#### Internet - impacto, virtudes e limitações de um modelo com mais de 40 Anos

José Legatheaux Martins
Departamento de Informática da
FCT/UNL

### Objetivos da Lição

- Quais os Fundamentos Tecnológicos da Internet
- Como a Internet e as aplicações evoluíram e que impacto isso teve nos seus fundamentos
- O modelo inicial ainda é válido?
- Que encruzilhadas enfrenta?

Não se pretende introduzir nada de novo. Apenas analisar qual a situação e como cá chegámos

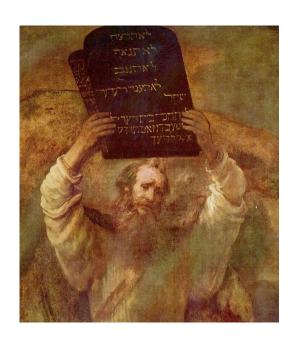
### Origem

- Em 1968 a USA DARPA (Defense Advanced Research Projects Agency) lançou um projeto para usar a tecnologia de comutação de pacotes de dados para ligar os computadores de várias universidades americanas com o objetivo de permitir a partilha de recursos computacionais entre investigadores.
- O projeto evoluiu depois para usar um <u>conjunto</u> <u>heterogéneo de redes</u> (linhas terrestres, satélite, rádio, ...)
- Em 1983 o modelo dessa rede estava definido, testado e em produção e esta recebeu o nome de <u>Internet</u>

#### Modelo Internet

#### Modelo

- Conjunto de princípios
- Conjunto de soluções tecnológicas que materializam esses princípios



#### Princípio 1 - Rede de Pacotes "Best Effort"

- A rede encaminha pacotes da origem até ao destino numa base "best effort", isto é, não há garantias de que o pacote chegue ao destino
- Um pacote pode perder-se, ou até chegar em duplicado e não é garantida a ordem da entrega de vários pacotes emitidos em sequência
- A rede não "pede desculpa" por ter falhas
- Compete ao emissor e ao recetor lidarem com as falhas e as deficiências e lidarem com a complexidade das aplicações (mais tarde teorizado no chamado "end-to-end principles in system design")

### Solução Tecnológica - Protocolo IP

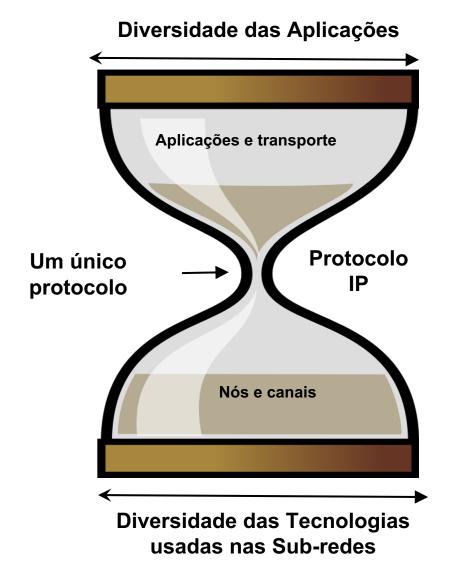
- Define o conceito de pacote de dados da Internet
  - Cabeçalho com endereços origem e destino, alguma pouca informação de controlo
  - Conteúdo (payload) opaco e não interpretado



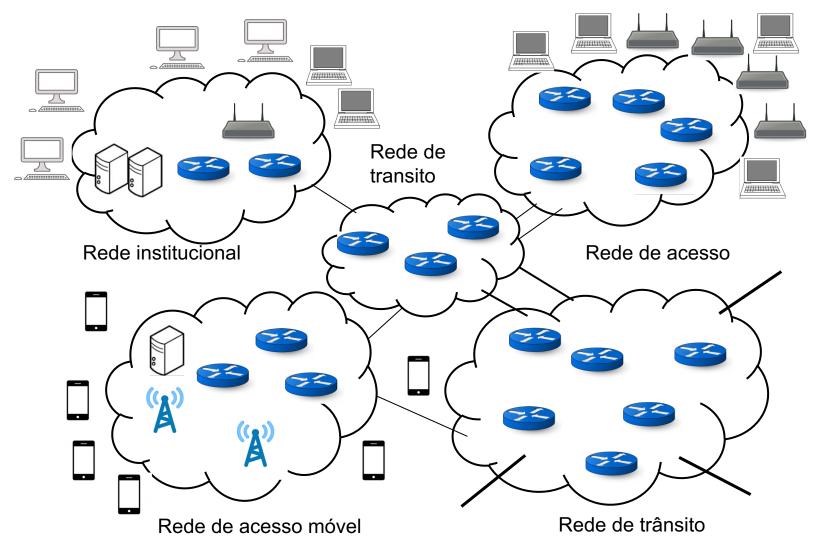
- Define em que condições este é encaminhado pela rede e com que garantias
- É a principal interface do modelo e é usada entre os dispositivos e a rede e entre sub-rede

