vértices tem o mesmo grau (Figura 2.8);
- Pseudografo ou Multigrafo – É um grafo não dirigido que pode ter arestas múltiplas ou paralelas, isto é, arestas com os mesmos vértices finais (Figura 2.9).

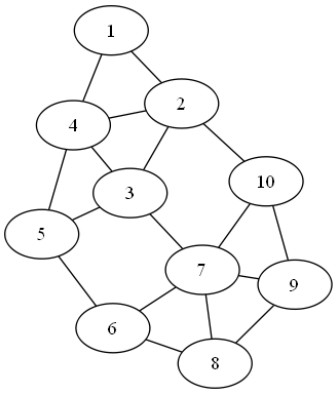


Figura 2.6: Grafo Simples.

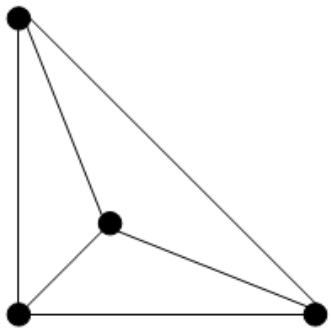


Figura 2.7: Grafo Completo.

Árvores

As árvores são grafos simples acíclicos e conexos. Um dos vértices da árvore é distinto e chama-se de raíz, e tem ligação com os outros objectos que são denominados de ramos ou filhos. Estes ramos levam a outros elementos que também possuem ramos. O elemento que não possui ramos é conhecido por folha ou nó terminal. Este tipo de grafo é normalmente usado como estruturas de dados na área da Informática (Figura 2.10) [Nascimento, 2005].

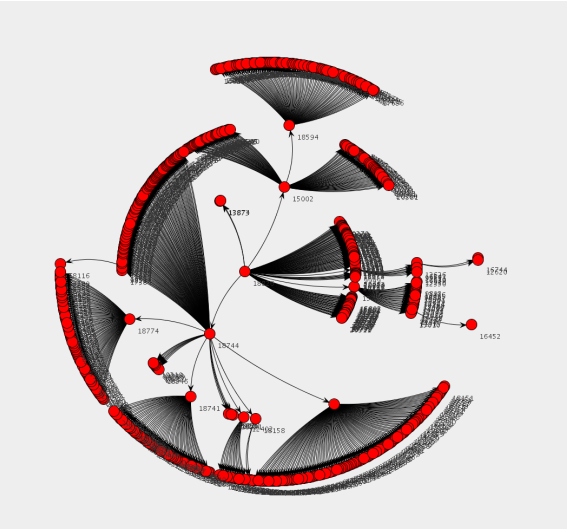


Figura 2.8: Grafo de 3 níveis.

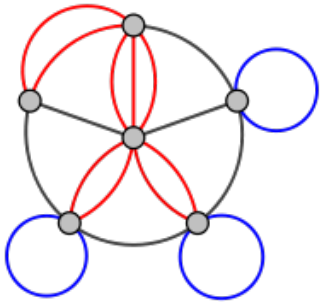


Figura 2.9: Pseudografo ou Multigrafo.

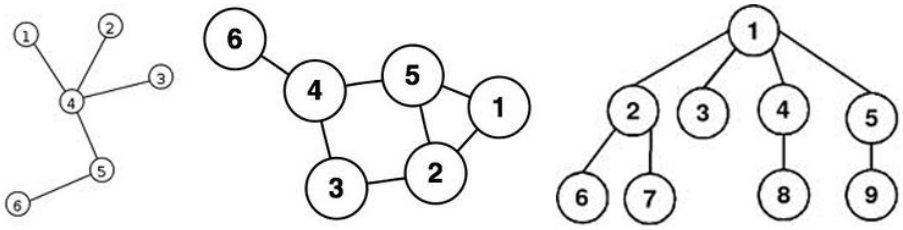


Figura 2.10: Grafo de Árvores.

Mapas

Neste paradigma de visualização é muito usado o *treemap* para mostrar dados hierárquicos através de rectângulos alinhados parecendo azulejos de várias cores (Figura 2.11). O *treemap* é um modelo de árvore estruturado com dados distribuidos por um conjunto de rectângulos alinhados. Cada ramo ou nó da árvore é um rectângulo de dados e os restantes menores representamos sub-ramos. Quando a côr e o tamanho estão correlacionados de alguma forma pode-se ver facilmente padrões que com outro tipo de árvores é difícil de detectar. A segunda vantagem é que a sua construção permite o uso eficiente de espaço, resultando de numa exibição legível de milhares de items simultâneamente [Page, 1994].

Gráficos Temporais

Os gráficos temporais representam dados com indicação de tempo, compreendendo uma sequência de dados numéricos em ordem sucessiva, observados ao longo do

Figura 2.11: Treemap.

tempo, e normalmente, com um intervalo uniforme. A sequência de dados podem dar origem a séries temporais ou a uma série não temporal. O que difere uma da outra é que a primeira, mostra, por exemplo, a temperatura diária de uma cidade ao longo de vários dias, existindo uma sequência. A segunda, define um conjunto de dados sobre a temperatura de uma certa cidade, em vários locais ao longo de um dia. A característica mais relevante deste tipo de dados é que permite ver as vizinhanças que são dependentes e analisar a forma como se moldam (Figura 2.12) [Ehlers, 2009].

A particularidade destes dados permite-nos:

- observar dados correlacionais que são mais difíceis de analisar e que necessitam de técnicas específicas;
- ter em conta o espaço temporal dos dados;
- seleccionar modelos que podem ser bastante complicados de analisar e as ferramentas podem não ajudar;

No entanto, é difícil de fazer observações em que não haja uma sequência, devido à discrepância dos dados a nível temporal.

O objectivo de estudar este tipo de dados é permitir:

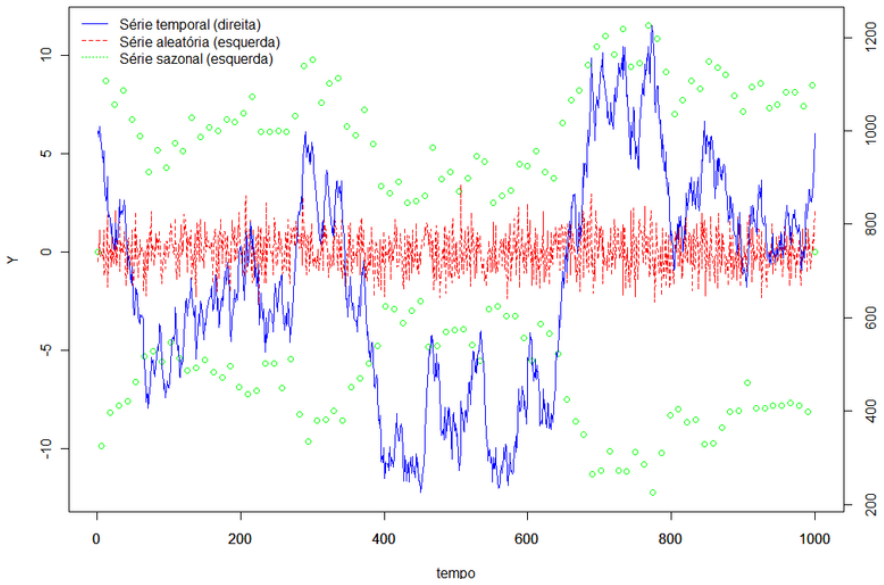


Figura 2.12: Séries Temporais.

- descrever – descrever as propriedades da série, como por exemplo, o padrão de tendência, as variáveis sazonais ou cíclicas, observação das discrepâncias entre os dados;
- explicar – consegue usar a variação da série para explicar a variação da outra série;
- prever – consegue prever uma tendência de valores no futuro, com base nos dados passados;
- controlar – os dados são propensos a seguir uma tendência, logo é possível medir a "qualidade" dos dados de forma a controlar o seu processo.

Este tipo de abordagem permite recorrer a alguns tipos de técnicas de visualização tais como: descritiva, gráfica, identificação de padrões, modelos probabilísticos, entre outras. É usada nas mais diversas áreas, desde a Ecomonia, Medicina, Meteorologia, etc..

Scatterplots

Os *scatterplots* são gráficos de dispersão, considerados como sendo um diagrama baseado em coordenadas cartesianas. São visíveis os valores em duas ou três variáveis, sendo a correlação entre elas feita por pontos. Cada ponto representa um valor que

determina a posição, tanto no eixo horizontal como no eixo vertical ou de profundidade [Cleveland, 2007]. Estes gráficos exibem dados contínuos, tendo em conta uma escala comum, que mostra a evolução dos dados em intervalos irregulares ou agrupados. São gráficos favoráveis para mostrar dados e serem usados como ferramentas de qualidade na análise dos mesmos, pois mostra uma possível relação de causa e efeito. Os scatterplots são muito parecidos aos gráficos temporais [office, 2010] (Figura 2.13).

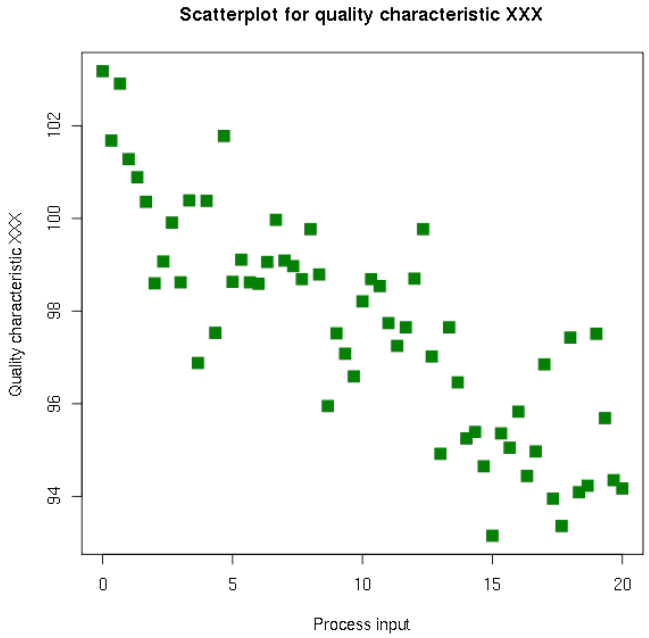


Figura 2.13: Exemplo de Scatterplots.

2.3 Resumo

A visualização de dados é uma abordagem complexa, compreendo várias opções e um workflow bem definido. Quando bem dimensionada permite construir vistas que, devido a características inerentes do sistema perceptivo humano, facilita a aquisição e interpretação de informação. Desde a aquisição de dados, passando pela construção de estruturas de dados, definição de elementos gráficos e construção das vistas, todas as fases exigem capacidade computacional por vezes assinalável, tendo, como resultado final, uma perspectiva do mundo que seria, de outra forma, difícil de contruir.

Capítulo 3

Sistemas e Redes

As organizações dependem cada vez mais das tecnologias de informação. A revolução que se iniciou há umas décadas com a utilização de sistemas informáticos nas organizações escalou para o uso de variadíssimos sistemas interdependentes suportados por redes de comunicação complexas. Neste capítulo debate-se a complexidade dos sistemas e das redes que necessitam de ser geridas hoje em dia pelos seus administradores. Inicia-se com a secção dos sistemas onde se caracteriza a heterogeneidade destes, as operações normais que asseguram bem como as actividades dos administradores. Segue-se a secção das redes onde se definem as topologia de rede mais utilizadas, bem como desafios que os administradores podem encontrar. Apresentam-se conceitos sobre segurança da informação e alguns mecanismos de segurança na secção seguinte, culminando com a secção de monitorização e registos, onde se discutem formas de monitorização de sistemas e redes e como tirar partido dos registos que resultam do funcionamento dos mesmos.

3.1 Sistemas

Os sistemas são uma parte integrante da actividades diárias nas organizações. O número de sistemas em cada organização aumentou e estes diversificaram-se. As actividades do negócio dependem do correcto funcionamento destes. Uma monitorização activa e um controlo do que acontece em cada sistema torna-se fundamental.

Os sistemas são normalmente divididos tendo em conta a sua função, de um lado estão as estações de trabalho do outro os servidores. Os servidores são de especial importância, pois são eles que disponibilizam serviços que podem ser consumidos por clientes que se encontram em execução em estações de trabalho. Falhas neste tipo de sistemas normalmente têm um impacto mais elevado do que nas estações de trabalho pois afectam diversos utilizadores simultaneamente. As estações de

trabalho têm normalmente instaladas ferramentas de escritório que são utilizadas pelo utilizador de cada sistema, alguns deles usam aplicações cliente que fazem uso de serviços disponibilizados pelos servidores.

O número de serviços disponibilizados e utilizados pelas organizações têm aumentando, o que obriga os administradores a controlarem e lidarem com mais serviços. Também passou a ser comum ter diversos sistemas operativos em funcionamento em cada organização. É normal encontrar sistemas Windows, Linux, MacOSX e também os sistemas operativos móveis.

Cada tipo de sistema operativo tem abordagens diferentes para a instalação, configuração e monitorização. Similarmente cada aplicação e serviço tem particularidades. Esta conjugação dificulta o trabalho do administrador de sistemas.

