

A INTERNET PARA PRINCIPIANTES

JOSÉ MANUEL PIRES MARQUES¹

FERNANDO COSTA PARENTE²

Departamento de Física da Faculdade de Ciências da Universidade de Lisboa
Campo Grande, Edifício C1, 4.º Piso, 1700 LISBOA

Neste artigo, são apresentados e definidos os conceitos fundamentais para a compreensão da Internet por parte daqueles que querem utilizar as extraordinárias possibilidades desta rede mas não podem perder muito tempo a apreender tudo o que ela envolve.

1. Introdução

A Internet (de *Interconnected Networks* ou *Internetwork System*) é uma rede global constituída por um conjunto de sub-redes, com milhões de computadores ligados entre si e que hoje se designa, em geral, simplesmente por *rede*. As suas origens remontam a 1969 com a criação por parte do Departamento de Defesa dos Estados Unidos de uma rede de computadores, a ARPANET, com a finalidade de facilitar a troca de informações militares em caso de ataque nuclear. Foi com base nessa rede, que em 1984 a *National Science Foundation* (NSF) criou a NSFNET para ligar instituições de ensino e investigação científica. Sendo a sua utilização financiada pelo governo americano, o uso desta rede por parte de entidades privadas esteve sempre bastante limitado. A partir de 1990 foram levantadas parte das restrições ao seu uso por entidades privadas bem como teve lugar nos EUA a aprovação de um conjunto de medidas tendentes à criação das chamadas auto-estradas da informação *HPCA High Performance Computing Act* (1991) e *NIIAA - National Information Infrastructure Agenda for Action* (1993) - o que provocou o crescimento exponencial na utilização da rede, que ainda hoje se verifica.

Ao contrário do que acontece em alguns sistemas *on-line* (em tempo real) em serviço há já algum tempo, onde a informação reside num grande computa-

dor, como acontece por exemplo na *Compuserve*, *America Online* (EUA) ou nas muitas e populares *BBS* (*Bulletin Board System*) existentes, na Internet a informação encontra-se distribuída por milhares de computadores em todo o mundo. Além da informação, também as comunicações funcionam de um modo distribuído, isto é existe mais do que um modo de ligação de um ponto (*nó da rede*) a outro. Tecnicamente, cada mensagem é dividida em várias partes ou pacotes (*packets*) para facilitar a comunicação, seguindo cada um o caminho mais conveniente. No nó de destino, os diferentes pacotes são de novo reunidos pela ordem original. Assim, é aumentada a eficiência da transmissão e, no caso de parte da rede deixar de funcionar, os pacotes podem atingir o destino através dos nós que se mantenham em funcionamento.

Actualmente a Internet é muito mais do que uma rede utilizada essencialmente por instituições de investigação científica, universidades e organismos governamentais ligados sobretudo à defesa. A Internet tornou-se num emblema cultural representativo daquilo que o futuro representa já hoje e do que é possível acontecer quando a utilizadores comuns é dada a capacidade de comunicar entre si sem limitação. Caracterizada como super auto-estrada da informação constitui, sem

¹ ze@alf1.cii.fc.ul.pt

² parente@alf1.cii.fc.ul.pt

**Conceitos básicos
sobre redes**

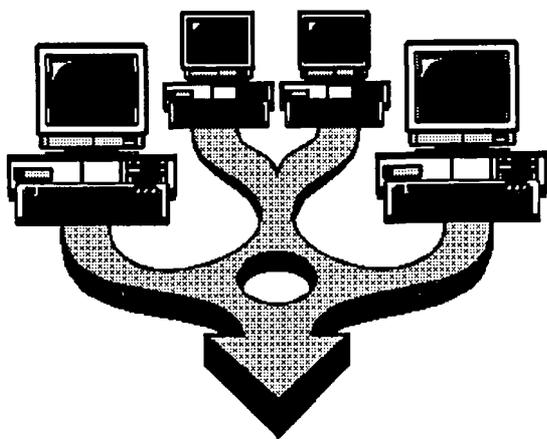
Ligação à Internet

**O modelo
Cliente/Servidor**

Serviços da Internet

dúvida, uma das melhores razões para se ter um micro-computador em casa.