#### O Protocolo IP é o Coração do Modelo

Vitória absoluta da Especificação Minimal ("Under Specification") e da Simplicidade



#### Princípio 2 - A Internet é uma Rede de Redes ou Sistemas Autónomos



#### Outra Base da Escala

 O funcionamento interno de cada <u>sistema autónomo</u> (AS) é escondido ao exterior

- Cada AS apenas se compromete a participar no encaminhamento dos pacotes para os destinos que diz poder alcançar
- Não existem barreiras a tornar-se um sistema autónomo basta "estabelecer contratos" com os vizinhos e obedecer a convenções mínimas comuns
- O sistema pode expandir sem restrições regulatórias intrínsecas significativas (mas não impede que existam barreiras políticas ou económicas a essa expansão)

### Solução Tecnológica - Protocolo BGP

- Os sistemas autónomos (ASs) comunicam uns aos outros os destinos finais que conhecem e que outros sistemas autónomos usam para chegar a cada destino.
- Cada AS usa essa informação para decidir como encaminhar os pacotes para o destino
- A informação passada de AS em AS é como um "boato" (está certo em 99,9% dos casos).
- Se há razões para achar que o vizinho mente, ignore o que este lhe diz
- Se não quer que um vizinhe não o use para chegar a um destino, não lhe diga que conhece o destino
- Permite a implementação de políticas comerciais entre ASs

### Princípio 3 - Endereços

- Os computadores ligados à rede têm um endereço único (uma espécie de número de telefone)
- Esses números estão organizados em prefixos hierárquicos
- Qualquer instituição, independentemente do país pode receber um prefixo e tecnicamente pode haver "roaming dos endereços sem limites" (ainda que na prática o mesmo esteja comercial e tecnicamente limitado)
- Para facilitar a descoberta dos endereços foi introduzido o sistema DNS (Domain Name System), uma base de dados descentralizada de mapeamento de nomes em endereços

### Implícito no Modelo

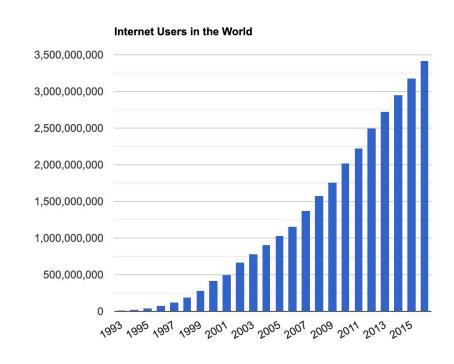
- Autenticação e a Identificação remetidas para os computadores ligados na periferia da rede
- Gestão minimalista e sem fronteiras
  - Um sistema autónomo não pode dar garantias que envolvam todos os outros sistemas autónomos
  - Qualidade de serviço minimalista
  - Gestão transnacional apenas tenta manter o conjunto operacional dentro dos princípios enunciados atrás
- Os USA decidiram explicitamente não regular as aplicações e o mundo seguiu (exceto a China)

# Evolução da Internet

#### 40 Anos Depois - 100 Milhões de Vezes Maior

#### • 2021

- Cerca de 4 biliões de utilizadores
- Cerca de 20 biliões de dispositivos computorizados ligados
- 1983 Modelo e protocolos TCP/IP definidos, testados, normalizados e operacionais
  - Rede com 400 computadores
- 1969 Protótipo
  - 4 computadores
- 1968 Lançamento do projeto

