

Complex Networks

Lecture 4: Internet

-- Topology and Modeling

EE 6605

Instructor: G Ron Chen





QQ Networking

Internet

❖ **Internet** is a global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide (Wikipedia)

❖ **Structure:**

- Autonomous Systems (AS)
- Wide Area Network (WAN)
- Metropolitan Area Network (MAN)
- Local Area Network (LAN)

❖ **Software:**

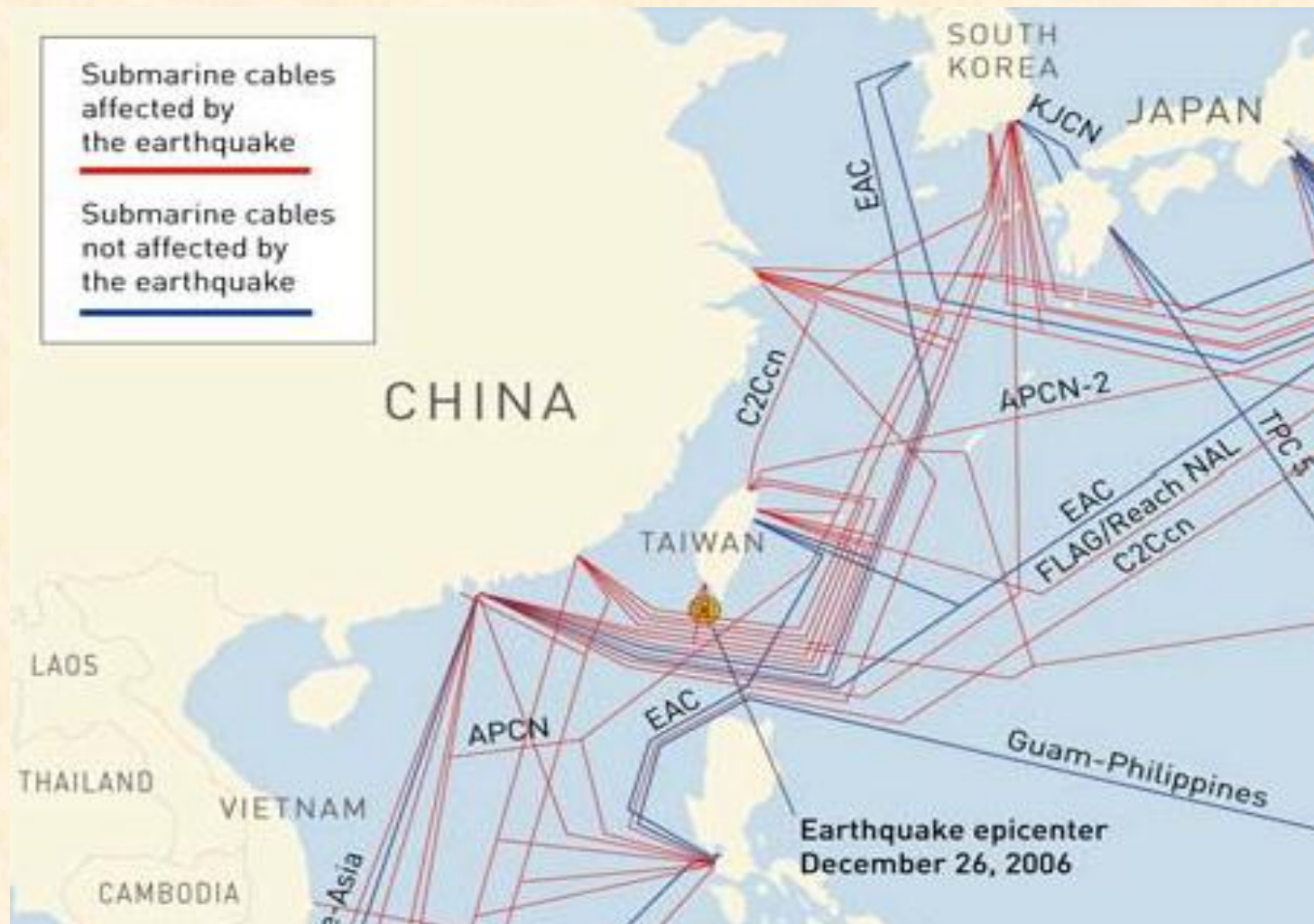
- Transmission Control Protocol (TCP)
- Internet Protocol (IP)
- Other protocols at the application level:
 - Internet Control Message Protocol (ICMP)
 - File Transfer Protocol (FTP)
 - Simple Mail Transfer Protocol (SMTP)

Internet is a network of networks

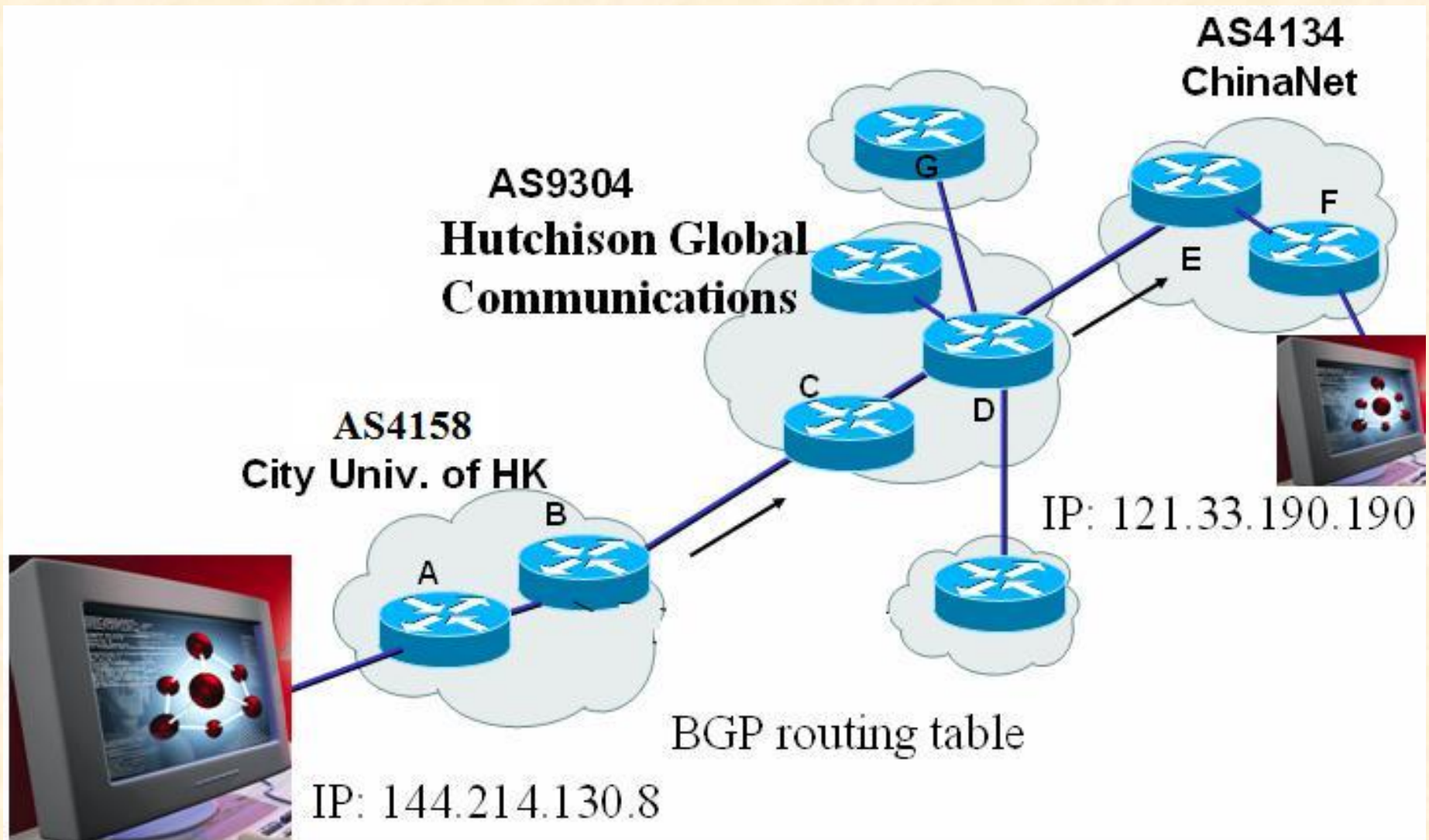
(structure)

- ❖ Internet is the inter-connections of networks:
 - Internet Service Provider (ISP) networks
 - Corporate networks (e.g. Cisco and Microsoft)
 - Campus networks (e.g. CityU)
 - Mobile operator networks (e.g. one2free)
 - Government networks
 - Internet Exchanges (e.g. HKIX)
- ❖ through optical fibers (telephone lines)

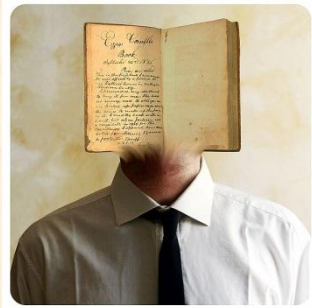
Submarine Cables in East Asia



A Real Example



A Typical Application Example - Facebook



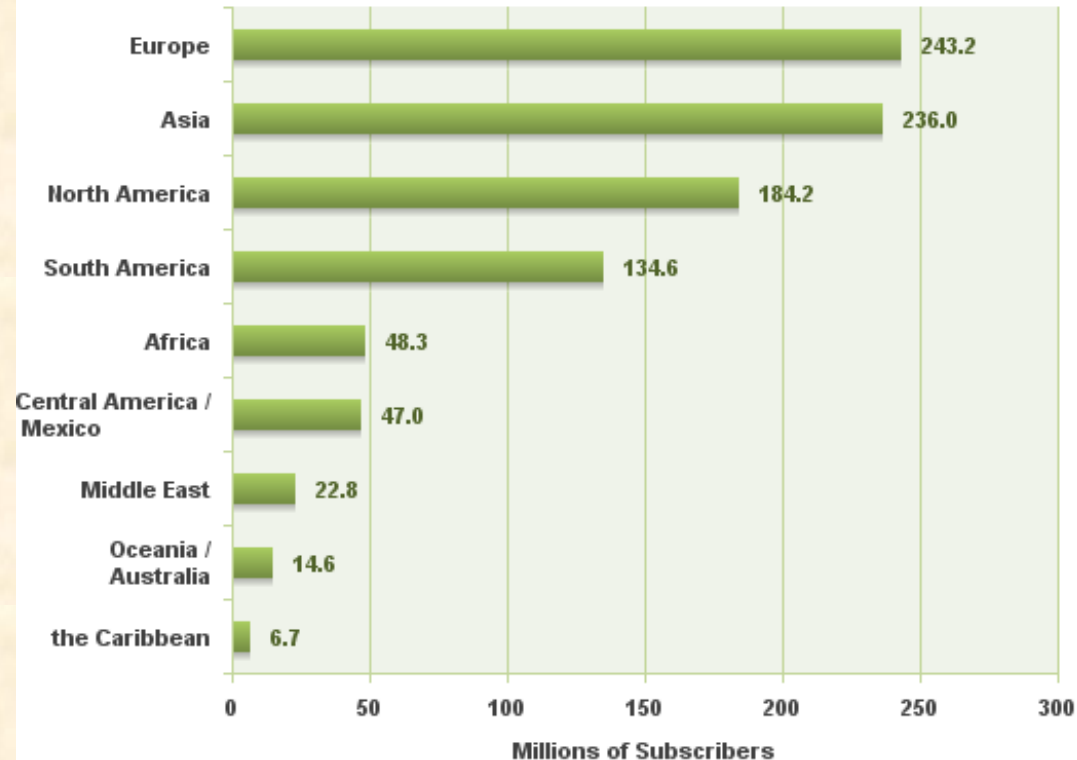
World Map



www.bbc.co.uk

Data Source

**Facebook Users in the World
by Regions - September 2012**



Source: Internet World Stats - www.internetworldstats.com/facebook.htm
The total number of Facebook subscribers in the world is estimated to be 937,407,180 on September 30, 2012.
Copyright © 2012, Miniwatts Marketing Group