O que é que as pessoas fazem então na Internet? Trocam mensagens por correio electrónico, navegam no chamado *ciberespaço*³, obtendo informações sobre praticamente qualquer assunto, consultam horários de aviões, compram bilhetes para espectáculos, encomendam livros, trocam programas entre si, etc. No entanto, o crescimento acelerado do tráfego na REDE, a sua importância estratégica no desenvolvimento técnico-científico bem como o seu potencial económico, levaram os EUA a criarem recentemente algumas medidas tendentes ao controlo do seu desenvolvimento e exploração comercial. Algumas destas medidas têm gerado bastante controvérsia. Estão em causa sobretudo implicações económicas e sociais. Há quem receie o potencial de vigilância e controlo destas tecnologias, essencialmente no que se refere à vigilância electrónica por parte dos estados sobre o conteúdo da informação que circula na rede e a capacidade de fichagem individual dos cidadãos. Há no entanto que salientar o indelével potencial democrático da REDE na concretização da ideia de **Aldeia Global** no que diz respeito à difusão da informação. Outro ponto polémico é a comercialização da REDE, que pode vir a limitar o acesso a cidadãos e organismos com menos recursos económicos.



A Internet está já hoje a alterar o nosso comportamento. Vemos cada vez mais nos meios de comunicação social referências a endereços electrónicos e páginas WWW, já para não falar nos casos de pessoas que se conheceram, namoraram e cujos casamentos aí começaram. Num futuro próximo, o acesso à REDE será tão banal como é hoje o acesso à rede telefónica. Os sectores mais tradicionais das actividades económicas começaram já a aperceber-se das potencialidades deste mercado emergente. Assim, estão já a surgir no mercado computadores mais baratos e de utilização simplificada (*Net PC*) e televisores capazes de navegar na Internet (*Net TV*).

Em Portugal o acesso à Internet começou, tal como na generalidade dos países, na década de 80. De início possibilitava apenas acesso remoto a computadores de universidades através da rede telefónica. A instalação de um nó EARN (*European Academic and Research Network*) em 1986 na Faculdade de Ciências da Universidade de Lisboa e a criação nesse mesmo ano da FCCN (*Fundação para a Computação Científica Nacional*) deu origem à instalação da primeira Rede da Comunidade Científica Nacional (RCCN). Inicialmente circunscrita à comunidade académica e científica nacional, tem vindo a generalizar-se essencialmente devido ao aparecimento no mercado de diversos fornecedores comerciais de acesso através da rede telefónica (*Service provider*). A ligação típica é efectuada através de uma linha telefónica normal, dedicada ou não, ou por uma linha digital RDIS (*Rede Digital com Integração de Serviços*). A diminuição dos custos do equipamento informático aliada à facilidade de utilização dos computadores actuais tornou bastante acessível a ligação individual à Internet.

2. Conceitos básicos sobre redes

Uma rede é constituída por um conjunto de computadores ligados entre si por um canal de comunicação. Uma rede que ligue computadores de uma dada Instituição designa-se por Rede Local (LAN — *Local Area Network*). Caso os computadores se encontrem distribuídos por uma área maior (cidade, país, etc.), denomina-se Rede Alargada (WAN — *Wide Area Network*).

2.1 Protocolos de comunicação

Na Internet estão interligados vários tipos de redes, quer quanto à cobertura geográfica quer quanto à tecnologia utilizada. Para que seja possível comunicar entre elas é necessário que existam regras comuns — protocolos — que regulem a forma como os pacotes de dados são organizados e enviados entre computadores. A Internet é baseada num conjunto ou família de protocolos designados por TCP/IP (*Transmission Control Protocol/Internet Protocol*).

2.2 Endereçamento

Cada nó da rede tem um endereço IP constituído por uma cadeia numérica de 4 bytes⁴ (32 bits) separados

³ Do título do livro de ficção científica de William Gibson "Cyberspace". Neste livro o ciberespaço surge como um ambiente integrado onde computadores e pessoas vivem e trabalham num espaço virtual — *Matrix* — para elas tomado como real.