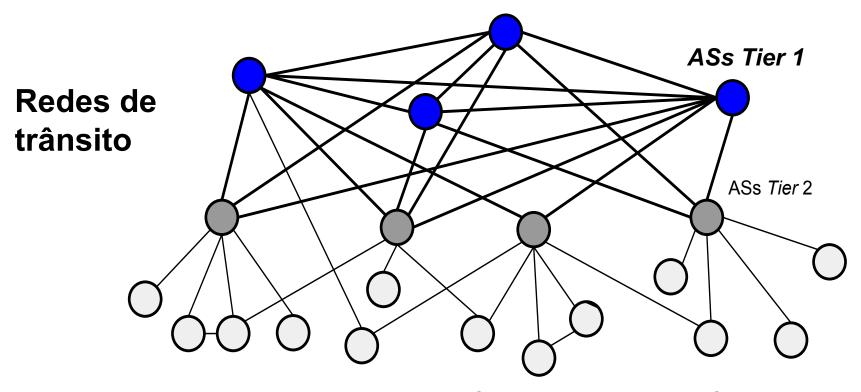


Fonte: www.internetlivestats.com

### Evolução Marcada Por Fases

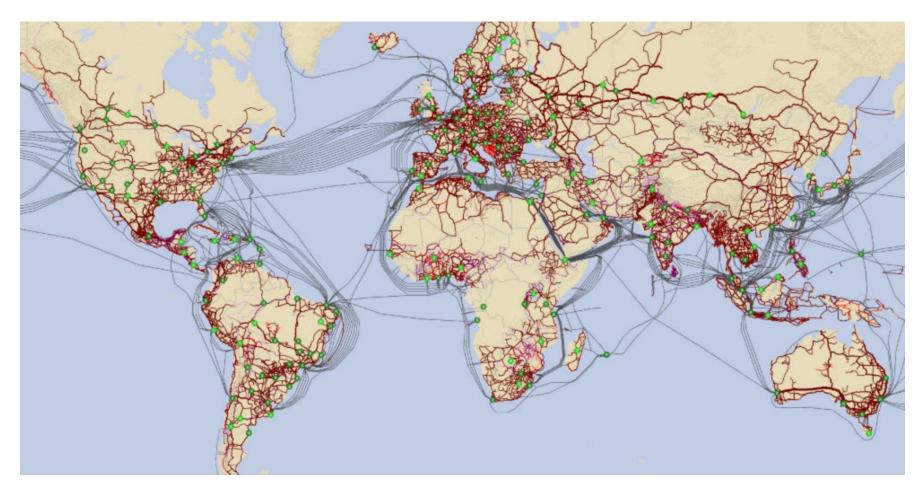
Época	Aplicações Introduzidas	Traços fundamentais Novos	Capacidade dos canais (Acesso / Core)	Suporte Computacional às aplicações
(1) Internet dos pioneiros (<1990)	Transferência de ficheiros, Email, Login remoto	Pessoas acedem a computadores remotos	< 56 Kbps / 1 Mbps	Mainframes e time-sharing
(2) + Internet = Web (1990 – 2000)	Browser acede a dados estáticos ou a aplicações interativas individuais	Pessoas acedem a informação e a alguns serviços	1 Mbps / 1 Gbps	Computadores pessoais (PCs)  Servidores ou clusters de servidores

# Estrutura Lógica da Internet (por volta do ano 2000)



Redes de acesso (e institucionais)

### "Core" - Operadores Tier 1 e 2



Fonte: ITU

# 2000 - 2010: Despontam as Aplicações Planetárias

Época	Aplicações Introduzidas	Traços fundamentais Novos	Capacidade dos canais (Acesso / Core)	Suporte Computacional às aplicações
(3) + Internet dos serviços planetários (2000 – 2010)	Web search  Partilha de conteúdos e difusão de conteúdos.  Aplicações para milhões	Serviços interativos generalizados, comunicação pessoa a pessoa	< 1 Mbps / 10 Gbps	+ smartphones (2007)  Servidores em centros de dados