Brief History of the Internet

- ◆ **1968** - DARPA (Defense Advanced Research Projects Agency) contracts with BBN (Bolt, Beranek & Newman) to create ARPAnet
- ◆ **1970** - First ARPAnet had 5 Nodes:
UCLA, Stanford, UC Santa Barbara, U of Utah, BBN
- ◆ **1974** - TCP specification
- ◆ **1984** - January 1, 1984, the Internet with its 1000 hosts converts altogether to using TCP/IP for messaging
- ◆



Leonard Kleinrock (1934 -)

Leonard Kleinrock conducted early research in queuing theory, which proved important in packet switching, and he later played a leading role in building and managing the world's first packet-switched network - ARPAnet

1961 MIT PhD Thesis:

Information Flow in Large Communication Nets

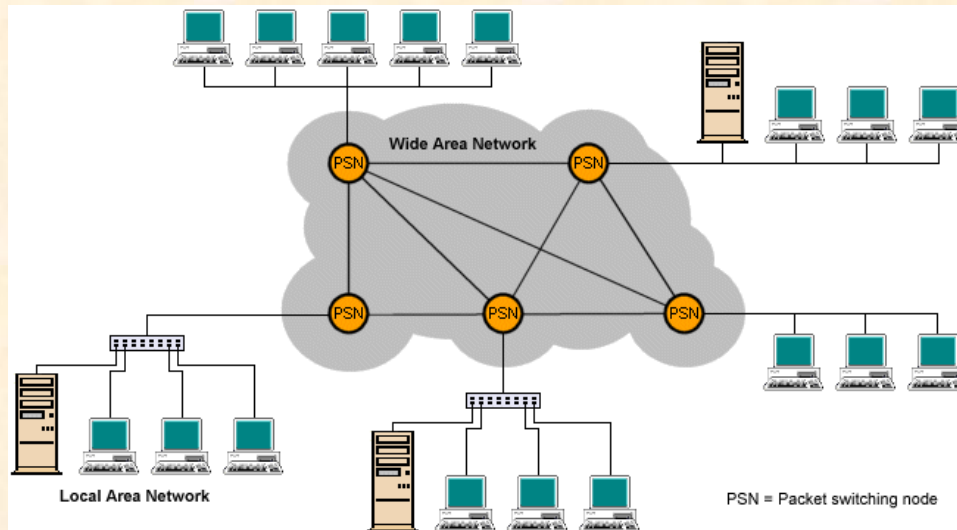
(on queuing theory and packet-switching networks)



Packet Switching is a digital networking communication method, which groups all transmitted data – regardless of their content, type or structure – into blocks, called **packets**

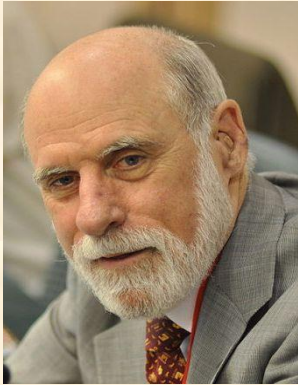


Paul Baran
(1926-2011)



Donald Davies
(1924-2000)

The concept of **switching small blocks of data** was first explored independently by **Paul Baran** at the RAND Corporation in the US and by **Donald Davies** at the National Physical Laboratory (NPL) in the UK in the earlier 1960s



Vinton Gray Cerf
(1943 -)

“Fathersu of Internet”



Robert Elliot Kahn
(1938 -)

TCP (Transmission Control Protocol) was invented by Vint Cerf and Bob Kahn in 1974, which was split to TCP and IP (Internet Protocol) in 1978

DoD used TCP/IP to inter-connect all networks in 1982, thus the real Internet was considered to be born

By the way, in 1967-1972, Vint Cerf was a graduate student in Kleinrock's lab, working on application level protocols for the ARPAnet (File Transfer Protocol, Telnet Protocol, etc.)

Protocols

- ❖ **Protocol** originates from the Greek word *protocollon*, which was a piece of paper glued to a manuscript to describe its contents.
- ❖ In computing and communication systems, a **protocol** is a set of rules, a convention, or a standard, that controls the connection and data transfer between endpoints.
- ❖ **Telecommunication systems:**
 - data interchange protocols at the hardware device level
 - data interchange protocols at the application program level
- ❖ **Internet:**
 - Transmission Control Protocol (TCP), for exchanging messages with other Internet points at the information packet level
 - Internet Protocol (IP), for sending and receiving messages at the Internet address level
 - Other protocols: Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP), Border Gateway Protocol (BGP), and Dynamic Host Configuration Protocol (DHCP), ...



“Father of WWW”



Sir Timothy John Berners-Lee
(1955 -)

Tim Berners-Lee (“TimBL”) developed a hypertext system with initial versions of HTML and HTTP and first GUI web browser called **World Wide Web** in 1990

Tim Berners-Lee is the director of the World Wide Web Consortium (W3C) and he is also the founder of the World Wide Web Foundation, established at MIT in 1994

WWW: Social Media

Global Social Media

Facebook

YouTube

Linkedin

Social Media in China

Weibo 微博

Renren 人人

Kaixin 开心

Youku 优酷

Tudou 土豆

Search Engine Marketing

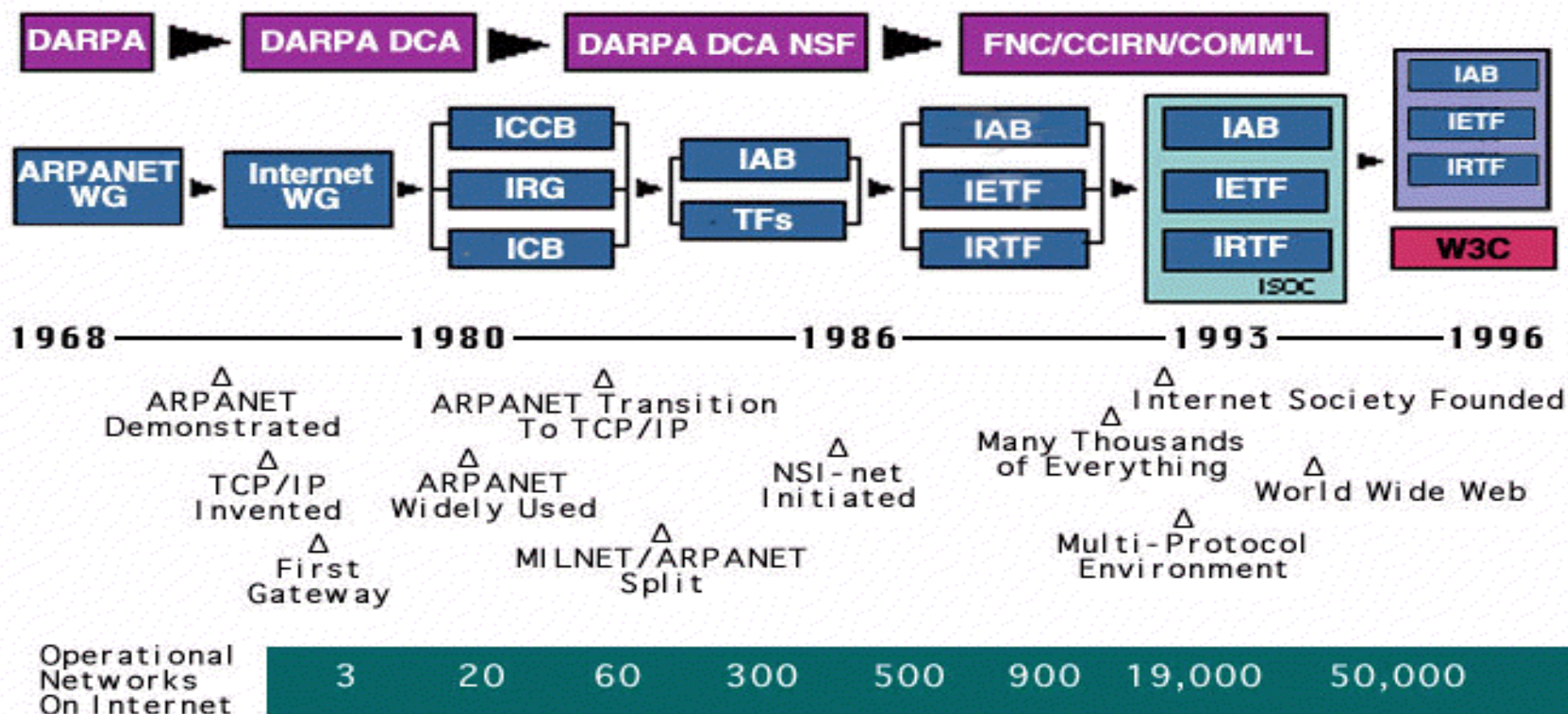
Google

Yahoo!