Uma actividade muito importante que os administradores têm que fazer é verificar se os sistemas estão a ser utilizados e executados da forma que pretendem. Este trabalho passa principalmente pelo estudo dos registos que cada sistema, serviço e aplicação gera. Considerando que o volume de registos tem aumentado, associado ao facto de que cada um tem a sua sintaxe própria, é de extrema importância ter ferramentas que permitam facilmente detectar e correlacionar os eventos registados.

3.2 Redes

As redes representam um conjunto de sistemas computacionais (também conhecidos por nós) interligados. A topologia de rede consiste numa forma de descrever o *layout* de uma rede informática através da qual circula informação. As topologias podem ser descritas fisicamente, considerando a disposição e interligação dos nós, e logicamente, considerando o fluxo de informação que circula na rede. Basicamente, a topologia física esta relacionada com a distribuição do *layout* pelos terminais, isto é, a forma física em que os terminais estão dispostos no ambiente de rede. Enquanto, a topologia lógica mostra o funcionamento de rede – por outras palavras, qual o método de acesso ao meio que esta a ser usado pelos terminais.

De entre as várias topologias de configuração, podem-se considerar as seguintes [Halsall, 1995], [Tanenbaum, 2002]:

Barramento

A ligação entre vários computadores é feita através de um único cabo. Com esta topologia todos os computadores recolhem a informação destinada a si próprios (Figura 3.1). A comunicação é geralmente partilhada no tempo ou na frequência, podendo surgir colisões o que obriga a que a comunicação seja reiniciada. Esta topologia

assenta na utilização de um único cabo como meio de transmissão e encontra-se, actualmente, em desuso.

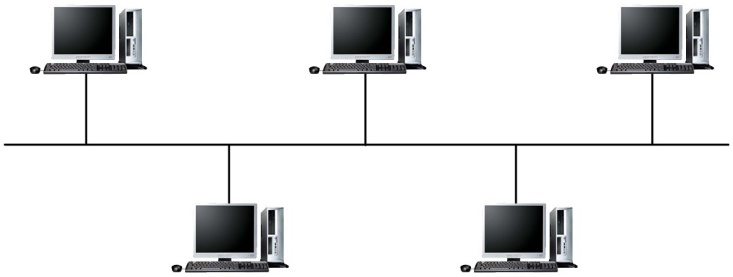


Figura 3.1: Topologia em Bus

Anel

Os dispositivos encontram-se ligados em série formando um circuito fechado (anel). Os dados transmitidos são unidireccionais de nó para nó até chegar ao seu destino. Quando é enviada uma mensagem, esta terá de passar por todas as estações até chegar ao destino final (Figura 3.2).

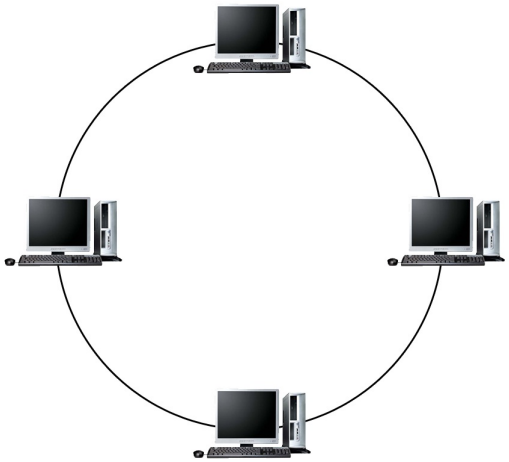


Figura 3.2: Topologia em Anel

Estrela

É a topologia mais usada, utiliza um transmissor (*hub* ou *switch*) como ponto central da rede para efectuar a transmissão de dados (Figura 3.3). Tem como vantagem a localização rápida de problemas, uma vez que consegue detectar qual a secção da rede que tem problemas. Este tipo de tipologia é usado apenas em redes pequenas.

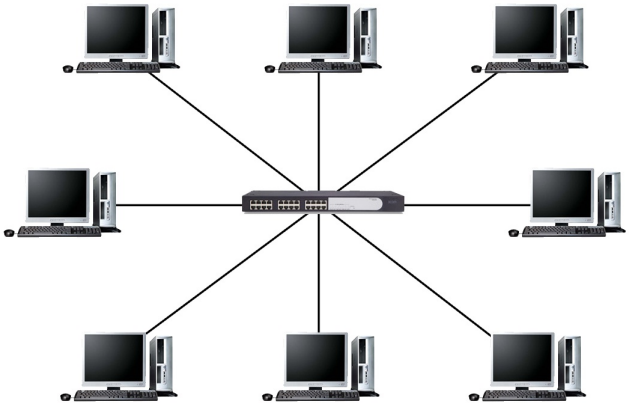


Figura 3.3: Topologia em Estrela

Estrela Estendida

A diferença entre esta topologia e a topologia em Estrela consiste que no centro da estrela existe um *hub* ou *switch* que permite a ligação às restantes estrelas. Permite um aumento da rede em tamanho e comprimento (Figura 3.4).

Hierárquico

Esta topologia é semelhante à topologia de Estrela Estendida, excepto no facto de ser um computador a controlar a rede e não um *hub* ou *swicth* (Figura 3.5).

Malha

A comunicação é feita ponto-a-ponto entre os computadores da rede, tendo como vantagem permitir que cada computador tenha uma linha privilegiada de comunicação com qualquer outro dispositivo na rede (Figura 3.6). Pode-se constatar uma redundância na comunicação entre os vários dispositivos, que poderá aumentar a resiliência a falhas de comunicação, pois será possível utilizar as outras ligações para chegar ao destino.

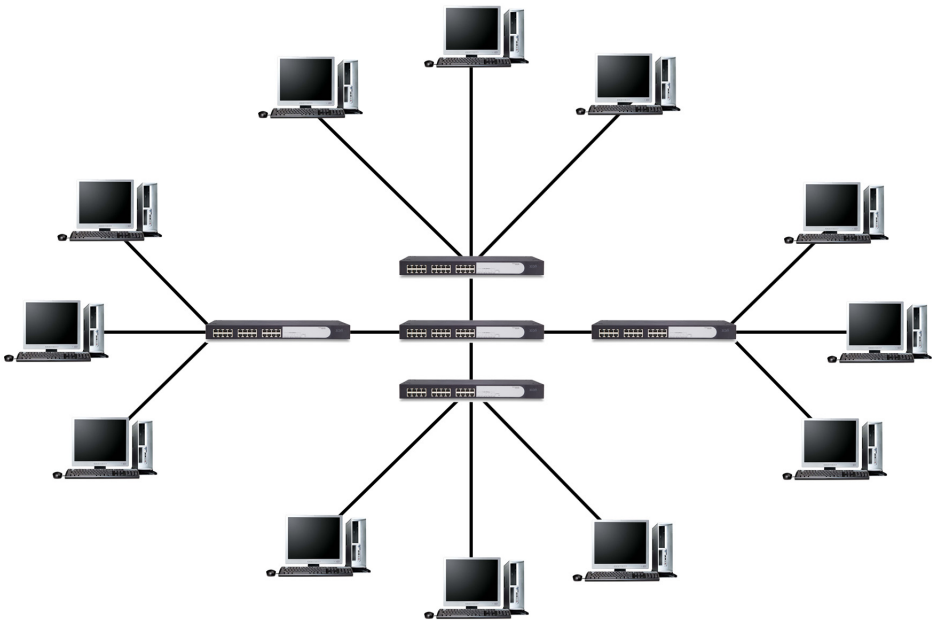


Figura 3.4: Topologia em Estrela Estendida

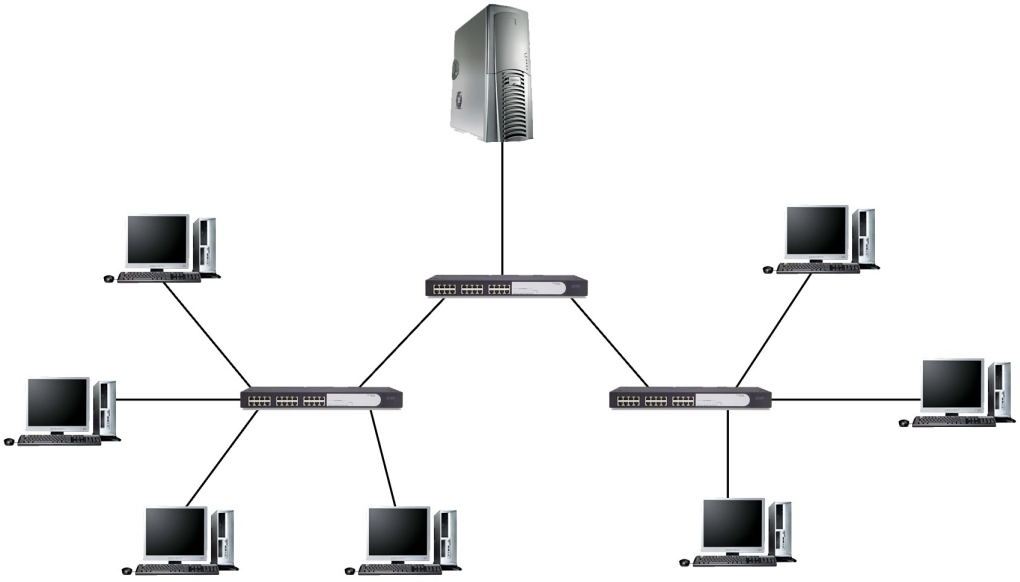


Figura 3.5: Topologia em Hieráquica

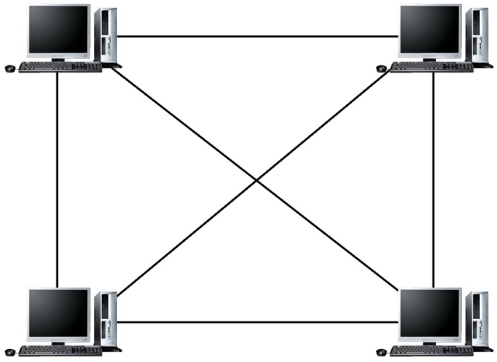


Figura 3.6: Topologia em Malha

A evolução das topologias permitiu a criação de sistemas e redes mais confiáveis e com grande capacidade de transmissão de dados. Claro que a sua complexidade também aumentou. Muitos equipamentos para gerir, coexistência de várias tecnologias e fabricantes. Rede cablada, rede sem fios, ligações redundantes, tornam difícil ter uma visão de toda a rede e das ligações lógicas de todo o equipamento de rede e dos sistemas que os usam. Existem configurações de redes avançadas que permitem a existência de ligações redundantes, que permite ter tolerância a falhas, pois a rede reorganiza-se para manter o seu funcionamento. Um solução deste tipo é o uso do *spanning tree* no *switchs*. Situações deste tipo, implica que a topologia da rede possa mudar de forma dinâmica e autónoma, o que pode dificultar as acções do administrador. Ferramentas de visualização que façam este mapeamento terão grande utilidade para o administrador de redes.

3.3 Segurança

A segurança pode ser vista como o procedimento usado para minimizar as vulnerabilidades dos sistemas de informação. Relaciona a melhor forma de possuir uma rede segura a partir de três conceitos básicos: defesa contra faltas/falhas previsíveis, defesa contra catástrofes e defesa contra actividades não autorizadas. As situações mais comuns para as faltas/falhas previsíveis são [Zúquete, 2008]:

- Falhas de energia;
- Falhas temporarias de conectividade na rede;
- Bloqueio das aplicações.

Para minimizar o impacto destas faltas/falhas destas situações podemos recorrer a sistemas de bateria (UPS), usando caminhos alternativos de rede e recorrendo a um conjuntos de máquina (*clusteres*).

A defesa contra catástrofes tem por objectivo defender o sistema contra situações ou ocorrências que podem pôr em causa a integridade física do mesmo. Podem ser considerados catástrofes os tremores de terra, incêndios, tempestades, ataques terroristas e perda de equipamento. Existem formas de combater estas situações, como por exemplo, usar sistemas de redundância (discos em RAID), fazer cópias de segurança da informação periodicamente e guardá-las em locais diferentes, ou correr o mesmo serviço em máquinas diferentes e em locais distantes um do outro.

A defesa contra actividades não autorizadas permite proteger o sistema de acessos internos e externos à organização. Os acessos externos são normalmente combatidos com recurso a *firewalls*. Os acessos internos são mais difíceis de se combater, pois estes são feitos por pessoas dentro da organização (funcionários), que conhecem o modo de funcionamento da rede e têm acesso físico à informação. Para minimizar estes acessos, normalmente são definidas políticas de segurança muito restritas, definindo explicitamente quem tem acesso ao quê.

A segurança assenta nos seguintes pilares [Junior and Menezes, 2005]:

Confidencialidade – garantir que a protecção da informação seja eficaz proporcionando medidas de controlo de acesso através de autenticação e por criptografia.

Integridade – evitar a corrupção de dados.

Disponibilidade – permite que a informação esteja disponível ao próprio utilizador.