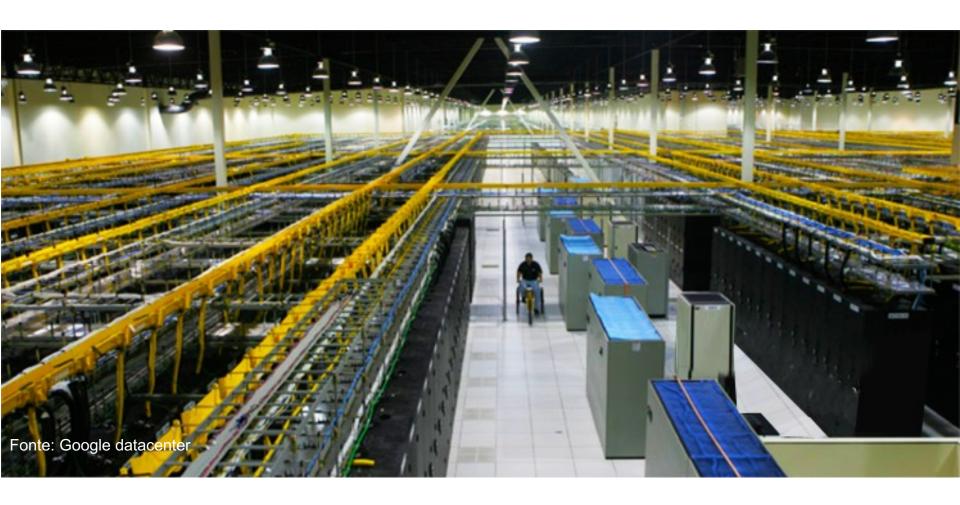
# > 2010 - Abrangência Total

Época	Aplicações Introduzidas	Traços fundamentais Novos	Capacidade dos canais (Acesso / Core)
(4) + comunicação interativa pessoa a pessoa, streaming de som e vídeo, aquisição de dados, IoT, sensores e atuadores, controlo remoto	Serviços interativos para milhões, difusão de mensagens, streaming de televisão e rádio,  Aquisição de dados para aprendizagem automática de suporte às aplicações interativas	< 1 Gbps / > 0,4 Tbps	+ dispositivos IoT  Centros de dados com centenas de milhar de computadores, com circuitos especializados para IA e paralelismo, complementados com centros de dados nas redes de acesso

### Duas Perspetivas de Evolução

- Responder aos novos desafios de forma gradual e reativa
- Responder redesenhando a arquitetura da Internet de forma pró-ativa ("clean slate approaches")
  - Information centric networking
  - Content centric networking
  - Named data networking
  - Multicast para suporte à difusão global
  - e muitas outras propostas de menor âmbito

### Cloud Computing foi a Resposta



#### Redes Privadas de Centro de Dados



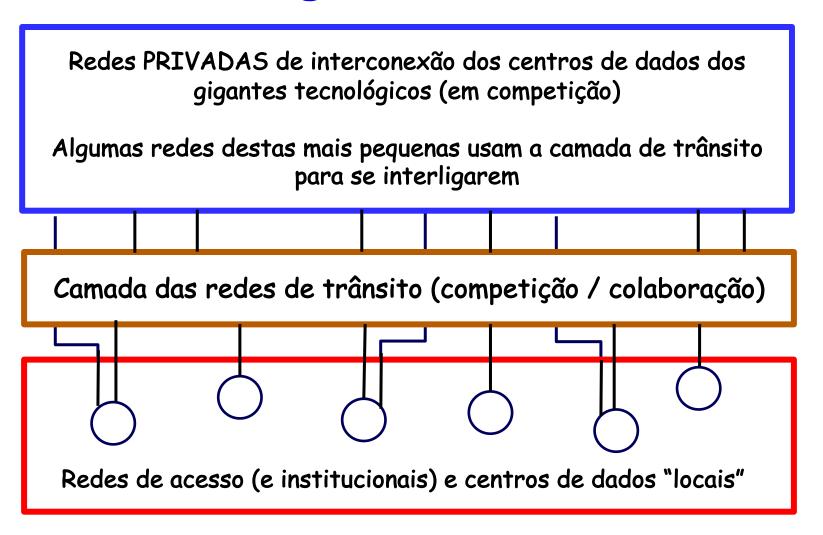
Fonte: Microsoft Azure Network

Solução comum aos grandes da Indústria (Google, Amazon, ...)

# É a Logística e a Distância, Estúpido!

- É verdade que os canais de longa distância têm hoje muito maior capacidade, mas trazer os dados e a computação para mais próximo dos utilizadores fez todo o sentido
- Na verdade a Internet atual funciona como uma cadeia de distribuição com grandes armazéns, os centros de dados, junto das concentrações de utilizadores (as redes de acesso)
- A dimensão e consolidação favorece a verticalização pois os gigantes da Internet otimizam as redes de longa distância - é como se cada um deles tivesse a sua companhia de transportes própria