⁴ 1 byte representa um conjunto de 8 bits (1bit, de *binary digit*, representa um dos 2 símbolos possíveis, 0 ou 1, do sistema binário). Com 1 byte é possível representar 256 números (0 a 255).

entre si por pontos (ex. 199.119.9.1). Este endereço identifica univocamente o computador (*host*) na rede. Com 4 bytes o número máximo de endereços disponíveis é de $2^{32} = 4\,294\,967\,296$. No entanto, para tornar as comunicações mais simples nem todos os endereços são utilizados. No IP a cada rede é atribuído um número de rede. A cada computador (*host*) dessa rede corresponde um nó identificado também por um número. Assim, cada endereço IP é constituído por duas partes: rede.nó. Existem várias formas de divisão dos quatro bytes do endereço entre o número da rede e o número do nó, dependendo da dimensão da rede.

2.3 Encaminhamento

Como vimos, uma mensagem é constituída por vários pacotes podendo cada um deles seguir o seu próprio caminho. Nesta acepção, a Internet é uma rede de comutação de pacotes de informação (*packet switching network*). Este modelo de funcionamento exige que cada pacote tenha embebida a sua própria informação de endereçamento. Exige também a existência de sistemas - *routers* - capazes de ler esta informação e enviar os pacotes para os respectivos destinos. Até uma mensagem atingir o seu destino passa por vários *routers*.

2.4 Nomes e Domínios

A identificação de computadores através de cadeias de números torna difícil a sua memorização. A solução encontrada foi a de utilizar nomes em vez de números. Tal como no endereço IP, os nomes encontram-se separados por pontos (ex. comp.cc.fc.ul.pt), mas agora sem qualquer limitação quanto à quantidade de nomes e pontos utilizados. Claro que a cada nome deve corresponder um endereço IP único. A ideia geral é a de que o utilizador não deve ter necessidade de conhecer o número de um determinado computador com o qual pretende comunicar, mas apenas o seu nome. Para isso torna-se necessária a existência algures de uma base de dados que associe, de forma transparente, a cada nome um número antes de enviar pacotes para a rede. Tal sistema denomina-se por Sistema de Nomes do Domínio - DNS (*Domain Name System*). Este sistema encontra-se, regra geral, distribuído por vários computadores, designando-se cada um deles por Servidor de Nomes do Domínio (*Domain Name Server*) ou simplesmente Servidor de Nomes. Um domínio identifica, regra geral, uma dada instituição ficando a primeira posição, em geral, disponível para o nome do computador em causa. No exemplo apresentado, o computador comp insere-se no domínio cc.fc.ul.pt sendo este domínio comum a todos os computadores nele inseridos e designando-se por *FQDN* (*Fully Qualified Domain Name*).

O domínio encontra-se organizado hierarquicamente. Por exemplo o domínio cc.fc.ul.pt encontra-se atribuído ao Centro de Cálculo (.cc) da Faculdade de Ciências (.fc) da Universidade de Lisboa (.ul), Portugal (.pt). O código do país designa-se por domínio geral geográfico (*top-level geographical domain*) ou simplesmente domínio geral. Eis os códigos de alguns países:

- .pt - Portugal
- .uk - Reino Unido
- .fr - França
- .at - Áustria
- .gr - Grécia
- .ca - Canadá
- .ch - Suíça
- .jp - Japão
- .de - Alemanha
- .us - EUA

O código .us é raramente utilizado e a razão deve-se ao facto de inicialmente existirem apenas seis domínios gerais, todos eles referentes a *sites* nos EUA e designados por domínios gerais por tipo de organização (*top-level organizational domains*):

- .gov - Entidades governamentais (ex. nasa.gov)
- .com - Organizações comerciais (ex. ibm.com)
- .edu - Instituições de ensino (ex. uoregon.edu)
- .mil - Instituições militares (ex. arpa.mil)
- .org - Instituições sem fins lucrativos (ex. aps.org)
- .net - Instituições ligadas à administração da rede (ex. internic.net)