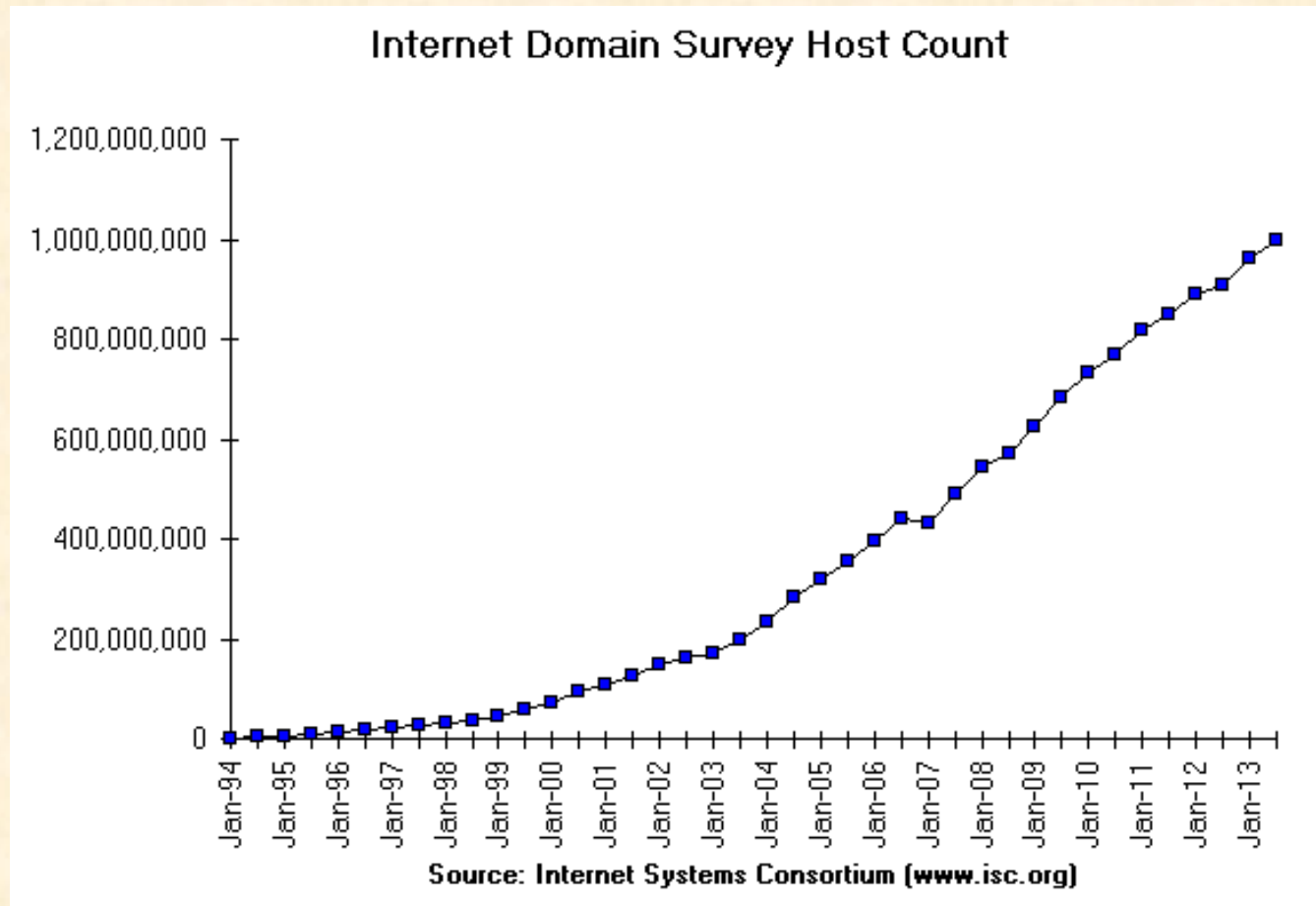
Baidu 百度

Earlier History of Internet

DARPA



Internet today

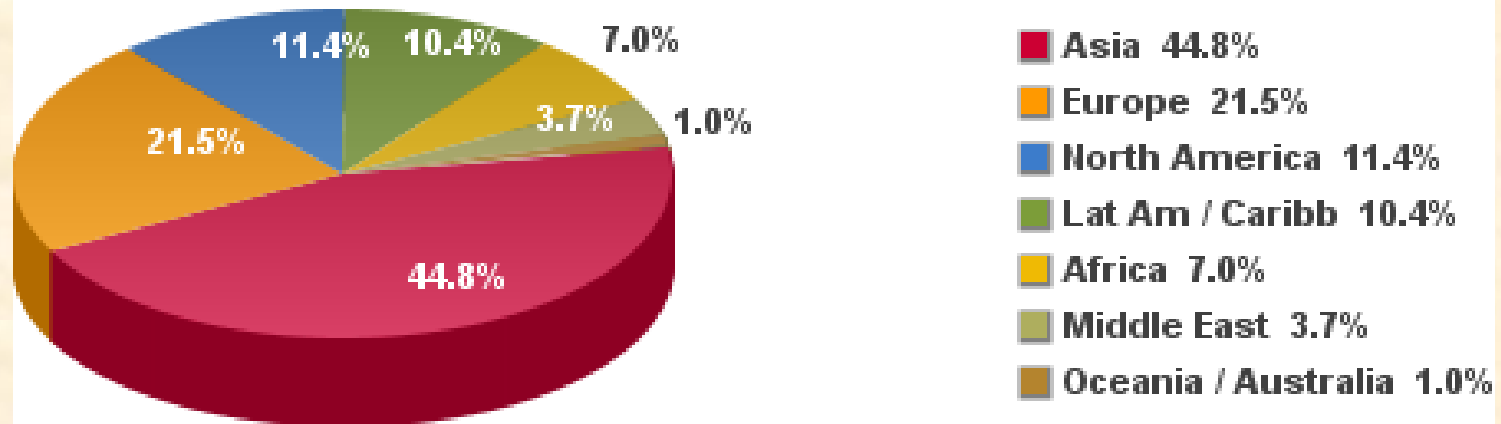


Host count was 768,913,036 in 2010
93,047,785 in 2000

↑ 826.36%

Data Source

Internet Users in the World Distribution by World Regions - 2012 Q2



Source: Internet World Stats - www.internetworldstats.com/stats.htm

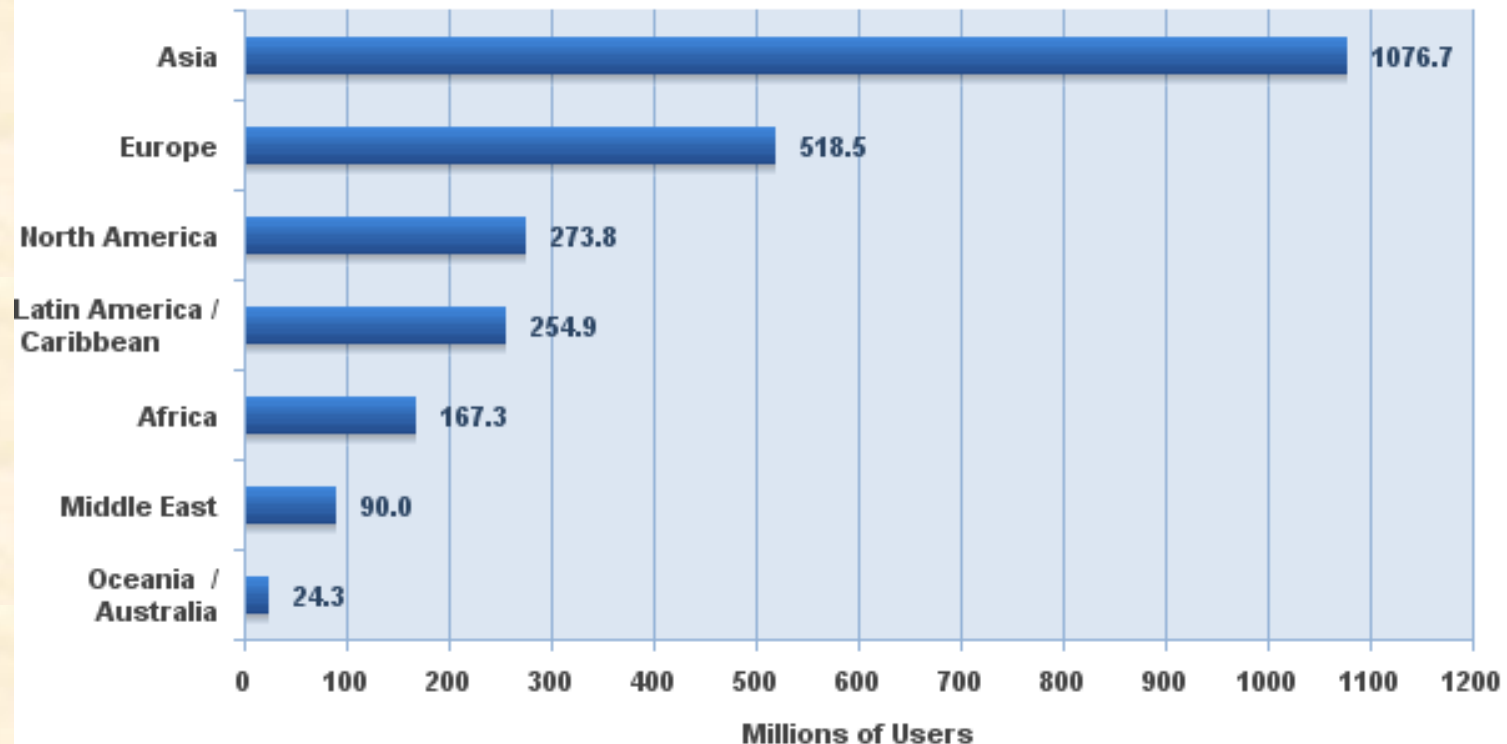
Basis: 2,405,518,376 Internet users on June 30, 2012

Copyright © 2012, Miniwatts Marketing Group



Recent Internet Statistics

**Internet Users in the World
by Geographic Regions - 2012 Q2**



Source: Internet World Stats - www.internetworldstats.com/stats.htm

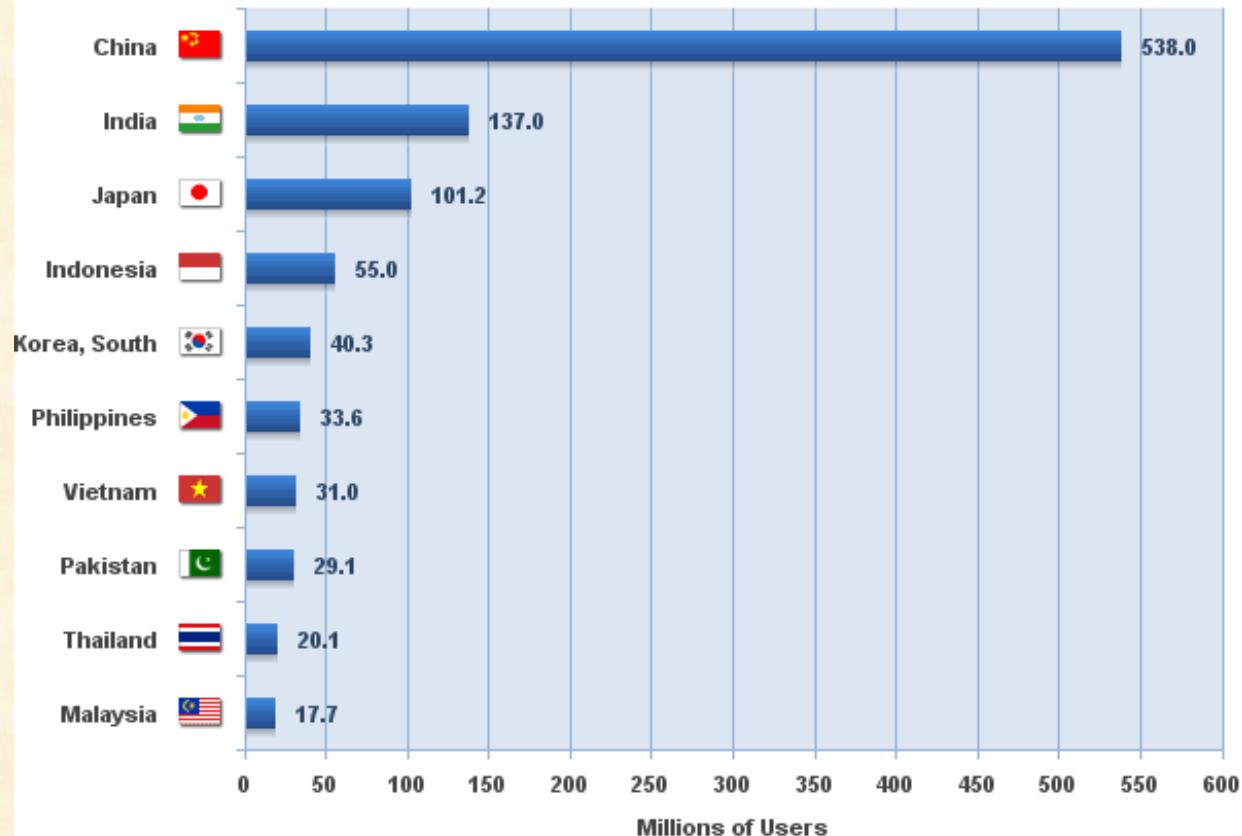
2,405,518,376 Internet users estimated for June 30, 2012