O objectivo da segurança é evitar vulnerabilidades do sistema, sendo estes susceptíveis a ataques [Zúquete, 2008]. Os ataques não são mais que um conjunto de actividades executadas para detectar fraquezas do sistemas. É necessário evitar esse risco ou ameaça que pode resultar em sucesso. A defesa do sistema necessita de medidas que contemplem [Junior and Menezes, 2005]:

- a diminuição das vulnerabilidades;
- detectar e contrariar/anular ataques passados ou futuros;
- minimizar os riscos decorrentes de ataques bem sucedidos.

Com estes três objectivos consegue-se projectar uma correcta defesa. Mas estas medidas necessitam da implementação de políticas e mecanismos de segurança [Zúquete, 2008]. As políticas vão definir requisitos de segurança que permitem ter um sistema seguro para garantir:

- confidencialidade;
- protecção;
- continuidade da prestação de serviços;
- ter capacidade de monitorizar acções passadas.

Os requisitos podem ser avaliados através de auditorias, monitorização e autenticidade dos utilizadores nos serviços. Os mecanismos não passam de um conjunto de máquinas, redes e pessoas que estão sujeitos à mesma política, estes normalmente são compatíveis entre si. Os mecanismos devem ser escolhidos e configurados de forma a não interferir com as respectivas políticas definidas, para não entrar em conflito com o sistema usado pelos utilizadores, e por sua vez conseguirem trabalhar sem anomalias. As políticas, com a ajuda dos mecanismos, geram o **princípio do privilégio mínimo** que consiste atribuir os direitos necessários aos utilizadores para executar as suas tarefas e a nível computacional tem a **capacidade funcional mínima**, em que todos os sistemas são configurados apenas com os componentes necessários e nada mais. Alguns dos princípios mecanismos são:

Norma ISO 17799 – surge baseada na norma britânica BS 177799-1:199 (ABNT ISO 17799, 2005)[ISO, 2005], fornece uma lista de gestão de risco mais pormenorizada, para garantir a segurança da Informação, como principais requisitos temos:

1. planeamento da prestação ininterrupta do sistema – Evita ou restringe as interrupções no fornecimento e prestação de serviços;
2. controlo de acesso – Controlar, impedir o acesso à informação;
3. desenvolvimento e manutenção do sistema – Assegura a segurança do sistema e impede a perda e alteração da informação;
4. segurança ambiental e física – Impede acessos não autorizados, danos e perdas;
5. conformidade – Impede lacunas da lei civil ou criminal, assegurando a conformidade do sistema com políticas e padrões de segurança;
6. segurança dos funcionário – Tem em atenção as acções das pessoas reduzindo o risco de erro, roubo ou fraude;
7. organização da segurança – Gere a segurança da informação interna;
8. gestão de computadores e redes – Assegura a correcta funcionalidade do sistema;

9. classificação e controlo de bens – Criar e manter um inventário detalhado e actualizado;
10. política de segurança – Normas desenvolvidas que têm em conta a responsabilidade, punição e autoridade.

Criptografia – Vem do grego *kryptós*, que significa escondido, oculto mais *graphein* (grafia). Consiste na arte ou ciência de escrever cifras de forma que a troca de informação seja só entendida pelos seus respectivos destinatários, baseia-se em dados a partir de algoritmos e chaves de criptografia [Junior and Menezes, 2005]. Existem dois tipos diferentes de algoritmos: simétricos e assimétricos. O algoritmo simétrico (Figura 3.7) utiliza o mesmo algoritmo para cifrar e decifrar a mesma mensagem. O algoritmo assimétrico (Figura 3.8) usa chaves públicas mas tem um conjunto de duas chaves, uma serve cifrar e a outra para decifrar a mensagem [kurose, 2003].

Figura 3.7: Criptografia Simétrica

Figura 3.8: Criptografia Assimétrica

Firewalls – Apareceram na década de 1990 [Cheswich, 2005], sendo consideradas uma solução eficiente na implementação de políticas de segurança definidas pelo

administrador. O objectivo principal da firewall (Figura 3.9) é regular o tráfego de dados entre máquinas ligadas à mesma rede e controlar a interacção entre as máquinas, impedindo a transmissão e/ou recepção de acessos novos ou não autorizados protegendo o software e o hardware.

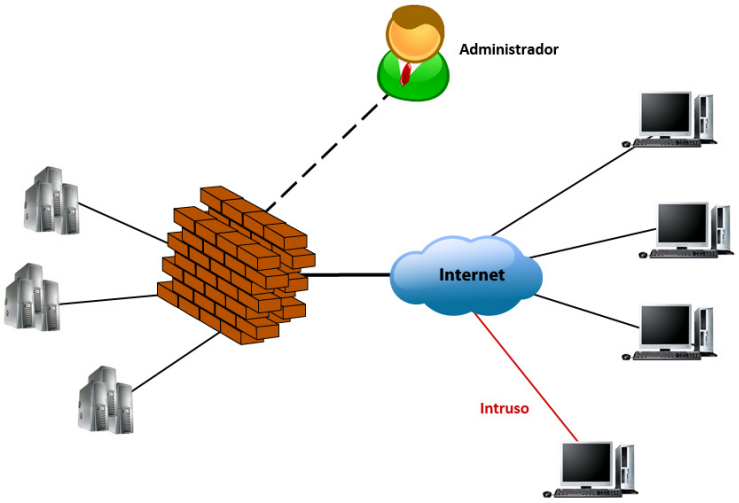


Figura 3.9: Firewall

Podemos destacar alguns tipos de firewall:

- filtro de pacotes – utiliza regras estáticas para filtrar os pacotes que tem origem de servidores externos, sendo simples a sua configuração;
- proxy – filtra os pacotes que são gerados na rede interna que por vezes pode impedir a conexão aos servidores externos que podem ser prejudiciais ao sistema de informação;
- firewall pessoal – é um software que intercepta as conexões de entrada e saída de um computador baseado em regras padrão ou definidas pelo utilizador que define o que pode aceder ou o que deve recusar;
- firewall reactivo – tem como função permitir reconhecer ataques e emitir alertas quando encontra sequências de pacotes IP bloqueando o acesso indevido automaticamente.

Sistema de detecção de Intrusos – IDS (Intrusion Detection Systems) são ferramentas de gestão que permitem detectar e contrariar as intrusões [Viega et al., 2002], pois é feita uma monitorização dos eventos ocorridos [Junior and Menezes, 2005], actuando em conjunto com o sistema de firewall. O IDS permite a verificação

dos conteúdos dos pacotes IP através do sistema de assinaturas. Emitindo alertas caso as assinaturas do IDS não sejam as compativéis com o conteúdo do pacote que chega. Assim, como os sistemas de firewall, o IDS tem diferentes configurações:

- Host-Based Intrusion Detection (HIDS) – monitoriza o sistema com base nos eventos registados no arquivode log;
- Network-Based Inrusion Detectiom (NIDS) – monitoriza o tráfego intermédio de captura dos pacotes IP e da análise do cabeçalho e conteúdo.

Estes mecanismos produzem registos de toda actividade da rede para posteriormente possam ser analisados e monitorizados.

Mais uma vez, a quantidade de informação que é preciso ligar nos registos que são criados pelos mecanismos de segurança é cada vez maior e com necessidade de se relacionar com outros registos.

3.4 Monitorização e registos

O funcionamento de sistemas e redes pressupõe que a topologia, os sistemas operativos e o hardware esteja devidamente ligados e configurados. Em regime estacionário, raramente será necessário proceder a intervenções, principalmente a nível correctivo. No entanto, devido à flutuação de carga, a alterações na topologia ou mesmo as alterações de software, é fundamental proceder à monitorização contínua do estado do equipamento e do funcionamento do sistema como um todo. A monitorização permite, também, adaptar a configuração e a topologia, no sentido de optimizar o funcionamento dos sistemas e das redes.

Com a complexidade das redes, dos sistemas, serviços e aplicações, bem como o seu número a aumentar a necessidade de uma monitorização de todos estes recursos torna-se findamental. Existem várias formas de fazer a monitorização, que serão apresentadas na próxima secção. A monitorização gera também registos que deverão ser estudados para prever falhas e compreender melhor o comportamento dos sistemas e das redes. Estes registos em conjunção com os registos gerados pelo equipamento de rede, sistemas, serviços e aplicações são a informação que o administrador tem que diciminar durante as suas actividades de manutenção e supervisão.

3.4.1 Monitorização

A partir dos anos 70 constatou-se uma convergência entre Tecnologias, Comunicações e Informática, provocando uma revolução visível no sector organizacional.

Os SI permitem às organizações obter uma oferta variada tanto no produto como nos preços, provocando alterações na competitividade. Estes avanços possibilitam um desenvolvimento dos computadores, na medida que passaram a ocupar menos espaço e simultaneamente serem mais velozes, ágeis e com maior capacidade de memória. Com esta evolução surgiu o conceito de gestão de redes, que consiste na monitorização de uma estrutura de recursos físicos e lógicos de uma rede que pode estar geograficamente dispersa ou próxima. A gestão de redes serve para garantir o funcionamento contínuo, assim como assegurar um grau de qualidade os serviços prestados.

A monitorização pode ser ad-hoc, ou seja, equipamento a equipamento, sistema a sistemas, que implica um consulta directa, quer remota, quer por intermédio da utilização de um protocolo de gestão de redes ou mesmo por CLI. Deste tipo de abordagem cada fabricante, tipo de sistemas operativo tem formas e comandos diferentes para verificar o seu estado. Num cenário de heterogeneidade e de grande número de equipamentos e sistemas esta tarefa demonstra-se muito pouco produtiva.

Para colmatar estas falhas surgiram as abordagens integradas. Estas procuram centralizar a informação e os registos de monitorização num sistema central que envia pedidos e recebe as respostas de todos os equipamentos, sistemas e serviços que se desejam monitorizar. Existem diferentes implementações e soluções, mas todas se baseiam neste paradigma de uma estação de gestão central que controla agentes instalados no equipamento de rede para fazer os seus pedidos e receber as suas respostas. Uma das abordagens mais conhecidas é o SNMP.

O SNMP [FSTALLINGS, 1999] foi criado pelo IEFF (Internet Engineering Task Force) em 1988, para monitorizar redes TCP/IP, da camada de aplicação.

Este protocolo permite aos gestores de rede um maior controlo de desempenho e resolver problemas que possam existir ou fornecer informação para um desenvolvimento posterior da rede. A gestão de rede através do SNMP permite um acompanhamento simples, fácil e em tempo real saber o estado da rede. O SNMP tem como princípio a flexibilidade e facilidade de implementação para produtos futuros. Sua especificação está contida em [K. McCloghrie, 1990] e [J. Case, 1990]. Os principais objectivos do protocolo SNMP são:

- reduzir o custo da construção de um agente que suporte o protocolo;
- reduzir o tráfego de mensagens de gestão pela rede necessária para gerir recursos de rede;
- reduzir o número de restrições impostas às ferramentas de gestão de rede devido ao uso de operações complexas e pouco flexíveis;

- apresentar operações simples de serem entendidas, sendo facilmente usadas pelos adinistradores para desenvolver ferramentas de gestão;
- permitir facilmente a introdução de novas características e novos objetos não previstos;
- construir uma arquitectura que seja independente e relevante somente em alguns casos e implementações particulares.

O SNMP esta baseado em três documentos:

- Structure of Management Information (SMI), definido por [M. Rose, 1990], descrevendo, essencialmente, a forma pela qual a informação gerida e definida.
- Management Information Base (MIB). Definida em [K. McCloghrie, 1990], a MIB principal do mundo SNMP (chamada MIB-2) define as variáveis de gestão que todo elemento gerido deve ter, independentemente de sua função particular. Outras MIBs foram posteriormente definidas para fins particulares, tais como MIB de interfaces Ethernet, MIB de nobreaks, MIB de repetidores etc.
- *Simple Network Management Protocol*, definido em [J. Case, 1990], é o protocolo usado entre agente e gestor para a gestão.

As mensagens usadas no protocolo SNMP, para comunicarem com os respectivos intermediários são:

- *get-request-PDU*: mensagem enviada pelo gerente ao agente pedindo o valor de uma variável;
- *get-next-request-PDU*: mensagem utilizada pelo gerente para pedir o valor da próxima variável depois de uma ou mais variáveis que foram especificas;
- *set-request-PDU*: mensagem enviada pelo gerente ao agente para pedir que seja alterado o valor de uma variável;
- *get-response-PDU*: mensagem enviada pelo agente ao gerente, informando o valor de uma variável que lhe foi pedido;
- *trap-PDU*: mensagem enviada pelo agente ao gerente, informando um evento ocorrido.

A estrutura do SNMP é baseada nas seguintes entidades, formado uma aplicação distribuída: o agente, o gestor e o protocolo (Figura 3.10).