### Estrutura Lógica Atual da Internet



## O Modelo Internet 40 Anos Depois

### Princípio 1 - Rede de Pacotes "Best Effort"

- Erros A latência média diminuiu pois os destinos estão mais próximos, logo consegue-se recuperar dos erros nos dispositivos terminais ainda mais facilmente
- Difusão Enviar os mesmos dados e pacotes de um emissor para muitos recetores pode ser feito a partir da estrutura dos centros de dados
- Enquanto o paradigma de comunicação ponto a ponto satisfizer os requisitos, o Princípio 1 mantém-se e revelou-se certeiro
- A exceção são as redes interplanetárias pois a latência é gigantesca e existem momentos em que a comunicação é impossível

#### Princípio 2 - A Internet é uma Rede de Redes ou Sistemas Autónomos

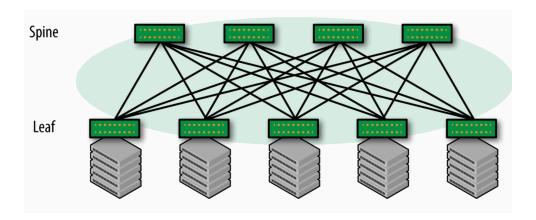
Dada a autonomia e opacidade do funcionamento das diferentes redes <u>só</u> <u>é possível oferecer serviços globais</u> <u>envolvendo vários operadores, com base</u> <u>no menor "consenso comum"</u>



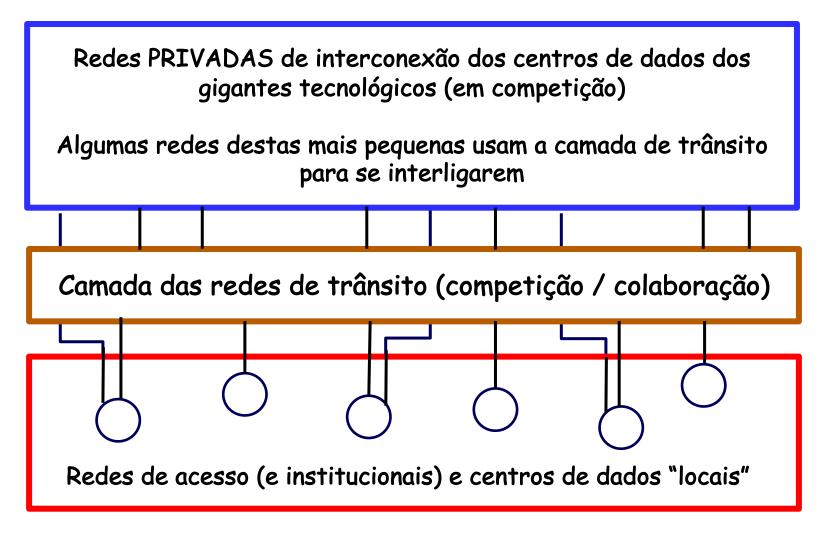
- Este princípio tem sido vital para a escalabilidade mas "congelou as soluções no core"
  - A distribuição de carga e a otimização do core são um pesadelo para os operadores
  - Os dispositivos não podem escolher o caminho dos seus fluxos de dados
  - O nível de segurança do core é baixíssimo
- Investigadores: qualquer nova arquitetura tem de coexistir com a existente e ter um modelo de negócio viável para a transição

#### Data Center Networking = Inovação Localizada

- Nas redes dos centros de dados e nas redes privadas que os interligam "nada ficou como antes"
  - Otimização e gestão logicamente centralizada da rede
  - Emergiu o paradigma "Software Defined Networking"
  - Configuração de sub-redes a pedido e adaptadas às necessidades das aplicações
  - Reconfiguração dinâmica e transparente
  - Software dos equipamentos de rede com base semelhante ao dos servidores



#### Conflito Entre Níveis



#### Propostas "New IP"

 Geralmente aparecem via o ITU e são feitas pela China ou pela Rússia, e inscrevem-se em manobras geopolíticas

- Outras, com origem também no IUT, são promovidas por operadores de telecomunicações de grande dimensão e inscrevem-se em competições sobre quem controla o "Core"
- O problema da "ausência de qualidade de serviço"
  - Exemplos de aplicações: cirurgias remotas e presença holográfica
  - Essas aplicações não são compatíveis com grandes distâncias pois não toleram latências significativas
  - Podem ser resolvidas dentro do mesmo AS ou por ASs contíguos

#### Como Evoluiu a Parte Implícita do Modelo?

- Segurança é um problema das aplicações e do "edge"
- Regulação mínima e sem fronteiras