Copyright © 2012, Miniwatts Marketing Group

www.internetworldstats.com/stats.html

Recent Internet Statistics

Asia Top Internet Countries
June 30, 2012



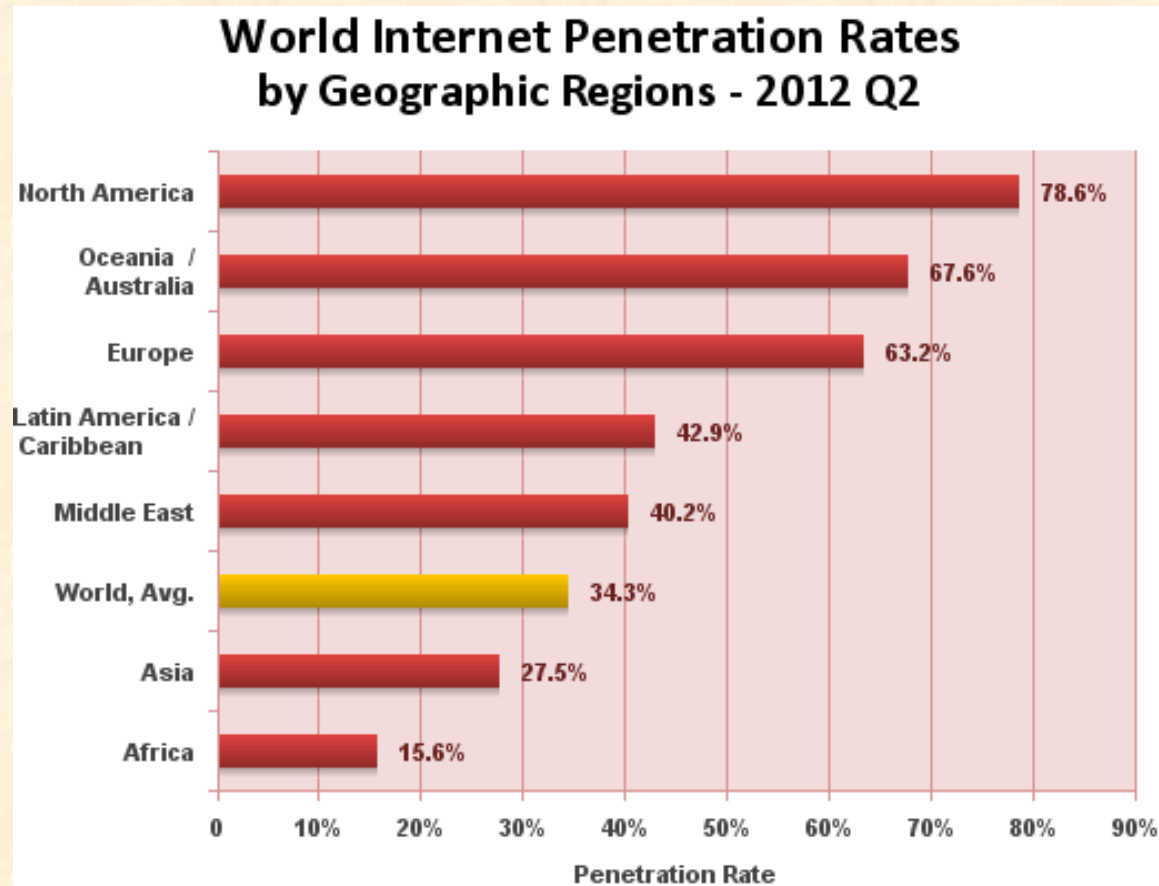
Source: Internet World Stats - www.internetworldstats.com/stats3.htm

2,405,518,376 Internet users in the World estimated for June 30, 2012

Copyright © 2012, Miniwatts Marketing Group

www.internetworldstats.com/stats.html

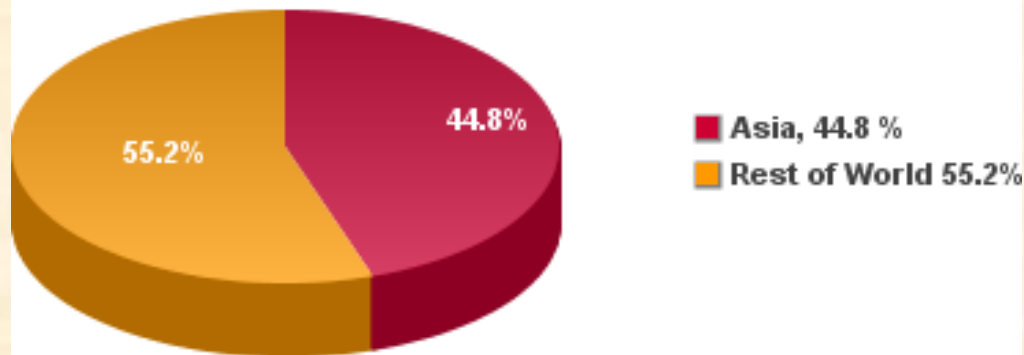
Recent Internet Statistics



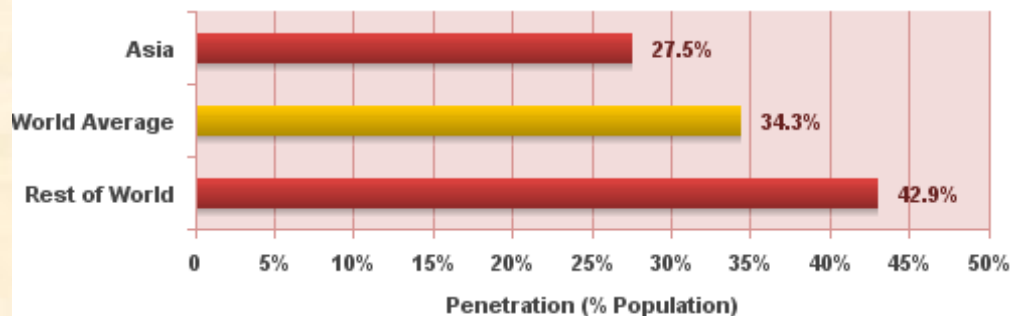
Source: Internet World Stats - www.internetworldstats.com/stats.htm
Penetration Rates are based on a world population of 7,017,846,922
and 2,405,518,376 estimated Internet users on June 30, 2012.
Copyright © 2012, Miniwatts Marketing Group

www.internetworldstats.com/stats.html

Internet Users in Asia 2012 Q2



Internet Penetration in Asia June 30, 2012



Source: Internet World Stats - www.internetworldstats.com/stats3.htm
1,076,681,059 Internet users in Asia estimated for June 30, 2012
Copyright © 2012, Miniwatts Marketing Group

www.internetworldstats.com/stats.html

Recent Internet Statistics

World Population: > 7,100,000,000

Number of Users: > 2,400,000,000

- Internet backbone traffic grew from one Tbit/sec in 1990 to 3,000 Tbit/sec in 1997. It grew > 40% in 2012 alone
- Number of Internet hosts more than doubled every year between 1980-2000

China (in millions): **CNNIC**

Internet users: 618

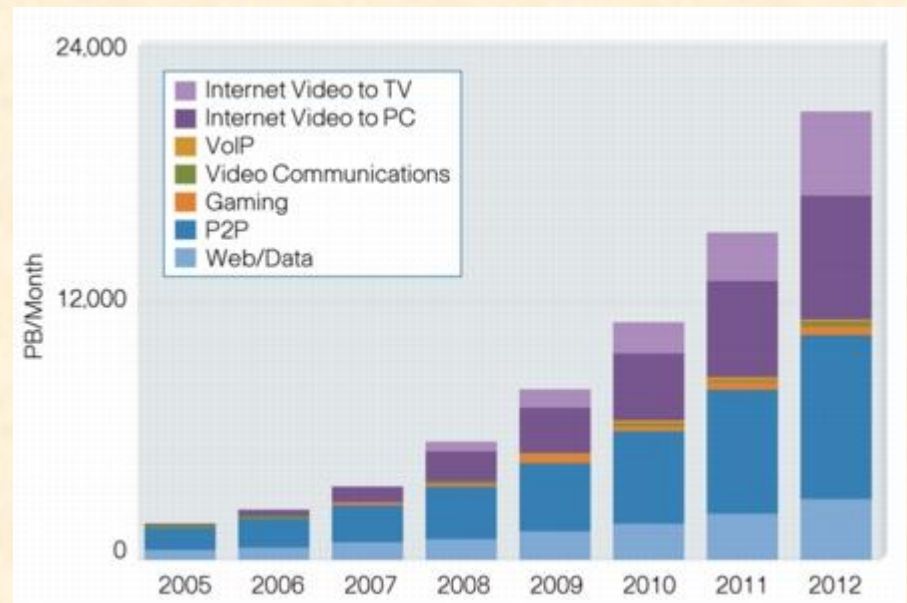
Broadband subscribers: > 450

IPv4 Addresses: 330

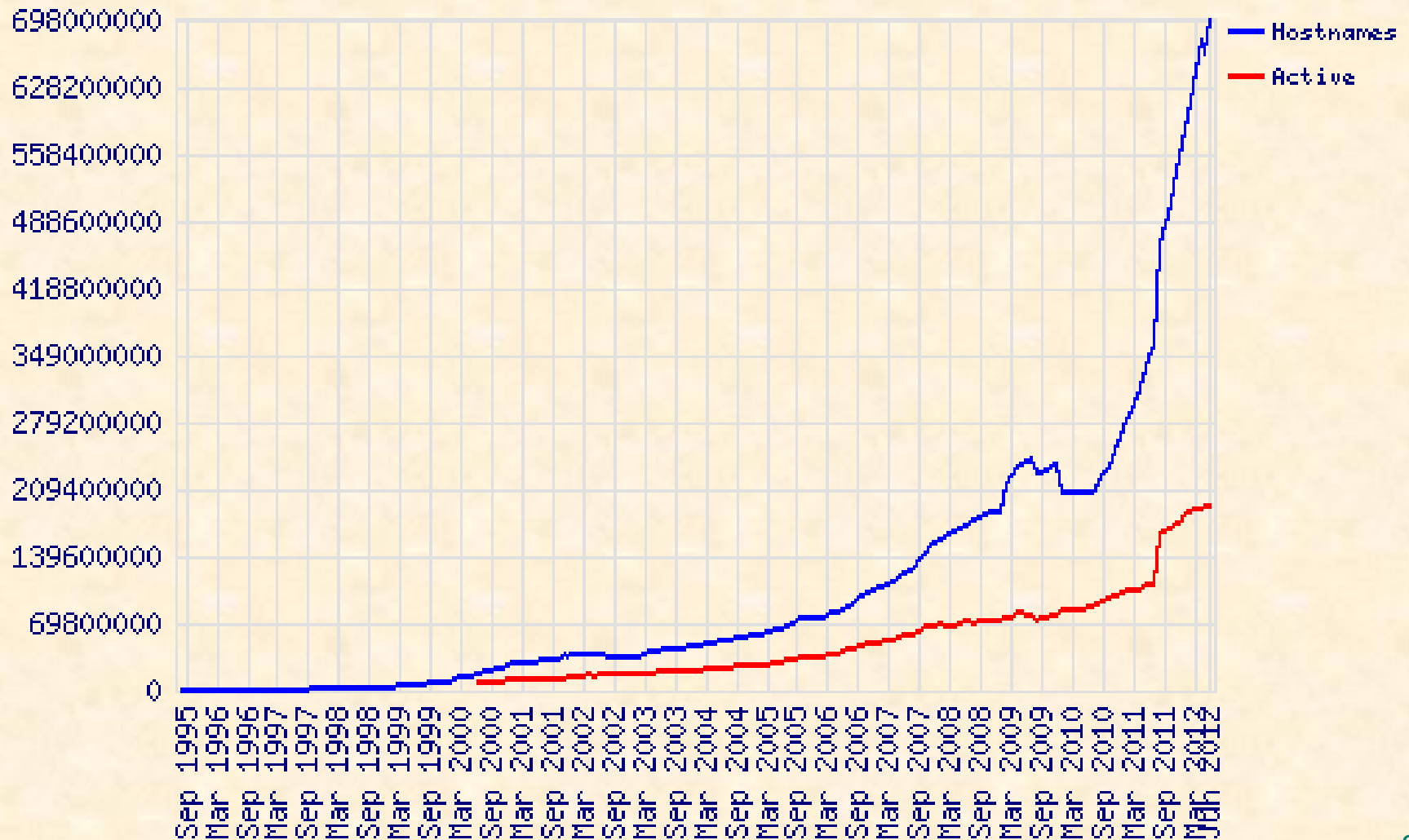
Domain Names: 87.3

International Bandwidth:

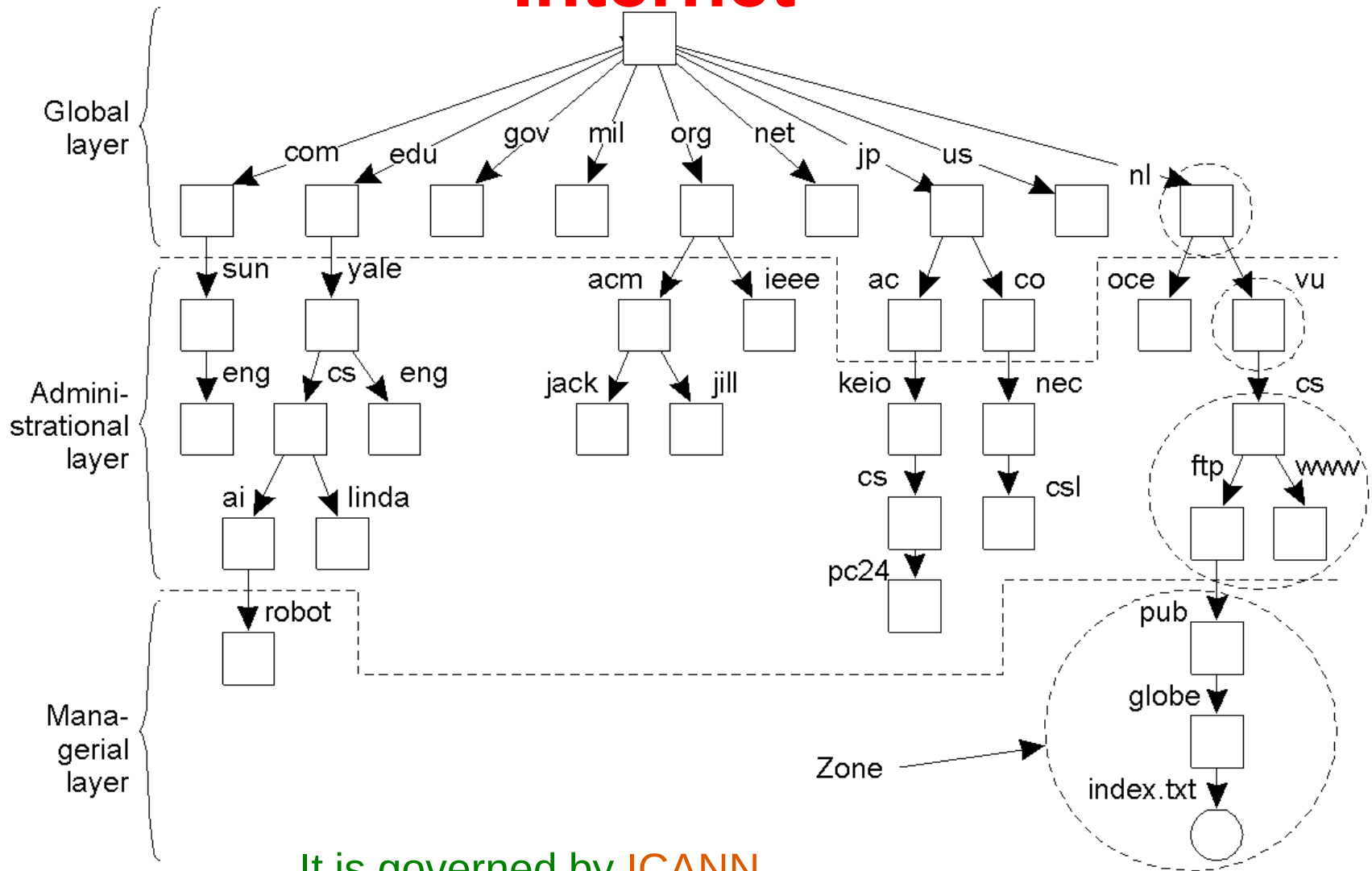
3,406,824 Mbps



Netcraft - an Internet services company based in Bath, England, providing web server and web hosting market-share analysis, including web server and operating system detection.



Internet



It is governed by ICANN

Internet

Applications

Web

FTP

Mail

News

Video

Audio

ping

napster

Transport Protocols

TCP

SCTP

UDP

ICMP

IP

It is governed by ICANN

Ethernet

802.11

Power lines

ATM

Optical

Satellite

Bluetooth

Link Technologies

Internet

Backbone (2001):

27.9% - UUNET/WorldCom/MCI

10.0% - AT&T

6.5% - Sprint

6.3% - Genuity

4.1% - PSINet

3.5% - Cable & Wireless

2.8% - XO Communications

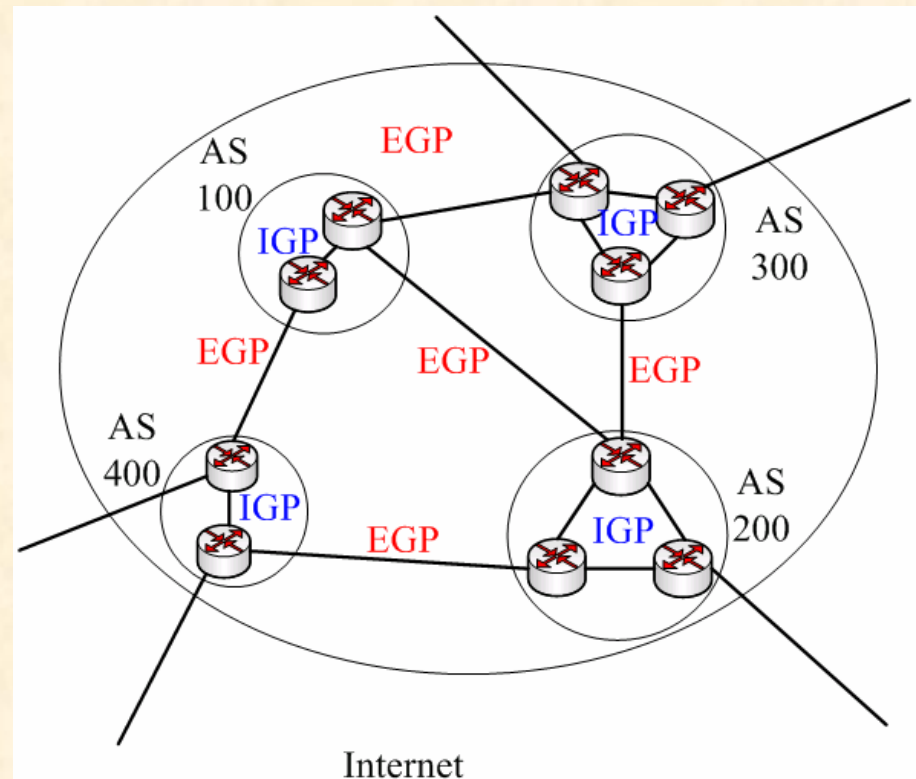
2.6% - Verio

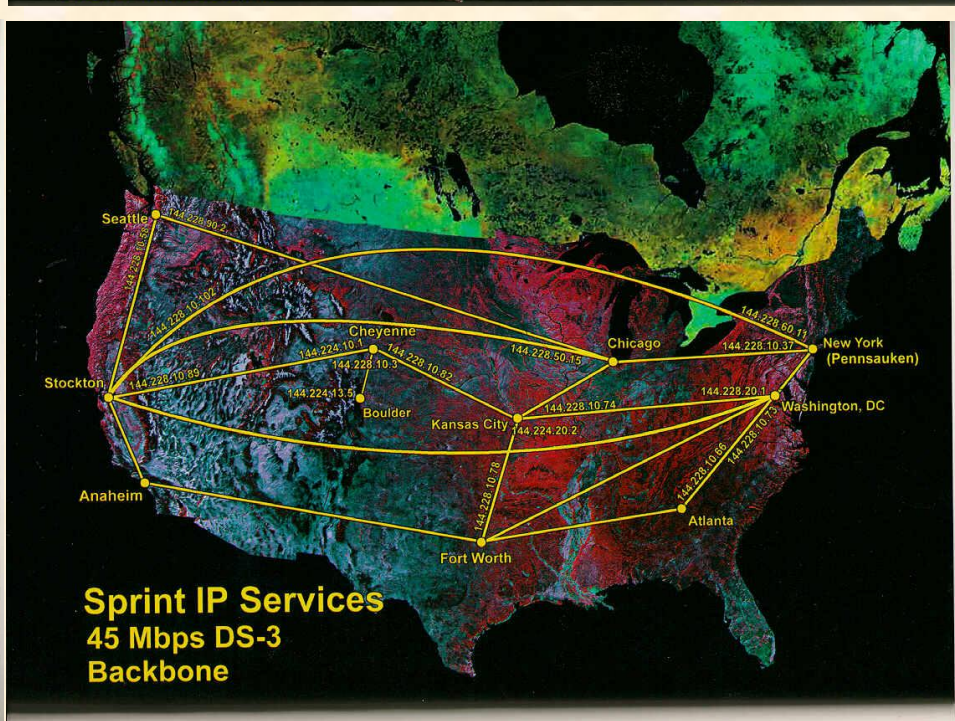
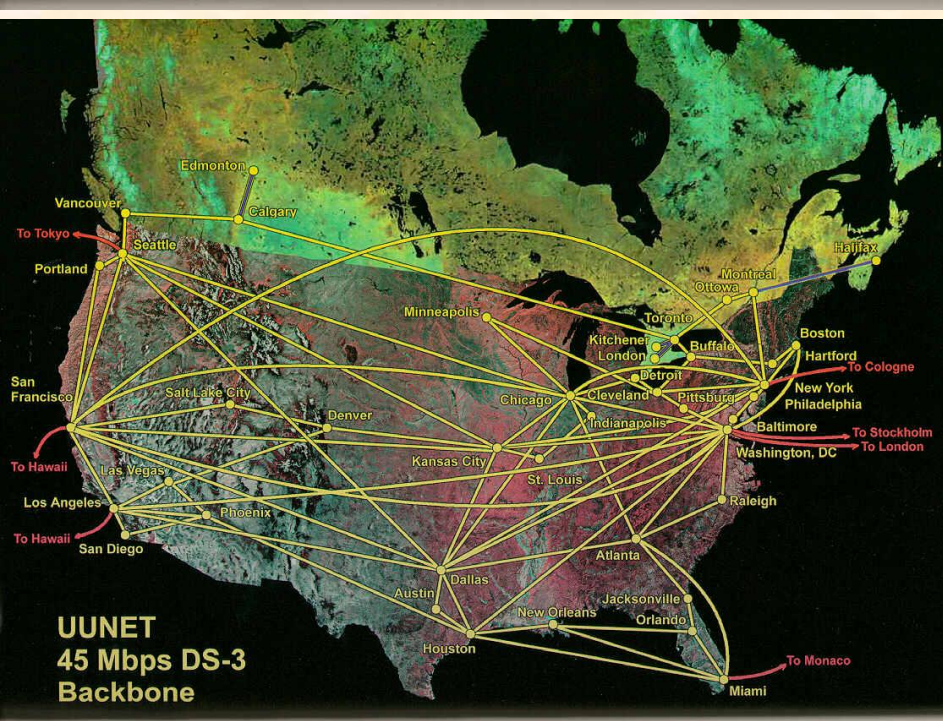
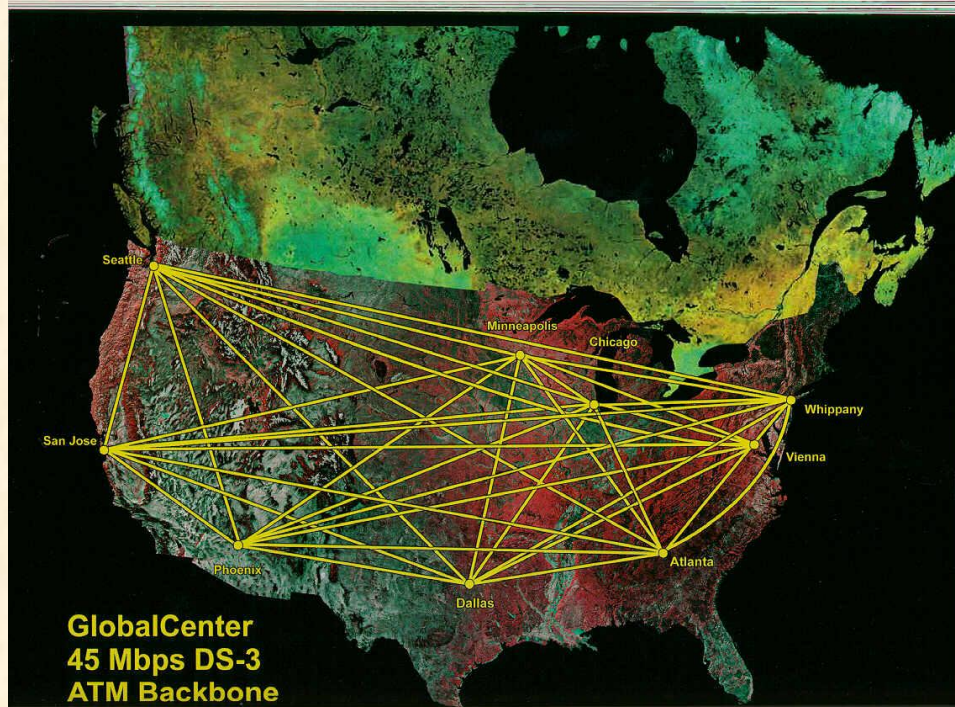
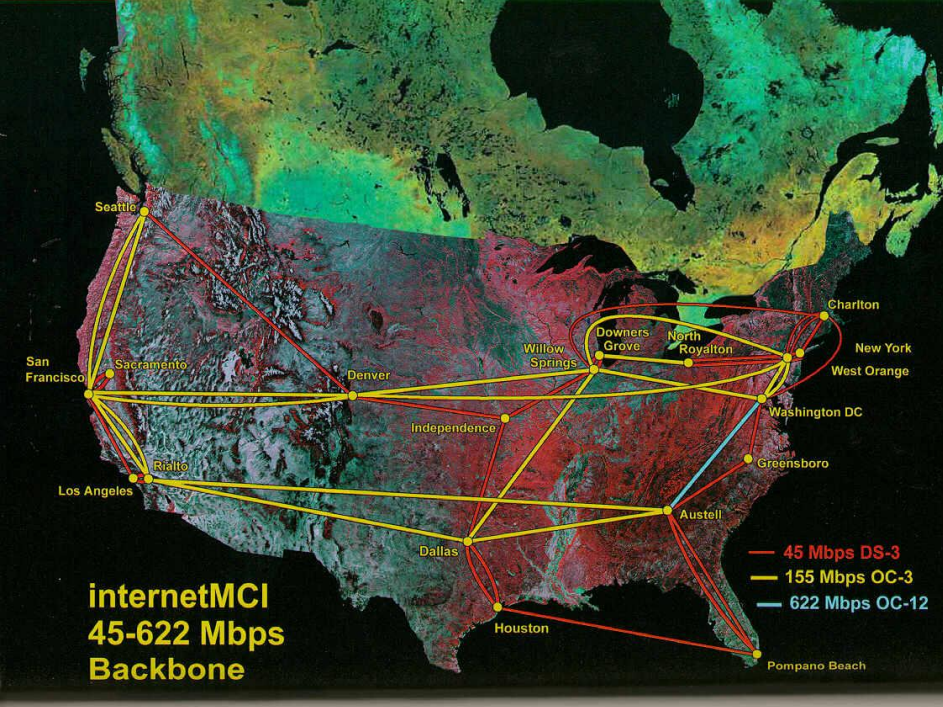
1.5% - Qwest