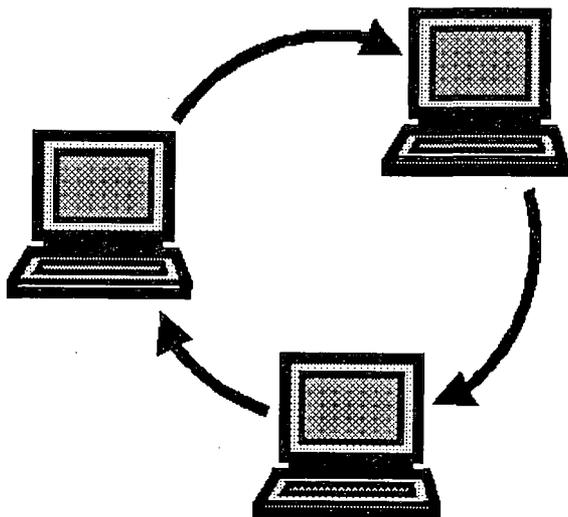
tendo os códigos dos outros países surgido mais tarde. Os domínios gerais .mil, .gov e .edu encontram-se reservados para uso exclusivo dos EUA. Os outros três apesar de representarem, em geral, instituições americanas, podem ser utilizados por organizações de outros países. O que é importante garantir é não existirem repetições, pelo que todos os nomes devem ser registados junto da InterNIC (*Internet Network Information Center*). Por norma as organizações pedem nomes que facilitem a sua identificação, tais como microsoft.com ou whitehouse.gov. Note-se, que nos exemplos anteriores não é indicado nenhum computador em concreto. Já no exemplo comp.cc.fc.ul.pt, o computador comp pertence ao Centro de Cálculo (cc) da Faculdade de Ciências (fc) da Universidade de Lisboa (ul) em Portugal (pt).

2.5 Endereços electrónicos pessoais

Os utilizadores têm endereços electrónicos de correio (e-mail) pessoais com a forma
id_utilizador@endereço_computador.

O nome ou id do utilizador é obtido junto do fornecedor de serviço e vem antes do símbolo @ (lê-se at). Por exemplo o utilizador joaquim que utiliza a Telepac como ISP terá como e-mail

joaquim@mail.telepac.pt.



3. A ligação à Internet

Para aceder à Internet necessitamos de um computador (geralmente um PC ou um Macintosh) e de um meio de ligação a um dos vários Fornecedores de Acesso existentes no mercado. A ligação pode ser efectuada através de uma linha telefónica normal, dedicada ou não, ou por uma linha digital RDIS — ISDN (*Integrated Services Digital Network*). Para isso necessitamos de:

- um *modem* (*modulator* — *demodulator*), caso o acesso seja efectuado através das linhas telefónicas normais. O que o modem faz é converter os dados digitais do computador em dados analógicos capazes de serem transmitidos através destas linhas telefónicas e, converter para digital os dados recebidos delas. Os modems podem ser internos (colocados no interior do computador) ou externos (ligados a este através da porta série). Quanto à velocidade de transmissão-recepção mede-se em bits por segundo (bps — *bits per second*) ou kilobits por segundo (kbps ou, simplesmente, k). As velocidades máximas mais correntes são 2400, 9600, 14400 (14.4 k), 28800 (28.8 k). Os modems de 56 k acabam de chegar ao mercado na altura em que este artigo está a ser escrito. Não confundir no entanto *bps* com *bauds*. O *baud* representa a taxa de variação ou de transições entre estados⁵ de um sinal num canal de comunicação. Um *baud* representa uma dessas variações. Dependendo do tipo de modulação, um *modem* pode emitir um ou mais *bits* por cada *baud*;

- uma placa *ethernet*, colocada no interior do computador, caso o acesso seja efectuado através de uma rede local (LAN) por um cabo próprio.

Além do *hardware*, necessitamos também de programas, *software*, que nos permitam aceder à rede. A maior parte destes programas é fornecida com o sistema operativo utilizado (Windows95/98, WindowsNT, MacOS, Unix).

A configuração específica do *software*, ou seja as opções concretas que têm de ser indicadas para que a ligação funcione depende obviamente do tipo de ligação e dos serviços prestados pelo Fornecedor de Serviço.