#### Segurança — O Elefante na Sala

- Inicialmente a rede era apenas usada por académicos
- Só a criptografia simétrica era conhecida e divulgada
- Usando apenas criptografia simétrica para resolver o problema da autenticação, provavelmente teria de se concentrar nos operadores das redes de acesso a autenticação e a distribuição de chaves criptográficas (como na redes móveis)



 Resultado provável: uma Internet menos escalável e aplicações menos inovadoras

#### Criptografia de Chaves Públicas e Privadas

- É a base da
  - Da <u>identidade</u> das entidades (uma entidade pode ter um número de entidades ilimitada)
  - Da <u>autenticação</u> das entidades
  - Da emissão de certificados e da sua verificação
  - Da <u>distribuição dinâmica de chaves criptográficas</u> entre os dispositivos finais
- Usada para estabelecer canais seguros e resistentes a ataques através da Internet (HTTPS)
- No entanto, ainda não temos grandes dificuldades em
  - Proteger os dispositivos ligados à rede
  - Mitigar os ataques de paralisação dos serviços (DDoS)

#### Cloud-based Content Distribution and Protection

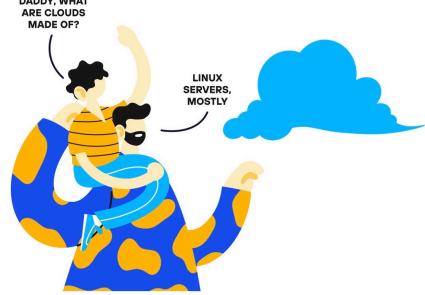


Exemplos: Akamai, Amazon CloudFront, Cloudflare, Fastly, Limelight e ainda Microsoft Azure, Google Cloud, Alibaba Cloud, Tencent Cloud

Fonte: https://www.cloudflare.com

#### Exemplo Cloudflare:

- Mais de uma centena de centros de dados no mundo
- Ligados diretamente a cerca de 10.000 redes de acesso e mega centros de dados
- Filtram o tráfego para os servidores aplicacionais dos clientes



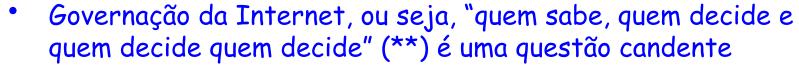
34

### Implicações

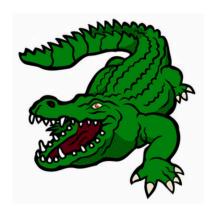
- Dependência cada vez maior dos Cloud Providers
- A segurança, o funcionamento e a estabilidade da Internet está muito dependente de uma ou duas dúzias de empresas
- Ausência de diversidade é um grande perigo qualquer erro ou problema tem consequências globais, eventualmente dramáticas

#### Regulação: "Wrestling with Alligators" (\*)

- Alguns desses "crocodilos" são:
  - Privacidade e Identidade
  - Segurança e uso da Internet para atividades ilegais
  - Propriedade intelectual
  - Uso da inteligência artificial e controlo social
  - Implicações nas relações de trabalho (Gig-economy)
  - Cyber-guerra
  - **–** ....



- As redes de trânsito só transportam pacotes, e as redes de acesso sempre estiveram enquadradas em regulamentação nacionais. Mas as aplicações evoluíram sem qualquer regulação e com uma concentração de dimensão planetária
- (\*) Parafraseando Vint Cerf
- (\*\*) Parafraseando Shohsana Zuboff



#### Valorizações Estratosféricas (Bolha?)

Das empresas mais valorizadas na bolsa de 2019 a 2021, 7 operam na Internet e estão quase sempre no topo

Nome	País	Valor (*) em Mil milhões \$US
Apple	USA	2 339 (**)
Microsoft	USA	2 119
Alphabet	USA	1 777
Amazon	USA	1 664
Facebook	USA	956
Alibaba Group	China	766
Tencent	China	637

<sup>(\*)</sup> O valor na bolsa vai variando. Cita-se o valor mais alto atingido.

<sup>(\*\*)</sup> Este valor é cerca de 9 vezes o PIB de Portugal.