1.3% - Global Crossing

EGP – Exterior Gateway Protocol

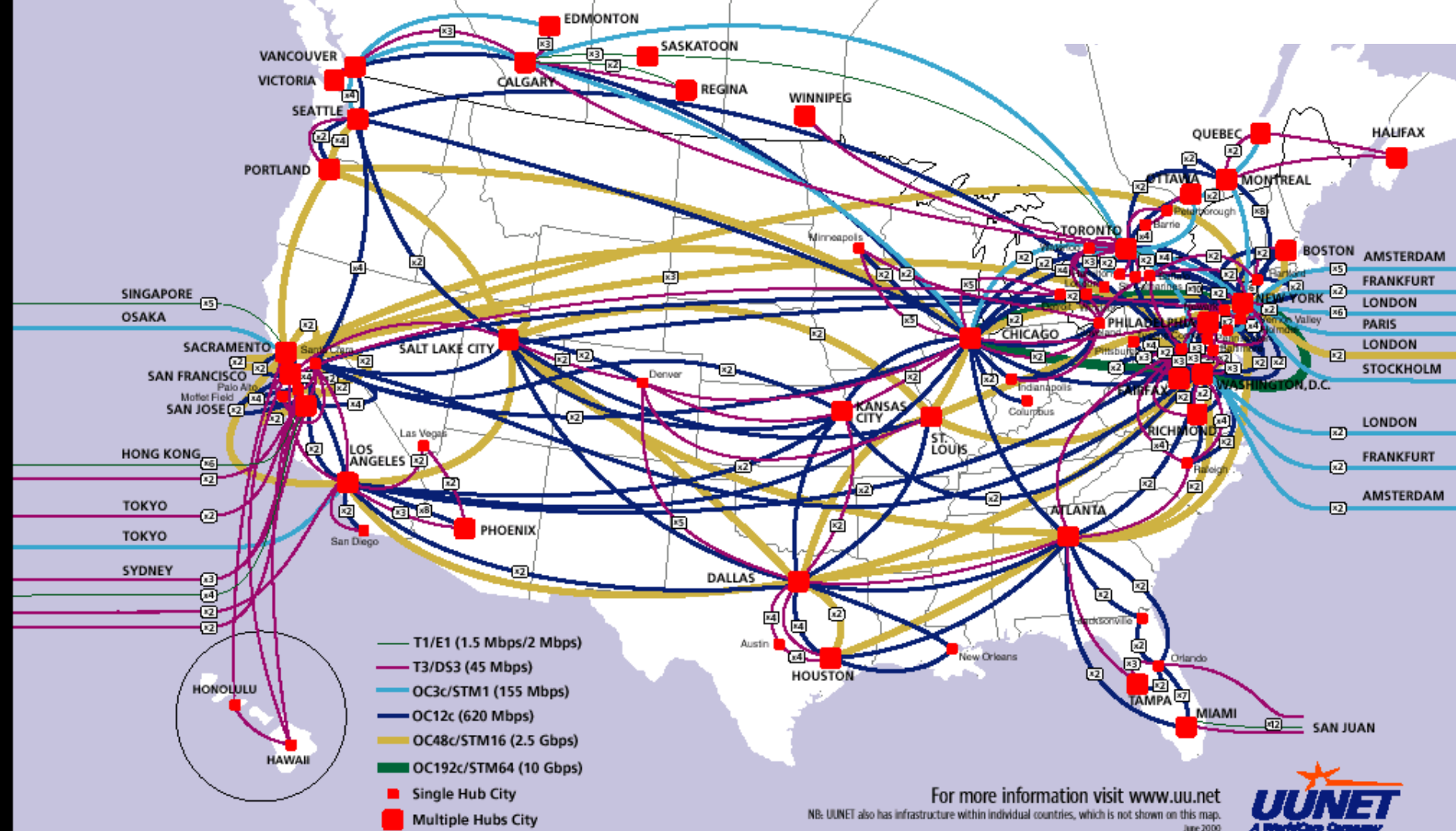
IGP – Interior Gateway Protocol



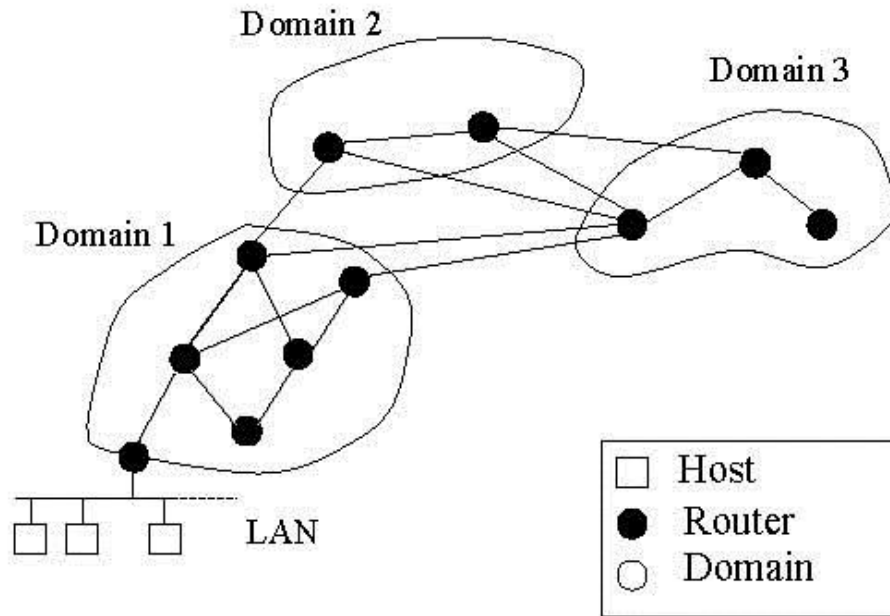


UUNET was founded in 1987: one of the largest Internet service providers and one of the nine Tier-1 networks

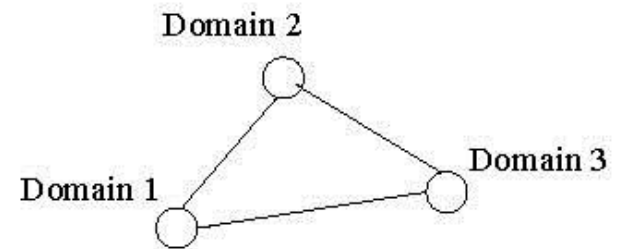
UUNET's North America Internet network



Internet at the AS Level



(a)



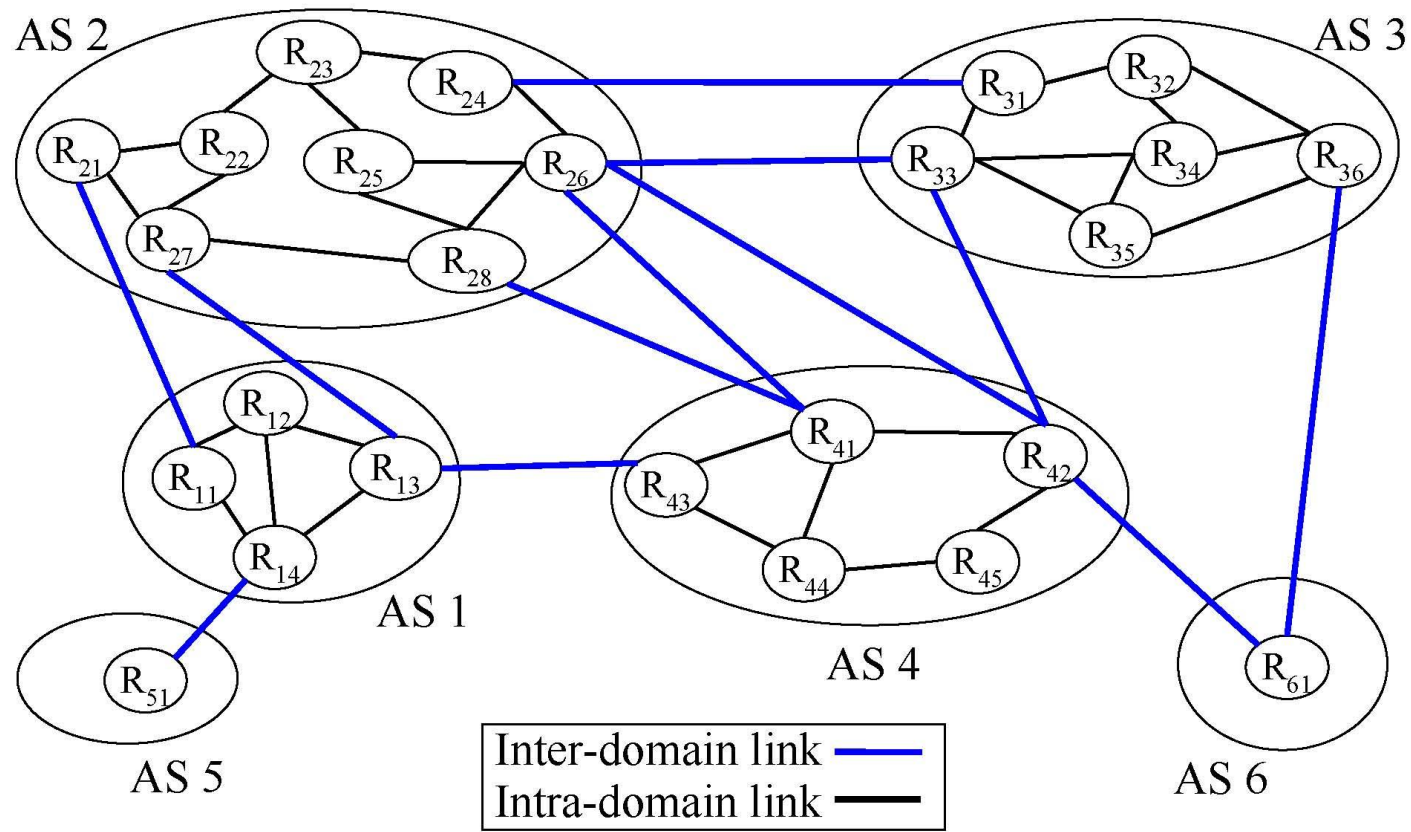
(b)