Para o modelo funcionar é necessário que a máquina gerida possua um Agente SNMP e uma base de informação de gestão – a MIB (Management Information

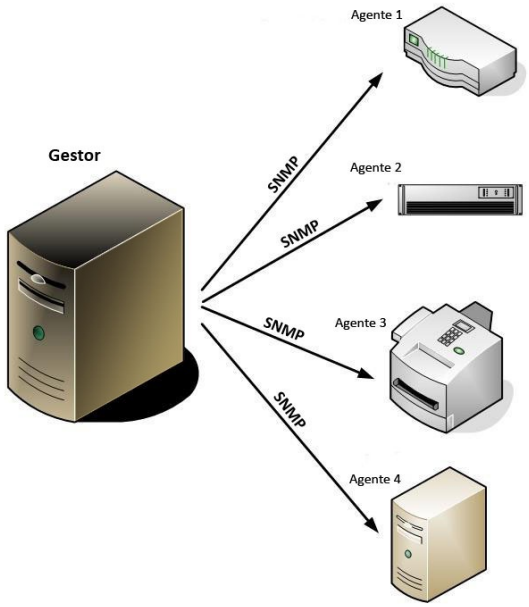


Figura 3.10: Gestão de SNMP

Base), O agente, não é mais que um dispositivo detentor de informação referente ao seu estado actual, permitindo ao gestor consultar e/ou alterar uma parte do sistema. O gestor é o responsável pela monitorização, relatórios e decisões no caso de correr problemas. Enquanto o agente têm a seu cargo o envio das alterações bem como a notificação de ocorrência de eventos programados (Traps).

3.4.2 Registos (*logs*)

Os sistemas informáticos possuem cada vez mais informação, o que provoca uma elevada quantidade de registos das actividades, os chamados *Logs*. Estes surgem com a necessidade de registar todas as actividades/interacções dos dispositivos, desde, SO, aplicações, equipamentos de rede e os respectivos mecanismos de rede. Através deles consegue-se monitorizar todo o tráfego de rede e procurar qualquer anomalia suspeita [Gregio and Santos, 2010]. Os registos normalmente fornecem os seguintes dados:

- Data e hora do evento;
- Hostname;
- O processo e/ou utilizador que gerou o evento;

- A descrição do evento.

Os registos são, normalmente, divididos nos seguintes tipos:

- Registo de Sistemas – guardam todos os registos gerados entre o SO, componentes de hardware e aplicações em execução, tentativas de acesso (local ou remotamente), falhas nos componentes e serviços ou mensagens do sistema que ficam registadas neste tipo de log;
- Registo de programas/serviços – guardam as operações efectuadas pelos programas/ serviços e eventos que acontecem, desde mensagens de início, fim ou reinício de um serviço, acesso a programas e erros ocorridos;
- Registo de sondas/equipamentos de rede – guardam os registos do estado do equipamento monitorizado através de sondas.

Devido à sua diversidade, os registos são considerados um conjunto de registos complexos, difíceis de visualizar e tirar conclusões. A diversidade dos registos, dependem do sistema operativo, dos serviços e das aplicações em execução.

Por exemplo em sistemas Unix é normal encontrar os seguintes registos:

- /var/log/message: Registo Geral do Sistema

```
Jul 17 22:04:25 router  dnsprobe[276]: dns query failed
Jul 17 22:04:29 router last message repeated 2 times
Jul 17 22:04:29 router  dnsprobe[276]: Primary DNS server Is Down...
 Switching To Secondary DNS server
Jul 17 22:05:08 router  dnsprobe[276]: Switching Back To Primary DNS server
Jul 17 22:26:11 debian -- MARK --
Jul 17 22:46:11 debian -- MARK --
Jul 17 22:47:36 router  -- MARK --
Jul 17 22:47:36 router  dnsprobe[276]: dns query failed
Jul 17 22:47:38  debian kernel: rtc: lost some interrupts at 1024Hz.
Jun 17 22:47:39  debian kernel: IN=eth0
OUT= MAC=00:0f:ea:91:04:07:00:08:5c:00:00:01:08:00 SRC=61.4.218.24
DST=192.168.1.100 LEN=60 TOS=0x00 PREC=0x00 TTL=46 ID=21599 DF
PROTO=TCP SPT=59297 DPT=22 WINDOW=5840 RES=0x00 SYN URGP=0
```

- /var/log/auth.log: Registos de autenticação

```
Jan 24 16:40:11 combo sshd[32564]: Failed password for root
from ::ffff:66.199.253.10 port 38579 ssh2
```

```
Jan 26 11:46:22 combo sshd[10375]: Accepted password for
root from ::ffff:63.126.79.65 port 32768 ssh2
Feb 26 01:42:59 combo xinetd[2013]: START: ftp pid=6533
from=83.27.214.35
```

- /var/log/kern.log: Registos do kernel

```
Mar 16 11:29:06 phantom kernel:
[   11.607336] ACPI: Power Button [PWRF]
Mar 16 11:29:06 phantom kernel:
[   11.611608] ACPI: acpi_idle registered with cpuidle
Mar 16 11:29:06 phantom kernel:
[   12.468131] parport_pc 00:09: reported by Plug and Play ACPI
Mar 16 11:29:06 phantom kernel:
[   12.468161] parport0: PC-style at 0x378, irq
7 [PCSPP,TRISTATE,EPP]
```

- /var/log/cron.log: Registos do serviço de agendamento

```
Jan 22 08:00:00 combo CROND[25108]:
(root) CMD (/usr/bin/mrtg /etc/mrtg/mrtg.cfg)
Jan 22 08:00:00 combo CROND[25110]:
(mailman) CMD (/usr/bin/python -S /var/mailman/cron/gate_news)
Jan 22 08:00:00 combo CROND[25113]:
(mailman) CMD (/usr/bin/python -S /var/mailman/cron/checkdbs)
Jan 22 08:01:00 combo CROND[25115]:
(root) CMD (run-parts /etc/cron.hourly)
Jan 22 08:05:00 combo CROND[25129]:
(root) CMD (/usr/bin/mrtg /etc/mrtg/mrtg.cfg)
Jan 22 08:05:00 combo CROND[25131]:
(mailman) CMD (/usr/bin/python -S /var/mailman/cron/gate_news)
Jan 22 08:10:00 combo CROND[25133]:
(root) CMD (/usr/lib/sa/sa1 1 1)
Jan 22 08:10:00 combo CROND[25136]:
(root) CMD (/usr/bin/mrtg /etc/mrtg/mrtg.cfg)
```

- /var/log/boot.log : Registos do arranque do sistema

```
Jan 26 12:22:06 combo syslog: syslogd startup succeeded
```

```
Jan 26 12:22:06 combo syslog: klogd startup succeeded
Jan 26 12:22:06 combo irqbalance: irqbalance startup succeeded
Jan 26 12:22:07 combo portmap: portmap startup succeeded
Jan 26 12:22:07 combo nfslock: rpc.statd startup succeeded
Jan 26 12:22:08 combo rpcidmapd: rpc.idmapd startup succeeded
Jan 26 12:22:03 combo network: Setting network parameters:  succeeded
Jan 26 12:22:08 combo random: Initializing random number generator:  succeede
Jan 26 12:22:03 combo network: Bringing up loopback interface:  succeeded
Jan 26 12:22:08 combo rc: Starting pcmcia:  succeeded
Jan 26 12:22:08 combo bluetooth: hcid startup succeeded
Jan 26 12:22:09 combo bluetooth: sdpd startup succeeded
Jan 26 12:22:09 combo netfs: Mounting other filesystems:  succeeded
Jan 26 12:22:09 combo apmd: apmd startup succeeded
Jan 26 12:22:10 combo autofs: automount startup succeeded
```

- Por cada serviço em execução é normal ter mais ficheiros de registo

Nos sistemas Windows os registos mais comuns são:

- Application Log Records – eventos registados por aplicações
- Directory Service Records – eventos registados pelo Active Directory e serviços complementares
- DNS Server Records – consultas DNS, respostas, e outras actividades de DNS
- File Replication Service Records – eventos de actividades de replicação de ficheiros
- Security Log Records – eventos que foram marcados para auditoria nas políticas locais ou de grupos
- System Log Records – eventos registados pelo sistemas operativo e os seus componentes

Estes registos podem ser exportados para ficheiros CSV, com as seguintes informações: Data, Hora, Origem, Tipo, Categoria, Evento, Utilizador, Descrição.

Exemplo:

```
9/7/99,9:43:24 PM,DNS,Information,None,2,N/A,ZETA,The DNS
Server has started.
9/7/99,9:40:04 PM,DNS,Error,None,4015,N/A,ZETA,The DNS
server has encountered a critical
error from the Directory Service (DS). The data is the error code.
```

Como se pode constatar existe uma grande variedade de registos com formatos diferentes e dependentes do sistema operativo. Assim sendo a aplicação de técnicas de visualização podem, quando correctamente utilizadas, ajudar os administradores de sistemas a interpretar a informação.

Capítulo 4

Visualização de Sistemas e Redes

O objectivo final da visualização de sistemas e redes é conseguir interpretar os dados resultantes da instrumentação do equipamento e aplicações de forma a conseguir uma análise eficaz e rápida, que permita optimizar o processo de tomada de decisão. Assim, a resposta a potenciais situações indesejadas será mais eficaz e permitirá dar resposta de forma adequada no sentido de manter os sistemas e redes em boa condição de funcionamento.

O administrador é um elemento fundamental para o processo de visualização e para o desenvolvimento do SI, permitindo que este se torne seguro e eficaz. Tal como descrito anteriormente, o procedimento de visualização inclui um workflow bem definido. O primeiro passo consiste na aquisição de dados, seguido de processamento e construção de estruturas de dados adequadas. Após a definição do mapeamento de elementos gráficos, é, finalmente, possível construir as vistas necessárias.

Neste contexto, é importante que a técnica de visualização mostre rapidamente os dados visualizados, se possível dispensando tarefas morosas de programação e parametrização. O acesso a um conjunto diversificado de dados exige um tipo de estrutura que garanta a leitura em qualquer formato de ficheiros provenientes das mais diversas fontes. Em suma, a visualização de sistemas e redes é um processo complexo que requer técnicas eficientes para a exploração os dados. O objectivo é procurar simplificar o desenvolvimento das organizações na área da segurança e dos SI.

4.1 Procedimento de visualização

O ponto de partida inclui a aquisição de dados, que, tipicamente, resultam da instrumentação das aplicações e do equipamento de rede. A grande diversidade de dados leva a inúmeros registos que vão desde os registos de sistemas (*logs*), de serviços e de

sondas/equipamentos de rede, que exportam os dados por intermédio de protocolos de gestão de redes (SNMP, por exemplo).

4.1.1 Dados

Os dados nem sempre estão imediatamente disponíveis ao utilizador sendo necessário adquiri-los para posteriormente serem interpretados. As fontes de dados são várias e o seu formato diverso. Várias aplicações geram dados sob a forma de registos (logs) que podem apresentar vários formatos: XML, JSON, CSV, TSV, entre outros. Outras possibilidades incluem dados obtidos da instrumentação do equipamento de rede, tipicamente obtido por intermédio de um protocolo de gestão de redes (SNMP ou outro). Adicionalmente, também é comum obter dados directamente de páginas Web ou do resultado de procedimentos de captura de pacotes.

Os dados apresentam, tipicamente, uma marca temporal, sendo, em sistemas correctamente configurados, sincronizados por intermédio de NTP. As informações obtidas fornecem dados, tais como, conexões externas, registos de utilização de serviços (arquivos transferidos via FTP, acessos a páginas web, tentativas de login sem sucesso, avisos de discos cheios, entre outros). Para obter estes registos é necessário configurar o sistema e possuir software específico.

Os logs facilitam o acompanhamento do que acontece na rede e no próprio sistema, e é necessário que sejam monitorizados com frequência para permitir a resolução rápida de eventuais problemas (Figura 4.1). Como mostra o exemplo, através dos LOGs consegue-se obter a seguinte informação: data e hora do acesso, o hostname, o que gerou a mensagem e a mensagem. Como se pode constatar à primeira vista não é fácil e requer técnicas que ajudem a filtrar essa informação de forma a ser visualizada rapidamente.

66.249.66.235 - - [26/Jul/2011:06:29:42 +0100] "GET /~ /index.html HTTP/1.1" 200 10073 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.235 - - [26/Jul/2011:06:49:02 +0100] "GET /~ /joomla/index.php?option=com_content&view=category&id=34%3Acontactos &Itemid=57&layout=default&format=feed&type=atom HTTP/1.1" 200 29266 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.235 - - [26/Jul/2011:07:07:01 +0100] "GET /~ / HTTP/1.1" 304 - "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"

Figura 4.1: Exemplo de um ficheiro de log

O objectivo final é, após obter os dados, construir visualizações que permitam construir um conhecimento mais adequado do estado dos sistemas e da rede. Como passo prévio à construção das vistas, é necessário processar os dados, quer para filtragem, quer para estruturação. Uma das técnicas mais comuns inclui a interpretação (*parsing*) que é um processo que analisa uma sequência de dados de entrada para determinar a sua estrutura gramatical [Niels Willems and van Wijk, 2003].

Como consequência, os dados são organizados segundo uma estrutura mais adequada à construção de vistas. A sua representação visual é importante para manter a coerência e a interacção dos dados mudando apenas o visual, como a côr, forma e os próprios recursos de visualização.

4.1.2 Visualização

O procedimento de visualização pressupõe a definição do problema concreto que se pretende resolver para, posteriormente, proceder à aquisição dos dados, o seu processamento e estruturação e, finalmente, visualização.