No caso da ligação se efectuar através da linha telefónica (*dial-up connection*), as principais opções habitualmente disponíveis são os protocolos SLIP (*Serial Line Internet Protocol*) e PPP (*Point-to-Point Protocol*) que fazem parte da família de protocolos TCP/IP. A ligação via PPP é actualmente a mais utilizada visto ser mais avançada e estável que o SLIP. Dependendo do método utilizado pelo Fornecedor de Serviço, podemos ter ou não necessidade de conhecer o endereço IP do computador através do qual acedemos à rede:

- Endereçamento estático (PPP estático) — a cada utilizador é atribuído um número IP permanente.
- Endereçamento dinâmico (PPP dinâmico) — cada vez que o utilizador se liga é-lhe atribuído um número IP, que se mantém enquanto durar a ligação. Esta é a forma mais vulgar de ligação.
- Endereçamento emulado (pseudo-PPP) — o número IP atribuído no início da ligação é variável. Nos tempos mortos, nos quais não há troca de pacotes de dados, o número IP é considerado livre e atribuído a outro utilizador que dele necessite. A principal vantagem deste sistema é poder ter um grande número de utilizadores ligados com um mínimo de números IP atribuídos. A desvantagem é não se saber qual o IP atribuído e tornar, no caso de haver muitos utilizadores ligados em simultâneo, as comunicações mais lentas.

Em vez de uma linha telefónica normal é possível, e cada vez mais acessível, ligar através de uma linha digital RDIS. Nestas linhas o acesso básico é feito através de dois canais de dados de 64 kbps. A existência de mais de um canal permite múltiplas transmissões na mesma linha e em simultâneo. Na prática a capacidade de transmissão por canal é de duas a três vezes maior que uma comunicação com um *modem* a 28.8k. Em relação a uma linha analógica convencional a linha digital apresenta ainda tempos de estabelecimento de ligação muito curtos, uma imunidade ao ruído e às interferências muito grande

⁵ Por estados entendemos frequências, níveis de tensão ou ângulos de fase.

e, além disso, a ligação à Internet pode ser realizada em simultâneo com uma conversação telefónica.

As grandes organizações, que pretendem ter um elevado número de pessoas a utilizar a Internet com ligações permanentes de grande velocidade utilizam linhas dedicadas. Estas linhas são, regra geral, alugadas às empresas de telecomunicações.

4. O modelo Cliente/Servidor

A partilha dos recursos existentes na Internet é baseada no chamado modelo Cliente - Servidor (*Client - Server Model*). Os muitos e variados serviços oferecidos por programas servidores a correr num determinado *host* podem ser acedidos por programas cliente adequados (como se o termo cliente abrangesse a pessoa que o utiliza) localizados noutra (ou no mesmo) computador. Quando alguém está, por exemplo, a transferir ficheiros (*download*) da rede, utiliza um programa cliente que lhe fornece a interface, gráfica ou não, para aceder ao serviço. O programa remoto que recebe e processa o seu pedido é o programa Servidor. A comunicação entre o Cliente e o Servidor obedece a um determinado Protocolo. Podemos assim falar

- do Protocolo — no exemplo, protocolo de transferência de ficheiros FTP (*File Transfer Protocol*);
- do Servidor — no exemplo, servidor de FTP (programa remoto que aceita o pedido);
- do Cliente — no exemplo, cliente de FTP (programa que permite pedir o serviço);
- do Serviço — no exemplo, serviço de FTP (efectuado pelo *host* servidor).

Assim, quando dizemos, por exemplo, FTP, há que dizer se nos referimos ao protocolo, ao Servidor, ao Cliente ou ao Serviço.

5. Os serviços da Internet

Nesta secção vamos analisar os principais recursos da Internet. O objectivo é dar uma ideia daquilo que esta rede actualmente oferece aos seus utilizadores.

5.1 O correio electrónico (E-mail)

Este serviço permite enviar e receber mensagens através de um computador para, ou de, qualquer utilizador ligado à rede. O *e-mail* não significa contudo apenas mensagens convencionais. Os actuais programas de *e-mail* têm verdadeiras capacidades multimédia, sendo possível enviar textos, imagens, sons, vídeos ou programas de computador. De notar no entanto que existe um ser-

viço próprio para transferência de ficheiros, FTP, que será referido mais tarde.