Internet at the **Autonomous Systems (AS)** level

(Faloutsos brothers, 1999)

Internet at the AS Level

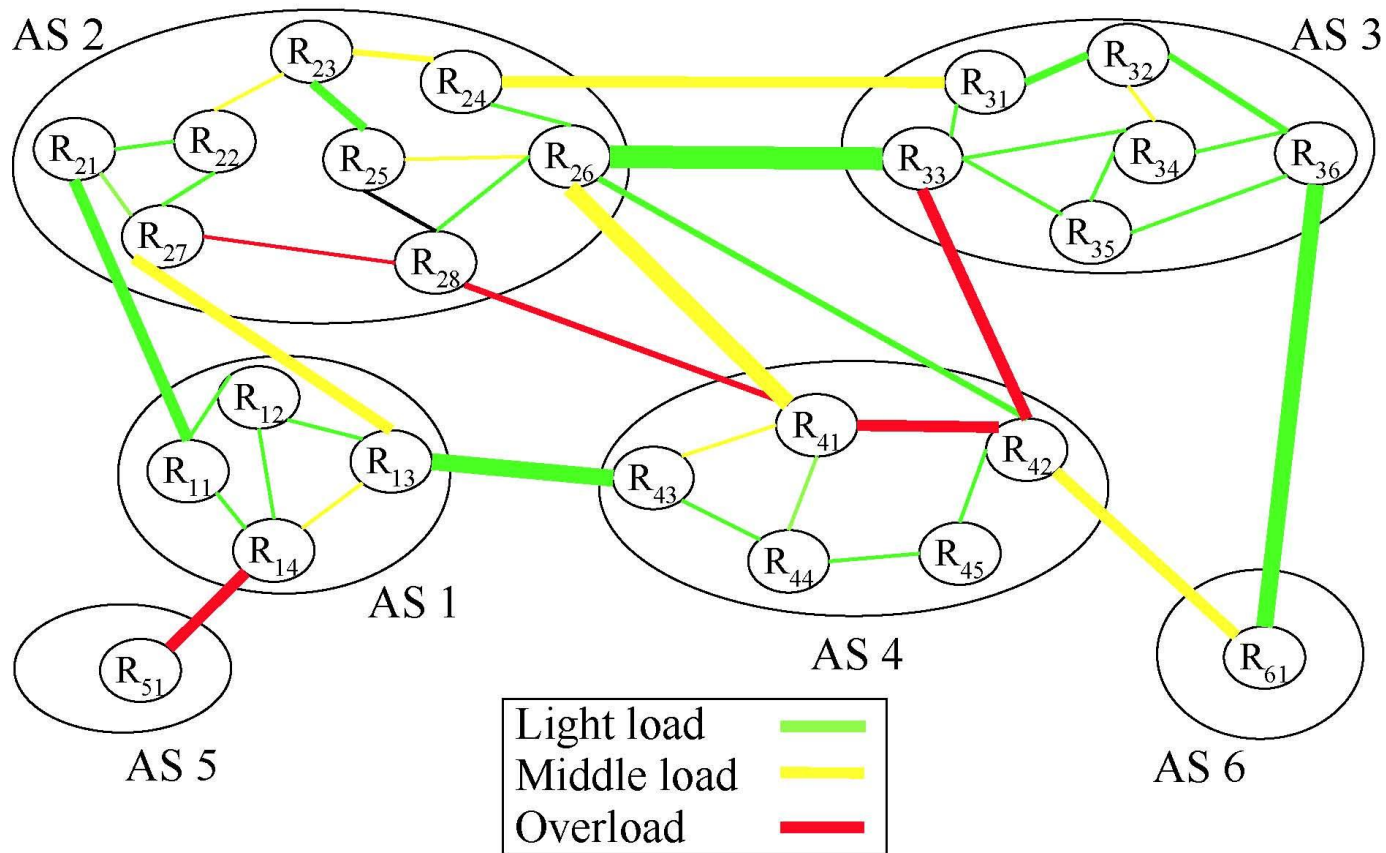
Routing viewpoint



(Steve Uhlig)

Internet at the AS Level

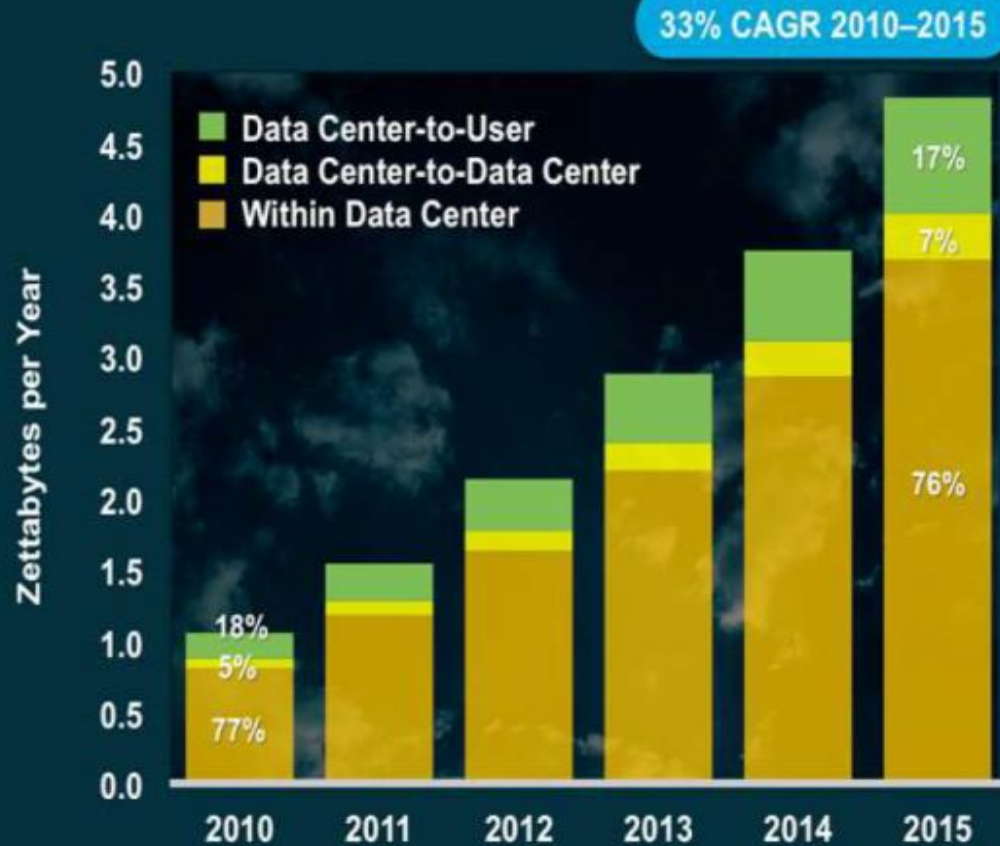
Traffic viewpoint



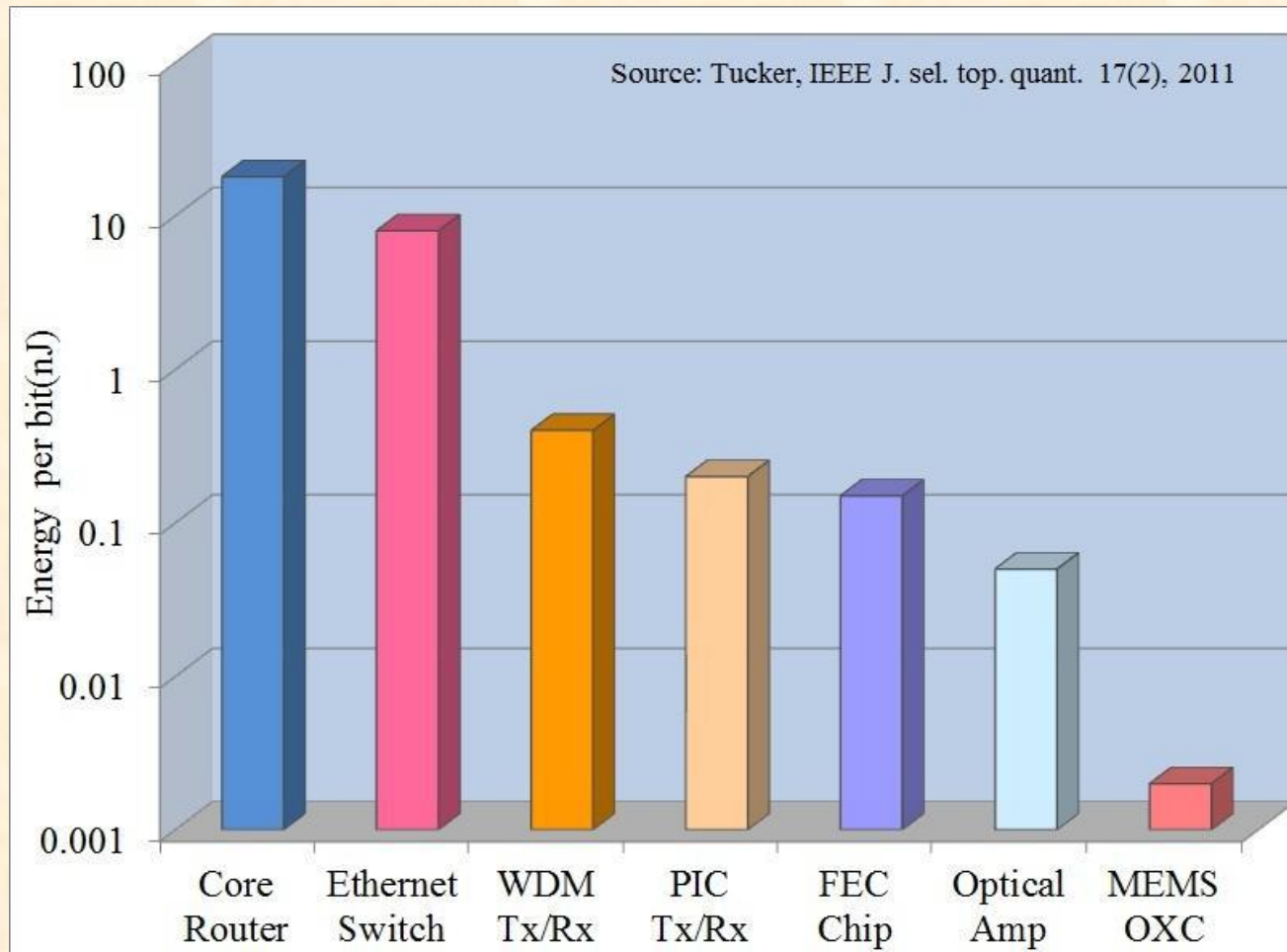
(Steve Uhlig)