No caso concreto dos sistemas e redes, há vários problemas que podem beneficiar da aplicação de técnicas de visualização, que abrangem desde a gestão corrente do funcionamento destes à detecção de eventos ou problemas. Estão já disponíveis várias ferramentas e aplicações que podem ser aplicadas directamente ou adaptadas para resolver estes problemas. O âmbito é tão abrangente que uma cobertura exaustiva é irrealista. Optou-se por identificar alguns problemas típicos mais comuns, descrevendo-se o procedimento na secção seguinte.

4.2 Cenários de aplicação

As redes de computadores sofreram mudanças para dar resposta à evolução das aplicações, que deixaram de ser sistemas isolados para se tornarem em sistemas abertos e distribuídos. Para acompanhar esta evolução, apostou-se na segurança da informação, que passou a ser encarada como uma gestão inteligente da informação priorizando recursos nesse sentido [Pinheiro, 2007].

Os sistemas abertos e distribuídos estão sujeitos a inúmeros cenários, desde a sua gestão ao controlo das ameaças acidentais ou intencionais. Assim, é necessário abordar alguns cenários como: conhecimento da topologia de rede, capacidade de resposta dos routers, controlo de acontecimentos, identificação de ocorrências, localizar a nível geográfico os utilizadores, os host, carga de rede nos equipamentos e segurança da rede (Tabela 4.1).

Possíveis Cenários
Como visualizar e manter o conhecimento actualizado de topologia de rede?
Em que medida a temperatura contribui para falhas nos servidores em *data centers* e como se relaciona com a altura do ano?
Como verificar se a rede instalada tem capacidade para lidar com a evolução da tecnologia e consequente aumento de utilização?
Como saber, de forma instantânea, o que está a acontecer na rede em termos de comunicação? Por outras palavras, quem está a comunicar com quem, com que frequência, se o responsável pelo tráfego é uma máquina e se algum nó está a ser mal comportado?
Como identificar a existência de network clustering de forma a melhor contribuir para a melhor concepção da rede? Por outras palavras, a existência de agrupamentos (clusters) de rede pode resultar na sobrecarga de troços de rede, com a consequente perda de pacotes e degradação de QoS. O conhecimento de clusters bem identificados na rede pode possibilitar a optimização e adaptação da rede de forma a conseguir lidar com estas situações.
Como saber, em tempo real, a localização geográfica dos utilizadores que estão, em cada momento, a aceder aos serviços electrónicos da empresa?
Como ter, em tempo real, conhecimento preciso das ligações feitas em determinados hosts, agrupando informação como origem, destino, porta e tipo de protocolo?
Como monitorizar, em tempo real, a carga de rede nos equipamentos de infra-estrutura?
Como saber, se o serviço de DNS está em ruptura, numa perspectiva de segurança?

Tabela 4.1: Exemplo de cenários de aplicação

4.2.1 Topologia de Rede

O conhecimento da topologia de rede é determinante nas diversas tarefas de gestão, tais como na relação de eventos ou a localização da origem de um problema. Devido à complexidade das redes de hoje, é difícil manter actualizado o mapa de forma manual pois é uma tarefa morosa, trabalhosa, tendo de ser repetida regularmente devido às alterações sofridas na rede.

As redes são geridas de uma forma descentralizada e por vários gestores locais e nem sempre estão informados, em tempo útil, de todas as alterações. O conhecimento da topologia de rede serve para dar ao administrador uma visão clara de toda a rede e os seus elementos, assim como toda a informação existentes nos equipamentos.

Problema

Como visualizar e manter actualizado o conhecimento de topologia de rede?

Dados

A topologia física da rede recorre à caracterização de relações físicas da conectividade que existe entre as várias entidades da rede de comunicação de dados. O conhecimento do layout físico e as interconexões dos elementos da rede é considerado como pré-requisito para muitas tarefas críticas, tais como a monitorização de falhas e análise de aplicações em execução.

A descoberta da topologia de rede passa, essencialmente, por identificar os equipamentos de infra-estrutura, nomeadamente, os routers e os switches. Enquanto que os routers providenciam informação sobre os endereços de nível 3 (IP), os switches possuem informação de encaminhamento de nível 2 (MAC). Os dados encontram-se disponíveis como resultado da instrumentação dos dispositivos, acessíveis por intermédio do protocolo SNMP. Em suma, a descoberta da topologia incide, essencialmente, na obtenção e análise das tabelas de encaminhamento (AFT) dos switches e o refinamento das ligações intermédias por intermédio de um proceso interactivo.

Os dispositivos de rede podem ser classificados como identifcados e não identificados. Os nós identificados são representados pelos switches que têm endereço IP e dos quais pode ser retirada toda a informação existente na tabelas AFT. Os nodos não identificados representam, por exemplo, hubs ou elementos que não suportam SNMP. As tabelas AFT possuem um mecanismo que limpa os endereços MAC, evitando que estes estejam sempre em cache, conseguindo assim uma visão correcta e actualiza dos vizinhos e da própria rede. Esta informação resulta do Spannig Tree Protocol, que vai capturar o caminho mais eficiente para os elementos que pertencem à rede.Os computadores, que têm ligação aos switchs, possuem endereço MAC que é analisado, tendo ocupado uma posição na AFT da respectiva porta. Assim, para detectar os computadores que estão ligados basta verificar as AFT que obtenham um único e simples endereço MAC.

Processamento

Alguns sistemas usam o método de traçagem para saber onde se encontra o equipamento e também das sub-redes [Breitbart, 2004]. O algoritmo de descoberta assenta em três princípios: cada domínio conter exactamente uma sub-rede, as vlans estarem presentes nos domínios de administração e as tabelas de encaminhamento conterem endereços IP completos. Começa por descobrir o conjunto de routers existentes no domínio e os seus vizinhos (Algoritmo 1).

De entre os IPs decobertos, são identificados os switchs em primeiro lugar, para cada router e para cada interface, e a sub-rede está directamente ligado ou não, ou

Algoritmo 1 Descoberta de routers de um domínio

```
1: procedure FindLeadConections(S1,S1,..., Sn)  ▷ S1,S2, ..., Sn são nós de uma subrede N
2:     Current ← {S1, S2, ... Sn}
3:     while Current is not empty do
4:         find AFT Aij with minimal number of entries
5:         if Aij = {St} then
6:             create a connection between Si and St
7:             eliminate node St from all remaining AFTs
8:             continue
9:         else
10:            Aij ← {St1, ..., Stk}
11:            create a hub and use it to connect all Si, St1, ..., Stk
12:            eliminate St1, ..., Stk from all remaining AFTs
13:        end if
14:    end while
15: end procedure
```

seja, o conjunto de endereços IP é calculado e associado na sub-rede correspondente na interface. Esta numeração tem em conta as máscaras de sub-rede e os formatos do endereço. Para cada endereço IP usa o MIB para obter o endereço MAC que permite saber qual a sub-rede, o nome do sistema e o número da portas. Depois, de calculado é determinado para cada endereço, o nó a que corresponde o switch (Algoritmo 2).

Visualização

A visualização de topologias passa pela representação de um grafo. Um grafo é uma abstracção matemática que relaciona vértices (V) por intermédio de arestas (E). Nos grafos não direccionados os vérices representam os elementos da rede (switchs), e as arestas que designam as respectivas ligações, tendo a notação de $G = (V, E)$, em que cada nodo será o vértice e as arestas representam as ligações [Breitbart, 2004] (Figura 4.2).

4.2.2 Efeito da temperatura ambiente na disponibilidade

Muitos edifícios não se encontram devidamente preparados para o aumento de carga consequente do incremento dos serviços electrónicos actuais, principalmente, ao nível da climatização nos *data centers*. Estes problemas podem originar falhas intermitentes ou lentidão na disponibilidade dos serviços fornecidos aos utilizadores. Para estes, bem como para eventuais utilizadores externos, este tipo de anomalias dificulta o desenvolvimento do seu trabalho, bem como o acesso a alguns serviços.

Algoritmo 2 Descoberta de switch e routers de um domínio

```
1: procedure FindInterConections(S_1,S_1,..., S_n, R_1, R_2, ..., R_m)   ▷ S_1,S_2, ..., S_n são switches de uma subrede S   ▷ R_1,R_2, ..., R_m são routers de uma subrede S
2:     Current ← {S_1, S_2, ... S_n}
3:     for switch S_i do
4:         for interface j of S_i do
5:             if S_ij has already been matched then
6:                 continue
7:             else if A_ij ⋃ A_kl = μ and A_ij ⋂ A_kl = φ then
8:                 match S_ij with S_kl
9:             end if
10:        end for
11:    end for
12:    for router R_k, switch S_i do
13:        for interface j of S_i do
14:            if S_ij is not matched and A_ij contains R_k then
15:                continue
16:            else if A_ij ⋃ A_kl = μ and A_ij ⋂ A_kl = φ then
17:                match S_ij with R_k
18:            end if
19:        end for
20:    end for
21: end procedure
```

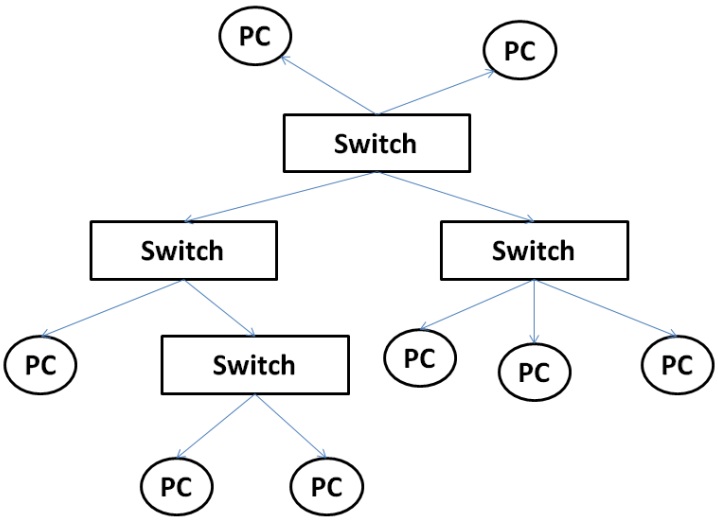


Figura 4.2: Representação gráfica de um domínio.

São exigidos aos *data centers* confiabilidade devido à grande quantidade de informação que armazenamento e é inaceitável que haja indisponibilidade destes. Por isso, é necessário combater os problemas associados com este sistemas, como a temperatura, problemas energético e erros humanos.

Problema

Em que medida a temperatura contribui para falhas nos servidores em *data centers* e como se relaciona com a altura do ano?

Dados

A má climatização da sala pode levar a que certas áreas sejam mais refrigeradas que outras, afectando, de forma imprevisível, o desempenho dos servidores. Neste sentido, para poder aferir a influência da temperatura na disponibilidade de serviços é necessário proceder ao registo e interpretação da temperatura bem como o número, duração e localização de falhas.

A temperatura é medida por intermédio de termómetros, quer isolados, quer em rede, e a contabilização de falhas é feita, normalmente, de forma manual. Estes valores devem ser registados em base de dados para poderem ser, posteriormente, processados.

Processamento

Os dados adquiridos deverão ser estruturados de forma tabular, contendo os seguintes campos: máquina, timestamp, duração, temperatura. O campo máquina deve estar associado a uma noção de localização, correspondendo à temperatura amostrada nesse instante (timestamp). A duração deve resumir o tempo em que durou a falha.

Visualização

A visualização pode ser feita por intermédio de um gráfico do tipo timeseries, em que um dos eixos representa a noção de tempo e o outro a temperatura. Associado ao ponto, deve estar o número de falhas (Figura 4.3).

4.2.3 Capacidade de processamento de routers

Os routers permitem que os computadores comuniquem e passem informação entre duas redes. Estes normalmente fornecem segurança incorporada, tal como numa firewall que actuam na camada 3 do Modelo OSI, tendo como principal missão

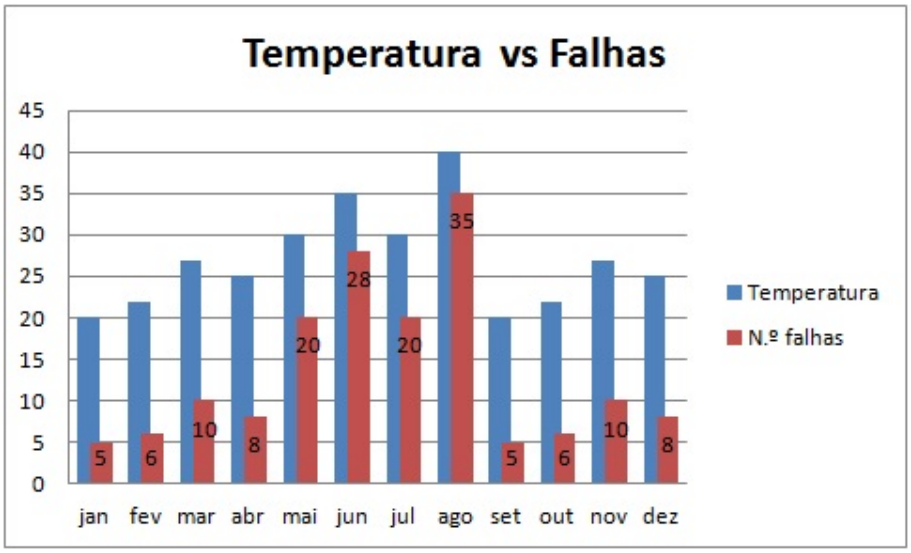


Figura 4.3: Temperatura vs Falhas.

seleccionar o melhor encaminhamento de pacotes para o seu destino final. Com a evolução da tecnologia e o aumento da utilização é necessário verificar se a rede têm capacidade de resposta para lidar com esta situação.