#### Um Grupo de Monopolistas?

Aplicação	Região	Empresas dominantes (faturação)
Search (*)	Mundo sem China	Google – 92%, Microsoft – 3%,
Search (*)	China	Alibaba + Tencent – 91%, Microsoft – 4%,
Browsers (*)	Mundo	Google – 65 %, Apple, Microsoft, Firefox,
Software operação dispositivos pessoais (*)	Mundo	Google 41%, Microsoft 32%, Apple 23%,
Redes sociais (*)	Mundo sem China	Facebook 76%, Twiter, Pintrest, Google,
Publicidade na Internet (**)	Mundo	Google 28%, Facebook 25%, Alibab 10%, Amazon 7%, Tencent,
Cloud (***)	Mundo sem China	Amazon – 32%, Microsoft – 20%, Google,
Correio	Mundo sem China	Google – 20%, Microsoft – 11%,
CDNs, Security (##)	Mundo	Akamai – 35%, Cloudflare 19%, Fastly, Verizon, Amazon,

(\*) Fonte: <a href="https://statcounter.com">https://statcounter.com</a>
(\*\*) Fonte: <a href="https://www.emarketer.com">https://www.emarketer.com</a>
(##) Fonte: <a href="https://www.t4.ai">https://www.t4.ai</a>

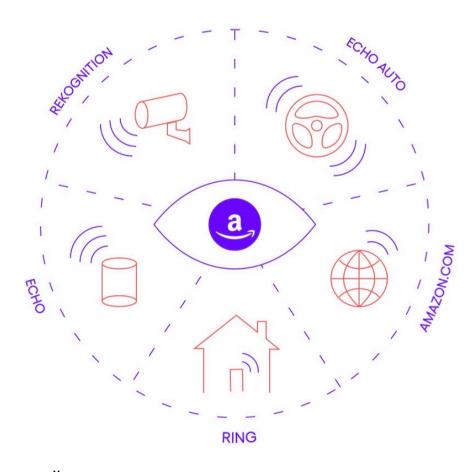
(\*\*\*) Fonte: <a href="https://www.statista.com">https://www.statista.com</a>
(#) Fonte: <a href="https://w3techs.com">https://w3techs.com</a>

### Expetativas de Ganhos Futuros

- Efeito de escala permite otimização e gera concentração (efeito "the winner takes it all") monopolização
- Modelo de negócio "serviço por dados"
  - O Modelo de negócios implícito nos serviços gratuitos
  - Dimensão dos dados que coletaram e coletam
  - Infraestrutura computacional uma ordem de grandeza à frente de qualquer concorrência pública ou privada
  - Domínio absoluto sobre a ciência dos dados e o capital humano (know how)

### Inteligência Artificial e Vigilância

#### What Amazon Sees and Hears



Fonte: https://onezero.medium.com

### Regulação e Geopolítica

- Os "campeões" são Americanos ou Chineses
- A China constituiu um bloco à parte onde existe uma estreita coordenação entre os seus campeões e o poder de Estado. Nos EUA isso só acontece no capítulo da Segurança Nacional
- A Europa quase não tem equipas que joguem neste campeonato debate "Soberania Digital da Europa"
- Mas apresenta-se como "<u>Campeã da Regulação"</u> com base na dimensão do seu mercado
- Será suficiente? Veja-se o exemplo do RGPD cuja aplicação não alterou a essência das coisas pois não tocou no modelo de negócio

### Regulação e os Cidadãos

- Na sociedade do conhecimento o controlo da informação e da decisão é a sede do poder
- Os gigantes tecnológicos estão a controlar
  - A aquisição de dados
  - A aprendizagem automática (AA)
  - Os investimentos no desenvolvimento científico em AA
  - A transformação desse processo em valor
- Nos desafios que se apresentam <u>são parceiros</u> <u>credíveis que não é preciso regular ou soluções privadas</u> <u>para problemas e serviços públicos a manter sob</u> <u>controlo</u>?

#### Todo o Conhecimento ao Alcance da Mão



### Regulação da Internet

- É de facto maravilhoso que todo o CONHECIMENTO esteja aparentemente ao "alcance da mão", mas é simultaneamente terrível que algumas entidades geridas segundo o "capitalismo da vigilância" decidam ou influenciem o que e como cada um o pode alcançar
- Qualquer regulação da Internet deve visar o bem comum e a defesa dos direitos humanos, mas necessita também de uma análise profunda do ponto de vista tecnológico e das suas repercussões.
- Ou seja, os decisores têm de perceber a tecnologia e os tecnólogos têm de perceber as implicações sociais da tecnologia

### Agradecimentos