Para enviar ou receber correio electrónico, o utilizador precisa de um endereço, uma caixa de correio, um serviço de entrega do correio recebido e de um sistema para enviar correio. O e-mail contém tudo isto:

- um serviço para levantar o correio recebido (*Post Offices*): programas que implementam o Protocolo POP (*Post Office Protocol*), denominados POP Servers;
- um sistema para transferir correio (*Delivery Systems*): programas que implementam o Protocolo SMTP (*Simple Mail Transfer protocol*), denominados SMTP Servers;
- caixas de correio (*Mailboxes*): áreas onde as mensagens são armazenadas até serem levantadas.

Os endereços concretos de e-mail são obtidos junto do Fornecedor do Serviço. Em geral é necessário um endereço para o Servidor SMTP, outro para o servidor POP (que pode ser o mesmo, dependendo do ISP) uma conta de e-mail e outra de POP. Vejamos um exemplo fictício para o utilizador xpto:

- Servidor SMTP: mail.telepac.pt
- Servidor POP: mail.telepac.pt
- Conta POP (*POP account*):
xpto@mail.telepac.pt
- Endereço de correio (*mail account*):
xpto@mail.telepac.pt

O correio electrónico pode também ser utilizado para distribuir mensagens sobre um determinado tema a um conjunto de subscritores. A lista de subscritores designa-se por *Lista de Distribuição (Mailing List)*. Todo o correio enviado (*posted* na gíria informática) para uma determinada Lista de Distribuição é distribuído por todos os assinantes da Lista através do seu endereço electrónico. Existem Listas de Distribuição sobre os mais variados temas, desde estações de rádio a clubes de futebol.

5.2 Ligação remota (telnet)

Este serviço TELNET permite aceder a um computador remoto independentemente do local onde este se encontra. Uma vez estabelecida a ligação, o utilizador actua como se estivesse directamente ligado ao computador acedido. Claro que para isso é necessário ter uma conta aberta nesse computador (isto é, possuir autorização para utilizar esse computador e uma área de memória atribuída), pois o utilizador é convidado a digitar o seu *username* ou *login* e a sua *password*, para se poder ligar à máquina em questão. O TELNET não é uma aplicação gráfica, pelo que a ligação não permite lançar programas que exijam uma interface gráfica.

5.3 FTP (File Transfer Protocol)

O serviço de transferência de ficheiros, FTP, tal como o seu nome sugere, serve para transmitir e receber ficheiros entre computadores (*hosts*), por exemplo entre o computador que estamos a usar e um noutro no qual tenhamos área aberta.

Existem também na Internet máquinas públicas destinadas a fornecer dados, usualmente programas e documentos, via FTP. Tais máquinas designam-se por servidores de FTP **anónimos** (*anonymous FTP Servers*). Anónimos porque ao estabelecermos a ligação, devemos introduzir como *username* a palavra *anonymous* e como *password* o nosso endereço de correio electrónico. Após a ligação estabelecida, o utilizador pode fazer o *download* (cópia para o seu próprio computador) dos ficheiros em que estiver interessado. Quase todas as grandes empresas de *software* e *hardware* possuem servidores anónimos de FTP onde disponibilizam vários produtos, por exemplo as actualizações (*update*) de programas. Além das grandes empresas existem pequenas empresas e particulares que utilizam este tipo de serviço para fornecer *software*, quer em regime de *freeware* (gratuito) quer em regime de *shareware* (regime em que o utilizador pode fazer o *download* gratuitamente enviando depois um pequeno pagamento ao autor do programa caso esteja interessado em continuar a utilizá-lo).

5.4 Grupos de discussão (Usenet newsgroups)

Estritamente falando, a Usenet (*Unix User Network*) vulgarmente conhecida por *news*, é uma rede independente e não faz parte de Internet. No entanto a maior parte dos sites da Internet aceitam tráfego da Usenet podendo desta forma os seus utilizadores utilizá-la de uma forma transparente. A Usenet é constituída por um conjunto muito grande de grupos de discussão temática que envolvem milhões de pessoas em todo o mundo. Estes grupos designam-se por *newsgroups* e os contributos individuais designam-se por artigos (ou *postings* ou *news*). A Usenet obedece ao Modelo Cliente/Servidor sendo por isso necessário um programa específico para ler os artigos (*news client*) e um Servidor para os armazenar (*news server*). O protocolo utilizado neste processo denomina-se NNTP (*Network News Transfer Protocol*).