Problema

Como verificar se a rede instalada tem capacidade para lidar com a evolução da tecnologia e consequente aumento de utilização?

Dados

A capacidade de encaminhamento de pacotes depende, em grande medida, da capacidade de processamento dos routers, pois todos os pacotes são analizados e o melhor caminho calculado. Quanto mais cálculos forem necessários, maior é a utilização média do processador, podendo haver problemas de rede no caso de o processador apresentar uma carga média elevada.

A informação é registada pelo sistema operativo do router como uma forma de instrumentação e, tipicamente, está acessível através de SNMP. A informação pode ser retirada a partir do objecto 1.3.6.1.2.1.25.3.3.1.2 (hrProcessorTable da HOST-RESOURCES-MIB) e representa a média da percentagem de tempo durante o último minuto em que o processador esteve activo.

Processamento

O valor obtido deve ser registado num formato tabular, com os seguintes campos: timestamp, percentagem. Desta forma é possível obter o histórico da ocupação do processador, bem como o instante de tempo em que tipicamente acontece maior sobrecarga.

Visualização

Este tipo de dados é facilmente visualizado através de gráficos do tipo *timeseries* que mostra o comportamento ao longo do dia da carga de processador (Figura 4.4).

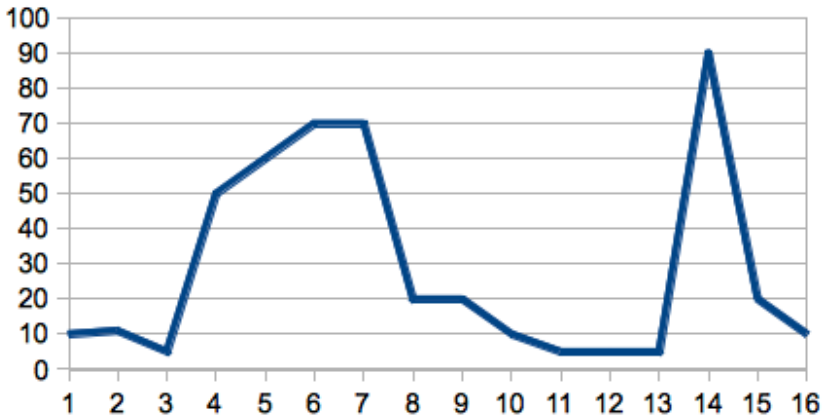


Figura 4.4: Carga média do CPU de um router

4.2.4 Controlo de acontecimentos na rede

As TI são responsáveis pela maioria dos pacotes de informação que circulam e pelo número de serviços imprescindíveis para o funcionamento da organização. As redes de comunicações são o pilar de suporte destes serviços, sendo vital garantir o seu funcionamento contínuo e adequado às necessidades. Por outras palavras, é necessário controlar tudo o que acontece na rede em termos de comunicação, obtendo informação de quem está a comunicar com quem, com que frequência, se o responsável pelo tráfego é apenas uma máquina e se algum deles está a ser mal comportado.

Problema

Como saber, de forma instantânea, o que está a acontecer na rede em termos de comunicação? Por outras palavras, quem está a comunicar com quem, com que frequência, se o responsável pelo tráfego é uma máquina e se algum nó está a ser mal comportado?

Dados

Os dados necessários para atacar este problema resultam da instrumentação da comunicação de dados. Na sua base, a comunicação é efectuada por intermédio da troca de pacotes (tipicamente IP) entre máquinas, sendo cada um deles encaminhado com o auxílio de equipamento específico de rede.

Há diversas aplicações que podem ser utilizadas para adquirir estes dados como, por exemplo, *tcpdump*[1] ou *Wireshark*[2]. Estas ferramentas escutam em determinada interface de rede e representam o tráfego capturado no ecran.

Os pacotes capturados têm a vantagem de conter toda a informação, incluindo os cabeçalhos e rodapés associados às diversas camadas da pilha conceptual de comunicação. No entanto, não há informação directa sobre os fluxos, informação pertinente para análise do estado da rede e da comunicação em si.

Processamento

Na camada acima da captura de pacotes é possível encontrar informação sobre o fluxo de dados. Esta informação encontra-se, tipicamente, nos routers. Por vezes, encontram-se computadores com estas funções, o que permite obter informação de fluxos directamente.

Adicionalmente, há protocolos especificados para a obtenção de fluxos a partir do equipamento de rede como, por exemplo, o IPFIX [Claise, 2008]. A visualização de fluxos pode ser uma ferramenta importantíssima na análise e interpretação de condições de operação da rede.

Visualização

A visualização de fluxos pode passar por recorrer a técnicas de visualização de grafos. A ferramenta Argus permite gerar relatórios detalhados dos fluxos que circulam na rede de forma instantânea. O Argus pode ser executado em sistema de auditoria,

[1] www.tcpdump.org
[2] www.wireshark.org

ou todos os tráfegos de rede que o host gera e recebe, ou funcionar como um sonda autónoma.

Pode fornecer controlo de acesso baseado em instalações de conecção usando a tecnologia TCP-wrapper (Figura 4.5). Para visualizar estes dados também é possível recorrer ao NVisionIP (V) que representa os dados atráves de grafos de dispersão 2D.

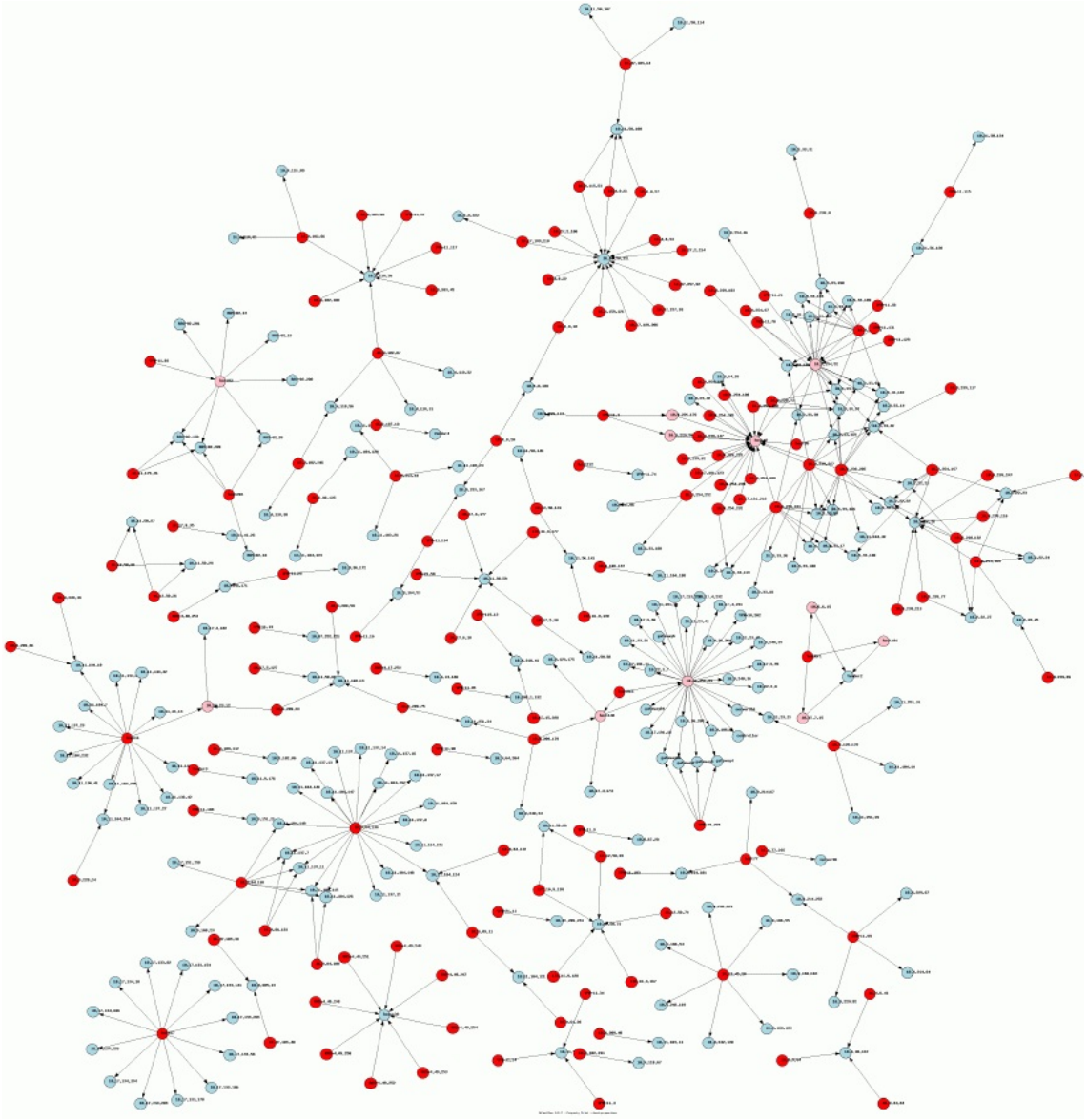


Figura 4.5: Grafo direcionado de sessões SSH intranet como registo por Argus

4.2.5 Identificação de desiquilíbrios de rede

A elevada informação que circula na rede pode provocar desiquilíbrios nos troços de rede. É necessário saber como os diagnosticar bem como identificar formas de os solucionar. O uso de clusters permite usar técnicas de agrupamento de dados e visualizar numa única interface.

Problema

Como identificar a existência de network clustering de forma a melhor contribuir para a melhor concepção da rede? Por outras palavras, a existência de agrupamentos (clusters) de rede pode resultar na sobrecarga de troços de rede, com a consequente perda de pacotes e degradação de QoS. O conhecimento de clusters bem identificados na rede pode possibilitar a optimização e adaptação da rede de forma a conseguir lidar com estas situações.

Dados

O processo de apoio às decisões tem como base a monitorização e integridade da fonte de dados, fazendo que seja importante obter as ocorrências de rede e analiza-las de forma rapida e completa. A ferramenta Cytoscape é um software gratuito, usado para visualizar as interacções de rede de forma a mostrá-las atráves de grafos.

Processamento

Para efectuar o processo dos dados podemos usar os clusters de filtragem ou de rede. Os clusters de filtragem são ferramentas que servem para "afinar" depois de um algoritmo de agrupamento. Em geral, os filtros servem para examinar os resultados do próprio algoritmo de agrupamento, com base em métricas, alterando arestas ou adicionando nós aumentado a qualidade do cluster. Os clusters de rede têm como objectivo detectar agrupamentos naturais de nós de rede. Estes são geralmente definidos como atributos de vantagem numérica que contém alguma semelhança ou distância métrica entre dois nós. os nós que são mais semelhantes (ao aproximar) são mais propensos a serem agrupados. Para identifir estas situações podemos recorrer à ferramenta de visualização Cytoscape, que permite exibir a rede em 3 partes: criar nova rede de atributos, criar nova rede de cluster e ciar nova rede com redes alinhadas de atributos.

Visualização

A ferramenta Cytocape suporta a visualização de redes padrão e dados de diversos formatos: SIF, GML, XGMML, etc,(Figura 4.6).

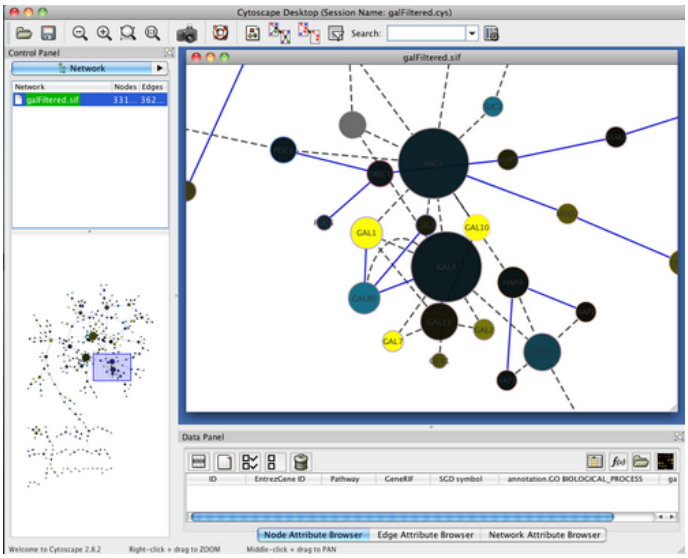


Figura 4.6: Rede criada por um clusterMaker 's.

4.2.6 Identificação da localização geográfica dos utilizadores

O conhecimento geográfico pode contribuir de forma assinalável para a prestação de serviços electrónicos, desde a oferta de publicidade específica, entrega de conteúdo adaptado (internacionalizável), cálculo de DRM (Digital Rights Management), entre outos. Pemite em tempo real, saber a localização geográfica exacta dos utilizadores que estão, em qualquer momento a aceder aos serviços da empresa.

Problema

Como saber, em tempo real, a localização geográfica dos utilizadores que estão, em cada momento, a aceder aos serviços electrónicos da empresa?

Dados

A localização geográfica dos utilizadores é elaborada com base nos dados que contém a sua geolocalização IP. Os dados recolhidos são inseridos numa BD, estes são partilhadas com o mesmo formato e APIs. Pois, o uso de base de dados binárias com APIs, é mais vantajoso que importar ficheiros CSV em SQL, pois o formato binário

é mais eficiente e é fácil de configurar e usar. As APIs são optimizadas para oferecer uma melhor velocidade, uso de memória e tamanho das BD. Assim, com localização geográfica dos utilizadores obtém-se através de dados recolhidos no acesso à rede e aos sites visitados.

Processamento

A aplicação GeoIP permite que os administradores de uma forma não evasiva solicitar diversas informações geográficas sobre os utilizadores e visitantes da Internet em tempo real. Por outras palavras, região, cidade, código postal, país, latitude/longitude, ISP, nome da empresa, nome do domínio e se o endereço IP é de um proxy anónimo ou de um ISP. O GeoIP é simples, mas é um processo complexo pois introduzem os dados dos utilizadores e em seguida é gerado uma serie de algoritmos que identificam, extraem e extropolam os pontos de localização de endereços IP.

Visualização

Com a aplicação GeoIP consegue-se visualizar os dados da localização actual (Figura 4.7).

GeoIP for your IP address

This is a comprehensive demo returning all of the data available from our various GeoIP databases. To obtain all the data returned below, you would need to purchase the GeoIP City, ISP, Organization, Domain Name, and Netspeed databases.

Your IP Address	85.138.50.219
Countries	Portugal
Region	05 (Braganca)
US Area Code	
US Metro Code	
Global Cities	Mirandela
US Zipcode*	
Latitude/Longitude	41.4833/-7.1833
ISP	ZON TV Cabo
Organization	ZON TV Cabo
Netspeed	Cable/DSL
Domain Name	netcabo.pt

Figura 4.7: Localização actual

4.2.7 Identificação de análises à rede

Um ataque a uma rede é precedida, geralmente, por uma análise exaustiva à rede, no sentido de averiguar que máquinas, sistemas operativos e respectivas versões e aplicações estão aí instaladas (network scan). Com este conhecimento, um atacante pode construir um caminho para a descoberta e exploração de vulnerabilidades, conseguindo, eventualmente, acesso à rede.

Adicionalmente, máquinas infectadas com vírus ou worms e actividade ilícita na rede pode ser detectada por intermédio da assinatura de rede, ou seja, de acordo com o padrão de transmissão de pacotes. A detecção rápida e atempada deste tipo de situações pode tornar a rede mais segura e providenciar um tempo de recuperação mais rápido [Irwin and van Riel, 2007].

Problema

Como ter, em tempo real, conhecimento preciso das ligações feitas em determinados hosts, agrupando informação como origem, destino, porta e tipo de protocolo?

Dados

Os dados necessários para resolver este problema resultam da auscultação dos pacotes que circulam na rede. São obtidos por intermédio de sondas, estrategicamente distribuídas, e agregadas de forma a conseguir um registo completo. A informação necessária inclui a origem, destino, porta e tipo de protocolo. O objectivo do administrador é monitorizar os padrões de transmissão que estão a circular na sua rede.

Processamento

Essencialmente, o processamento dos dados é feito de forma a fazer o levantamento dos seguintes tipos de padrões:

- *port-scan*: uma única fonte tenta várias ligações a um único destino. Este padrão constitui uma descoberta de serviços (ou descoberta vertical).
- *port-sweep*: uma única fonte tenta fazer várias ligações a vários destinos. Também é conhecido por *host discovery*, *vulnerability scanning* ou descoberta horizontal.
- *distributed scanning*: várias fontes (coordenadas) tentam fazer ligações a um ou mais destinos de forma a evitar serem descobertas.

Visualização

Uma forma relativamente fácil, é usar a ferramenta InetVis, que permite visualizar em 3-D o tráfego de rede através de gráficos de dispersão.

É útil para observar em tempo real a actividade de verificação e outros padrões de táfego anómalo. Esta ferramenta permite fazer a seguinte leitura (Figura 4.8):

- Eixo x (horizontal - azul) – Endereço de destino;
- Eixo z (profundidade - vermelho) – Endereço de origem;
- Eixo y (verde) – Portas TCP/UDP;
- Tráfego ICMP.

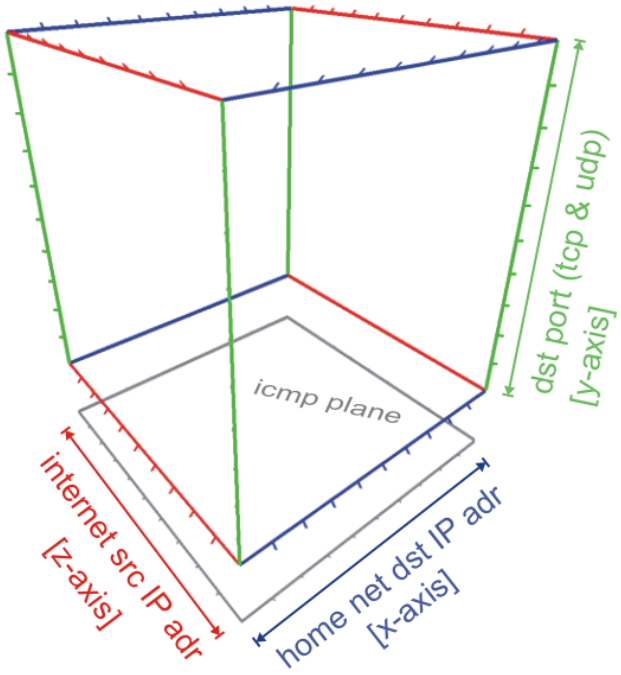


Figura 4.8: InetVis

4.2.8 Monitorização da carga de rede nos equipamentos

A transmissão de pacotes é efectuada por intermédio de routers, localizados em pontos estratégicos na rede e, como tal, constituíndo a infra-estrutura activa. Tipicamente, estes routers encontram-se na porta de ligação da rede interna com a rede

externa, pelo que é de suma importância a monitorização do tráfego, em termos de volume, através de diversos períodos de tempo.

Problema

Como monitorizar, em tempo real, a carga de rede nos equipamentos de infra-estrutura?

Dados

Os dados necessários são obtidos por intermédio de SNMP, através de consulta às MIBs de cada router. Em concreto, são vários os objectos de gestão envolvidos. No entanto, os mais significativos são ifInOctets, ifOutOctets, ifInUcastOctets, ifOutUcastOctets, entre outros.

Estes dados são acumulados localmente para ser possível medir o tráfego em qualquer momento, incluíndo um histórico detalhado.

Processamento

O processamento necessário inclui a redução e cálculo de médias em determinados instantes de tempo. Adicionalmente, para distinguir o tráfego de acordo com a sua natureza é necessário proceder à definição de uma codificação cromática:

- verde – tráfego de entrada em bits por segundo;
- azul – tráfego de saída em bits por segundo;
- verde escuro – tráfego máximo de entrada em cinco minutos;
- Rosa – tráfego de saída máximo em cinco minutos.

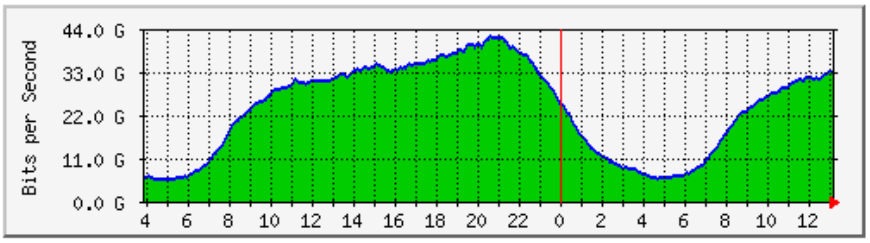
Visualização

Os gráficos, do tipo timeseries, permite alguma variabilidade em termos de dimensão temporal. Por exemplo, consegue-se obter a análise diária (em média 5 minutos Figura 4.9), semanal (em média 30 minutos Figura 4.10), mensal (em média 2 horas Figura 4.11), anual (em média 1 dia Figura 4.12).

4.2.9 Segurança para DNS

A internet é considerada uma rede de computadores interligados entre si, fornecendo informação constante, resilente e rápida aos seus utlizadores. O administrador de

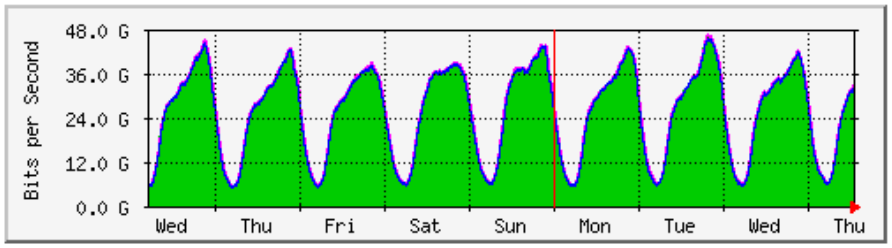
Gráfico Diário (5 minutos Média)



	Max	Média	Atual
Em	41,8 Gb / s	23,7 Gb / s	32,5 Gb / s
Fora	41,8 Gb / s	23,7 Gb / s	32,5 Gb / s

Figura 4.9: Gráfico Diário

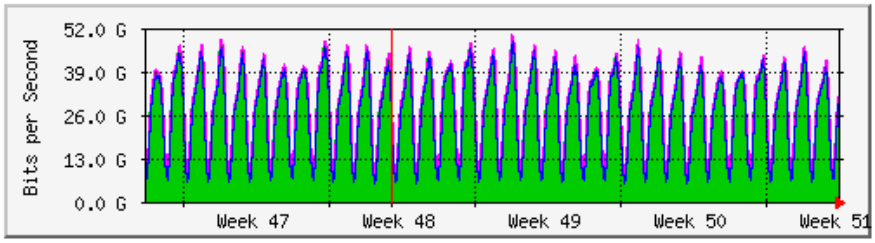
Gráfico Semanal (30 minutos Média)



	Max	Média	Atual
Em	46,2 Gb / s	25,2 Gb / s	32,2 Gb / s
Fora	46,1 Gb / s	25,2 Gb / s	32,2 Gb / s

Figura 4.10: Gráfico Semanal

Gráfico Mensal (2 horas Média)



	Max	Média	Atual
Em	49,7 Gb / s	26,2 Gb / s	31,2 Gb / s
Fora	49,6 Gb / s	26,1 Gb / s	31,2 Gb / s

Figura 4.11: Gráfico Mensal

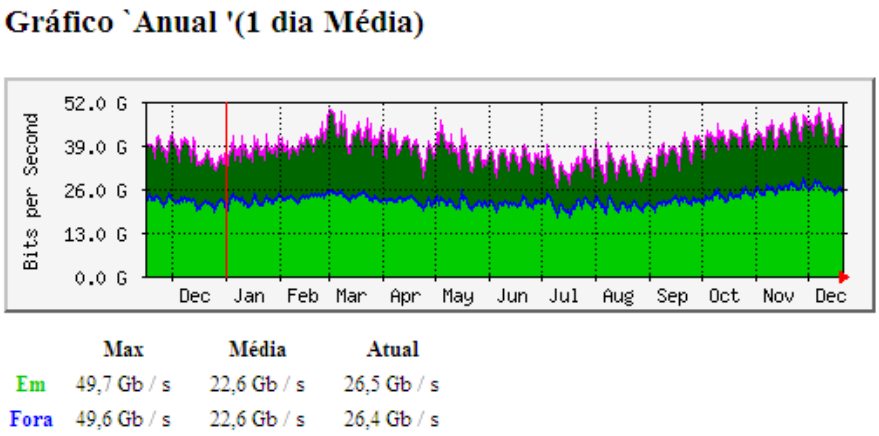


Figura 4.12: Gráfico Anual

sistemas tem de precaver actividades maliciosas. O protocolo de DNS permite dar resposta a um elevado número de sistemas que se encontram ligados. É um sistema maleável, flexível e robusto, em tudo semelhante aos restantes protocolos, contudo vulnerável no que diz respeito à resistência dos ataques e garantia da integridade dos dados. Surge a nível de segurança o DNSSEC (Domain Name System Security Extension) com o objectivo de combater estas falhas.

Problema

Como saber, se o serviço de DNS está em ruptura, numa perspectiva de segurança?

Dados

Para a segurança do DNS é necessário filtrar e analisar o seu tráfego, mas é quase impossível de fazer a separação de tráfego malicioso através de detecção e filtragem de trafego. É necessário certificar todas as transações, e para isso usamos o DNSSEC que se baseia na criptografia assimetrica.

Processamento

Um ataque ao serviço de DNS pode originar a sua ruptura. É enviado um elevado número de perguntas recursivas, a um ou mais servidores de DNS, provocando esgotamento de alguns dos seus recurso, por exemplo, largura de banda disponível, memória ou capacidade de processamento no servidor. Este esgotamento levará à degradação do serviço em si. Para evitar situações destas, o administrador terá de monitorizar todas as tentativas de intrusão, e analisá-las. Como, é uma tarefa

demasiado morosa e com elevado número de informação para analisar, necessita de recorrer a ferramentas, que em tempo real, analisem o tráfego e associar ao serviço de DNS para detectar padrões de desvio.