Existem grupos de discussão sobre praticamente todos os temas em que se possa pensar (por mais esotéricos que lhe possam parecer) e como são aos milhares, o melhor é **subscrever** apenas os grupos que lhe interessam. Desse modo cada vez que activar o seu leitor de *news*, receberá a informação sobre as mensagens que entretanto chegaram. Por norma as mensagens são mantidas cerca de duas semanas no servidor. Ao ler uma mensagem pode responder, participando assim no grupo em questão.

De uma forma análoga à estrutura dos nomes dos domínios também os nomes dos grupos de *news* obedecem a uma determinada hierarquia composta por palavras

separadas por pontos. No entanto, e contrariamente aos nomes dos domínios, o topo da hierarquia é constituído pela primeira palavra. Por exemplo no grupo *pt.desporto.futebol* o topo da hierarquia é *.pt*. Os topos de hierarquia mais comuns são

- .alt* – tópicos controversos ou inabituais
- .bionet* – investigação biológica
- .bit.listserv* – listas de distribuição da rede *Bitnet*
- .clarinet* – *news* comerciais da rede *Clarinet*
- .comp* – sobre computadores
- .news* – sobre a própria *Usenet*
- .rec* – divertimentos
- .sci* – sobre ciência

5.5 – WWW (World Wide Web) - A Teia global

O *World Wide Web* (WWW), frequentemente designado apenas por *Web*, é a mais recente aplicação da Internet. Muitas vezes confundida com a própria Internet, é a aplicação mais cobiçada, e uma das mais utilizadas da rede. É também a grande responsável pelo formidável crescimento de utilização da rede a que hoje assistimos.

A *Web* surgiu no *Centre Européen de Recherches Nucléaires* (Centro Europeu de Investigação Nuclear) – CERN, em Genebra. Tim Berners-Lee, um recém-licenciado pela Universidade de Oxford, pretendia criar no CERN um novo tipo de sistema de informações através do qual os investigadores pudessem colaborar e trocar informações entre si. Para tal usou a técnica do **hipertexto** para ligar entre si um conjunto de documentos. Um hipertexto é um texto com palavras-chave (hiperligações - *hyperlinks*) que nos permitem aceder directamente a outros documentos, independentemente da sua localização. Em 1994 foi fundado o Consórcio W3 (W3C) que tem como principal tarefa desenvolver normas (*standards*) para o uso da *Web*. Este Consórcio tem o apoio de muitas instituições governamentais de vários países.

A *Web* baseia-se no protocolo HTTP (*HyperText Transfer Protocol*). Este protocolo, tal como hoje existe, foi criado no CERN em 1990 e faz uso extensivo de aplicações multimédia, tais como imagem e som. A linguagem dos documentos colocados na rede, designados por **páginas**, é o HTML (*HyperText Markup Language*). O endereço específico em que uma dada página pode ser encontrada designa-se por URL (*Uniform Resource Locator*). Um URL tem a seguinte forma:

protocolo://servidor/ficheiro

Para a *Web* o protocolo é como vimos o *http*. Assim, para um documento de nome *texto.htm* localizado na directoria *dir* de um determinado servidor o URL será dado por:

http://servidor/dir/texto.htm

Esta forma de identificar a informação disponibilizada na rede é geral. Por exemplo se em vez de um servidor

de WWW tivéssemos um servidor de FTP, o seu URL seria:

ftp://servidor/ficheiro

Para podermos aceder às páginas da Web, necessitamos de um programa que possa aceder a um dado URL, e que entenda a linguagem HTML, em que o documento está escrito. Tal programa designa-se em gíria informática um *browser*. Os mais conhecidos são o *Netscape Navigator* da Netscape e o *Internet Explorer* da Microsoft. Estes *browsers* utilizam uma interface gráfica orientada para a utilização do rato. A navegação é feita apontando com o rato para determinadas zonas marcadas (habitualmente sublinhadas ou/e a outra cor) – *hiperlinks* – que ligam entre si documentos nos mais variados locais. É o que se designa por *surfing na net*. Esta “navegação” é cada vez mais um processo interactivo, através do qual é possível, por exemplo, consultar bibliotecas ou preencher formulários.

Existem actualmente milhões de páginas na Web e este número aumenta todos os dias. Estas páginas cobrem os mais variados temas. Com tantas páginas na rede, surge obviamente o problema de como procurar informação específica. Para isso existem vários catálogos ou índices temáticos e também os chamados Motores de Busca (*Search Engines*) que nos permitem fazer pesquisas temáticas, ou por palavras-chave, na rede. Em geral as duas funções estão associadas. Eis alguns dos mais populares:

AltaVista - <http://altavista.digital.com/>

Lycos - <http://www.lycos.com/>

Excite - <http://www.excite.com/>

Yahoo - <http://www.yahoo.com/>

Em Portugal existem também alguns Motores de Busca, dos quais destacamos

SAPO - <http://www.sapo.pt/>

AEIOU - <http://www.aeiou.pt/>

CUSCO - <http://www.cusco.pt/>

5.6 Outros serviços

Existem muitos outros serviços na Internet, mas menos utilizados do que os referidos até aqui ou mesmo em desuso. Eis alguns deles.

5.6.1 Conversação em tempo real (IRC – Internet Relay Chat)

O IRC é um sistema de teleconferência onde um conjunto de utilizadores pode “conversar” em tempo real sobre um determinado tema. Cada palavra é escrita e imediatamente transmitida à totalidade dos restantes participantes que pode atingir vários milhares distribuídos por todo o planeta.

5.6.2 Finger

Este é um serviço que permite obter informação sobre um utilizador específico num determinado *host* ou sobre os utilizadores desse *host* em geral. A informação obtida é do domínio público e disponibilizada para o efeito por cada utilizador.

5.6.3 Gopher e Veronica

O *Gopher*, é um serviço que permite aceder a informação através de menus. A cada opção desses menus podem estar associados ficheiros ou outros serviços (por exemplo telnet). Apesar de ser relativamente recente (1994), devido ao aparecimento do WWW, que permite muito mais opções de uma forma muito mais atraente, o *Gopher* encontra-se agora praticamente em desuso.

O serviço *Veronica* (*Very Easy Rodent-Oriented Netwide Index to Computerized Archives*) é acedido através do *Gopher* e destina-se a facilitar a navegação nos menus destes.

5.6.4 Archie

O serviço *Archie* serve para procurar um ficheiro nos milhares de servidores anónimos de ftp que existem em todo o mundo. Por exemplo imagine que se pretende obter o programa *xpto* que se pensa existir algures num servidor de ftp anónimo. Para isso entramos no serviço de *Archie* e digitamos a palavra *xpto*. Passados alguns instantes devemos obter uma listagem com os locais onde a informação que pretendemos se encontra.

Este serviço encontra-se também em desuso devido aos *Search Engines* do WWW, que permitem obter a mesma informação de uma forma mais fácil.

5.6.5 WAIS

O WAIS (*Wide Area Information Server*) é outro dos serviços que o WWW veio tornar obsoleto. Este sistema destina-se à pesquisa de documentos em bases de dados numa linguagem simples e com recurso a palavras chave.

Bibliografia

1. P. Syros and V. Terzopoulos, *Web for Schools Internet Starter Kit*, Context – European Education Magazine, N.º 17, 1997.
2. A. Costa, E. Rodrigues, F. Pinto, J. Macedo e M. Nicolau. *Internet – Guia prático do Cibernauta*. Ed. Campo de Letras, Porto, 1995.
3. Manuel Lemos. *Estar na Internet*. Ed. MacGraw-Hill de Portugal, Lisboa, 1997.

José Manuel Pires Marques é Professor Auxiliar da Faculdade de Ciências da Universidade de Lisboa.
Fernando Costa Parente é Professor Associado na mesma instituição e especialista em Física Atómica.