Visualização

A monitorização pode ser realizada atráves de algumas ferramanta, como por exempolo, SNORT, que permite em tempo real, detectar alterações no comportamento do tráfego e do próprio sistema (Figura 4.13)

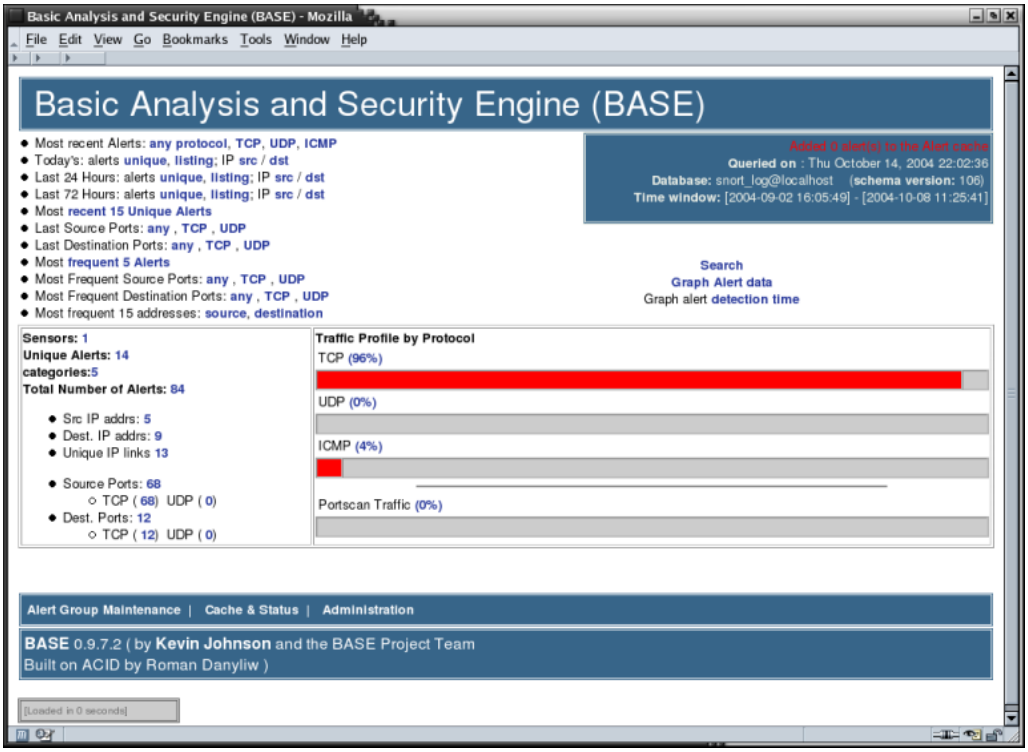


Figura 4.13: Snort

4.3 Conclusão

A tarefas dos administradores de sistemas não é fácil, devido à grande quantidade de técnicas e abordagens que estão envolvidas. Adicionalmente, há a necessidade de construir uma visão ampla de todo o sistema e rede para o manter seguro. Para isso recorrem aos registos armazenados (logs), dados de instrumentação e outros. Toda esta informação é, tipicamente, extensa e difícil de interpretar em estado bruto. Logo, é necessário recorrer a ferramentas externas, para ajudar a interpretar os

Possíveis Cenários	**Dados**	**Protocolo**	**Visualização**
Topologia de Rede	Tabelas AFT dos switches	SNMP	Grafo
Efeito da temperatura ambiente na disponibilidade	Termómetros ou sensores de temperatura	Manual	Scatterplot
Capacidade de processamento de routers	Carga média de CPU do Router	SNMP	Timeseries
Controlo de acontecimentos na rede	Tráfego na rede	SNMP	Grafos
Identificação de desiquilíbrios de rede	Tabela de interfaces	SNMP	Grafos
Identificação da localização geográfica dos utilizadores			
Identificação de análises à rede	Endereços e portas	Sondas de tráfego	Scatterplot
Monitorização da carga de rede nos equipamentos	Tabela de interfaces	SNMP	Timeseries
Segurança para DNS	Registo	Logs	Tabular

Tabela 4.2: Exemplo de cenários de aplicação

registos. Estes são processados, filtrados e visualizados. As ferramentas de visualização permitem tirar proveito de paradigmas visuais, facilitando a interpretação dos dados.

Como exemplo, pode-se destarcar o DAVIX – Análise de Dados e Visualização Linux, pois consiste num conjunto de ferramentas de visualização agrupadas num live cd. É fácil de usar, intuitivo e personalizavel, tendo como caracteristas gerais ser *out-of-the-box*, pois suporta hardware para placas gráficas e de rede, ser modular, o que facilita a sua personalização. O DAVIX inclui um manual de utilização bastante acessível aos seus utilizadores.

Foram apresentados e discutidos alguns cenários de aplicação, que se encontram resumidos na Tabela 4.2.

Capítulo 5

Conclusões

Este trabalho teve como principal objectivo estudar diversas ferramentas que possam contribuir para um diagnóstico correcto e adequado sobre determinadas situações de funcionamento de sistemas e de redes, possibilitando ao administrador uma rápida e adequada tomada de decisão.

Verifica-se, no dia-a-dia, que estamos rodeados de enúmera informação, sendo necessário tomar decisões por vezes sem ter uma ideia clara da situação em causa. Logo, é com a ajuda da visão que conseguimos compreender tudo o que nos rodeia. Através desta capacidade, conseguimos distinguir formas, padrões, cores, etc.. Assim, surge a possibilidade de utilizar visualização associada aos sistemas e redes através da utilização de programas que permitem analisar e processar dados de uma forma rápida e apresentá-los em imagens de mais fácil compreensão.

As organizações dependem cada vez mais das tecnologias de informação, provocando não só um aumento da sua utilização, mas também um aumento da complexidade destes, tornando fundamental a sua monitorização e registo para solucionar eventuais problemas de forma rápida e adequada.

A visualização de sistemas e redes resulta na interpretação dos dados resultantes dos equipamentos e aplicações de forma a garantir uma análise eficiente e correcta.

No estudo realizado foram abordados alguns cenários possíveis para aplicação de várias técnicas de vizualização. Para cada cenário foi descrito o problema, como efectuar a aquisição dos dados, o seu processamento e finalizando com uma possível visualização.

A tarefa de análise dos sistemas por parte dos administradores não é fácil, devido à enorme quantidade de informação que circula numa rede e devido ao facto de ao mesmo tempo a manter segura. Verifica-se, por outro lado, a existência de ferramentas de visualização que podem ajudar recorrendo a diversos paradigmas visuais.

Bibliografia

[sec, 2011] (2011). SecViz | security visualization. http://www.secviz.org/.

[Andery, 2010] Andery, G. d. F. (2010). Integrando projeções multidimensionais à analise visual de redes sociais. http://www.teses.usp.br/teses/disponiveis/55/55134/tde-06102010-111345/fr.php.

[Breitbart, 2004] Breitbart, Yuri; Garofalkis, M. J. B. M. C. R. R. S. A. (2004). Topology discovery in heterogeneous ip networks: The netinventory system(Objecto application/pdf).

[Card et al., 1999] Card, S. K., Mackinlay, J., and Shneiderman, B., editors (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1 edition.

[Carvalho and Marcos, 2009] Carvalho, E. and Marcos, A. (2009). Visualização de informação.

[Cerqueira, 2010] Cerqueira, R. (2010). Monitoramento online e visualização da informação. http://www.slideshare.net/renatacbc/monitoramento-online-e-visualizao-da-informao.

[Cheswich, 2005] Cheswich, w., B. S. H. A. D. (2005). Firewalls e sgurança na internet.

[Claise, 2008] Claise, B. (2008). Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information. RFC 5101 (Proposed Standard).

[Cleveland, 2007] Cleveland, William S.; Diaconis, P. M. R. (2007). scatter82.pdf (Objecto application/pdf).

[Ehlers, 2009] Ehlers, R. S. (2009). Análise de séries temporais.pdf (Objecto application/pdf).

[Fry, 2008] Fry, B. (2008). Visualizing data.

[FSTALLINGS, 1999] FSTALLINGS, W. (1999). Snmp, snmpv2, snmpv3 & rmon 1 & 2 (3rd ed.).

[Gregio and Santos, 2010] Gregio, A. R. a. and Santos, R. (2010). RCTI: análise e visualização de logs de segurança. http://repositorio.cti.gov.br/repositorio/handle/10691/153.

[Halsall, 1995] Halsall, F. (1995). *Data communications, computer networks and open systems*. Addison Wesley Longman Publishing Co., Inc.

[Hamann, 2011] Hamann, R. (2011). A evolução dos computadores. http://www.tecmundo.com.br/9421-a-evolucao-dos-computadores.htm.

[Hessen, 2000] Hessen, j. (2000). *Teoria do Conhcimento*. São Paulo_ Martins Fontes.

[Irwin and van Riel, 2007] Irwin, B. and van Riel, J. (2007). Inetvis: a graphical aid for the detection and visualisation of network scans. In *Conference on Vizualisation Security (VizSec2007)*.

[ISO, 2005] ISO, A. (2005). Iec 17799: 2005: Tecnologia da informação–técnicas de segurança–código de prática para a gestão da segurança da informação. *Rio de Janeiro, Associação Brasileira de Normas Técnicas*.

[J. Case, 1990] J. Case, M. Fedor, M. S. J. D. (1990). RFC 1157 - simple network management protocol (SNMP) (RFC1157). http://www.faqs.org/rfcs/rfc1157.html.

[Junior and Menezes, 2005] Junior, Raimundo Viegas;Martins, A. d. O. and Menezes, C. A. T. (2005). gcete2005.pdf (Objecto application/pdf).

[K. McCloghrie, 1990] K. McCloghrie, M. R. (1990). RFC 1156 - management information base for network management of (RFC1156). http://www.faqs.org/rfcs/rfc1156.html.

[Knight, 2000] Knight, C. (2000). 31b.pdf (Objecto application/pdf).

[kurose, 2003] kurose, J. F.; Ross, K. w. (2003). Redes de computadores.

[M. Rose, 1990] M. Rose, K. M. (1990). RFC 1155 - structure and identification of management information for tcp/ip-based internets. https://datatracker.ietf.org/doc/rfc1155/.

[Marty, 2008] Marty, R. (2008). Applied security visualization.

[Mazza, 2009] Mazza, R. (2009). *Introduction to Information Visualization*. Springer, 1st edition. edition.

[Nascimento, 2005] Nascimento, Hugo A. D. do e Ferreira, C. B. R. (2005). arq0285.pdf (Objecto application/pdf).

[Niels Willems and van Wijk, 2003] Niels Willems, H. v. d. W. and van Wijk, J. J. (2003). Visualization of vessel movements. http://www.visualcomplexity.com/vc/search.cfm?input=security.

[office, 2010] office, M. (2010). Apresentar os dados em um gráfico de dispersão ou de linhas - outlook - office.com. http://office.microsoft.com/pt-br/outlook-help/apresentar-os-dados-em-um-grafico-de-dispersao-ou-de-linhas-HA010227478.aspx.

[Page, 1994] Page, R. D. M. (1994). Tree map 1.o - user's guide. Division of Environmental and Evolutionary Biology Institute of Biomedical and Life Sciences University of Glasgow.

[Paulovich, 2009] Paulovich, F. V. M. R. (2009). Mapeamento de dados multidimensionais – integrando mineracao e visualizacao.

[Pinheiro, 2007] Pinheiro, J. M. d. S. (2007). Ameaças e ataques aos siatemas de informação-11.pdf (Objecto application/pdf).

[Silva, 2007] Silva, F. C. d. (2007). Unidimensional (x) bidimensionais (x,y) tridimensionais (x (2). http://www.slidefinder.net/U/Unidimensional_Bidimensionais_Tridimensionais_Multidimension

[Tanenbaum, 2002] Tanenbaum, A. (2002). *Computer networks*. Prentice Hall Professional Technical Reference.

[Targino, 1995] Targino, M. d. G. (1995). v1n1_1995_2.pdf (Objecto application/pdf).

[Teyseyre and Campo, 2009] Teyseyre, A. and Campo, M. (2009). An overview of 3d software visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 15(1):87–105.

[Van Ham et al., 2009] Van Ham, F., Wattenberg, M., and Viegas, F. B. (2009). Mapping text with phrase nets. *IEEE Transactions on Visualization and Computer Graphics*, 15:1169–1176.

[Viega et al., 2002] Viega, J., McGraw, G., and Online, S. T. B. (2002). *Building secure software: how to avoid security problems the right way*, volume 2. Addison-Wesley New York.

[Wikipédia,] Wikipédia, a. e. l. Teoria dos grafos – wikipédia, a enciclopédia livre. http://pt.wikipedia.org/wiki/Grafos#Defini.C3.A7.C3.B5es_de_grafos_e_digrafos.

[Zúquete, 2008] Zúquete, A. (Fevereiro, 2008). Segurança em redes informáticas.