Traffic within the Data Center

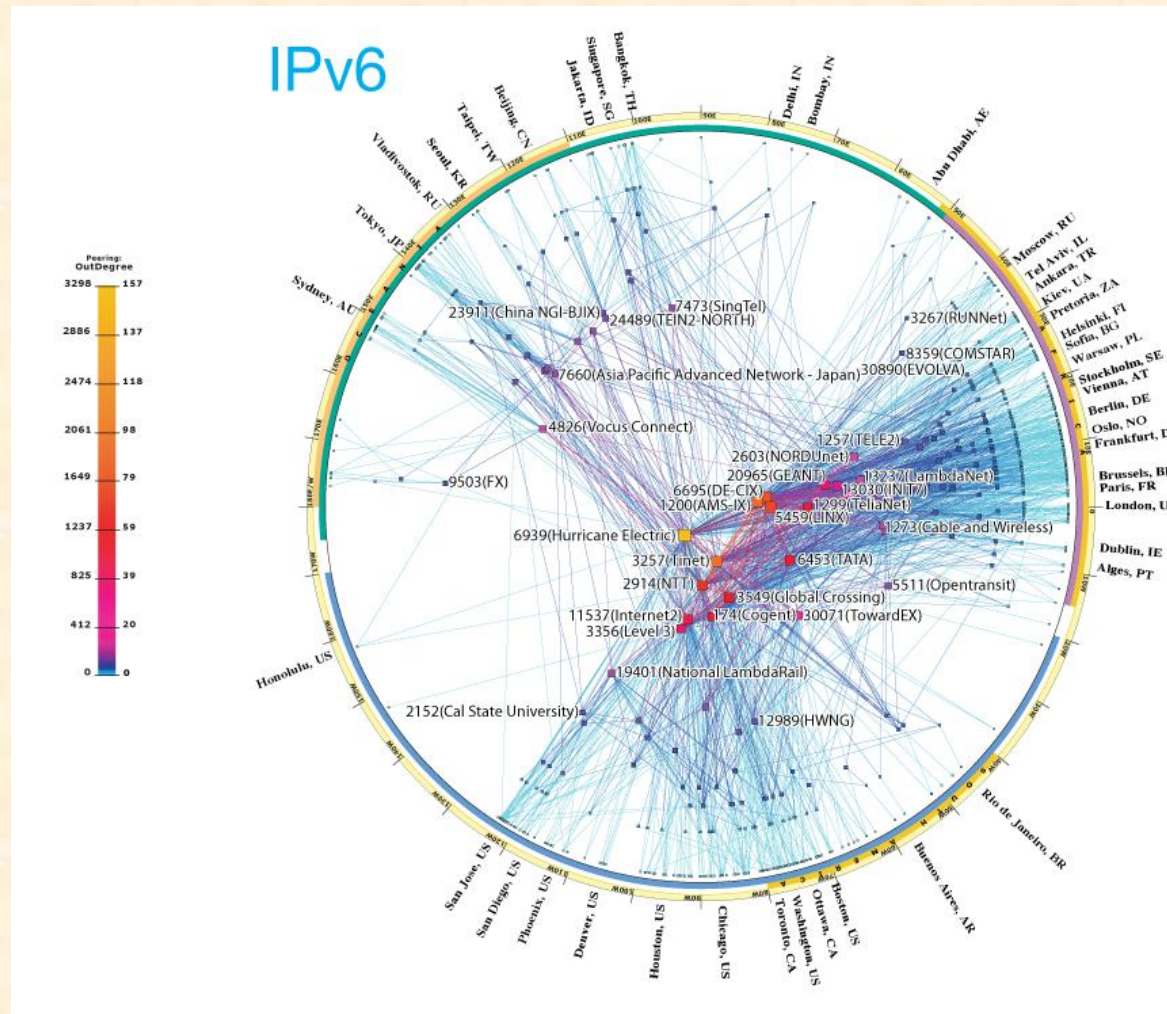
Majority of Data Center Traffic Stays Within the Data Center



Energy per bit consumed by 2010-era equipment



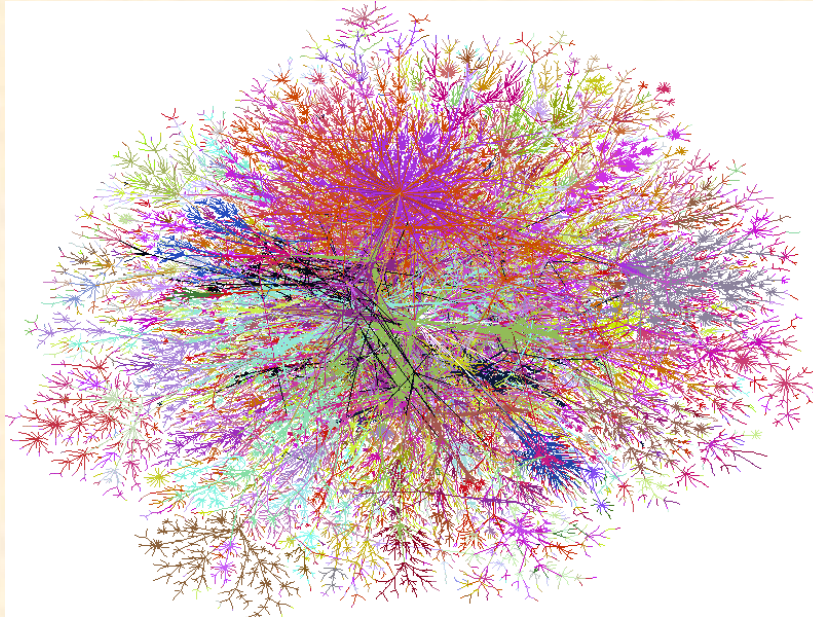
Internet Visualization



(2010)

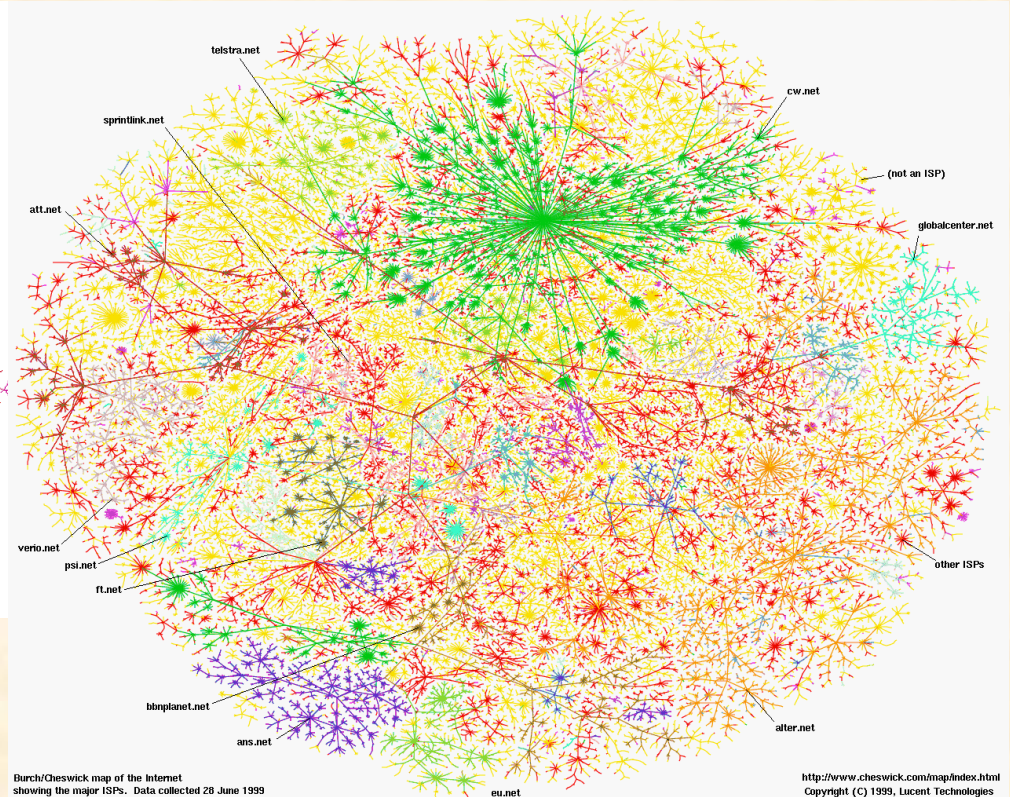
CAIDA Topology Mapping Analysis Team

Internet Visualization



1998

(William R. Cheswick)

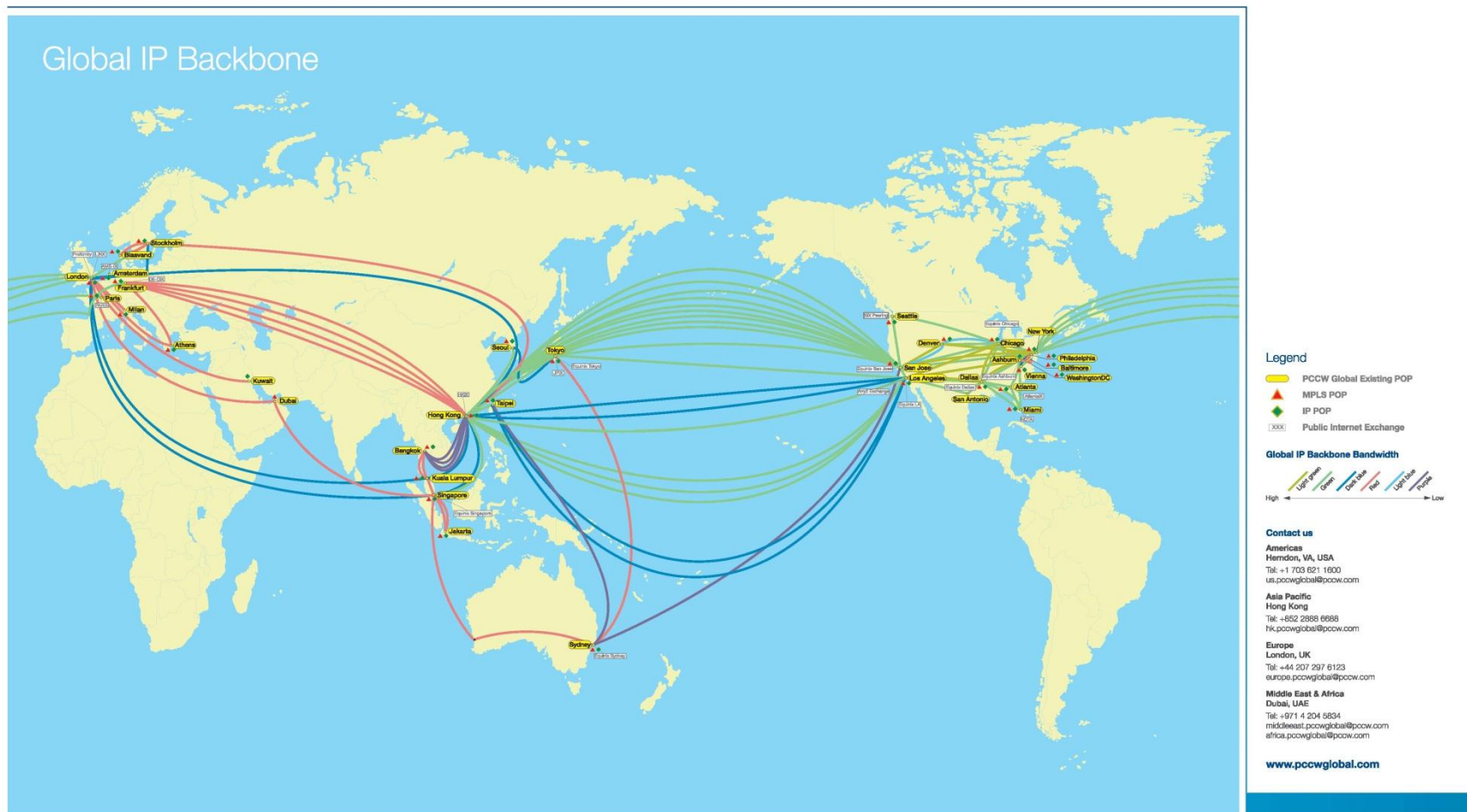


1999

North American Internet Backbone Has 134,855 Routers

IP addresses (router level)

Hong Kong



BREAK

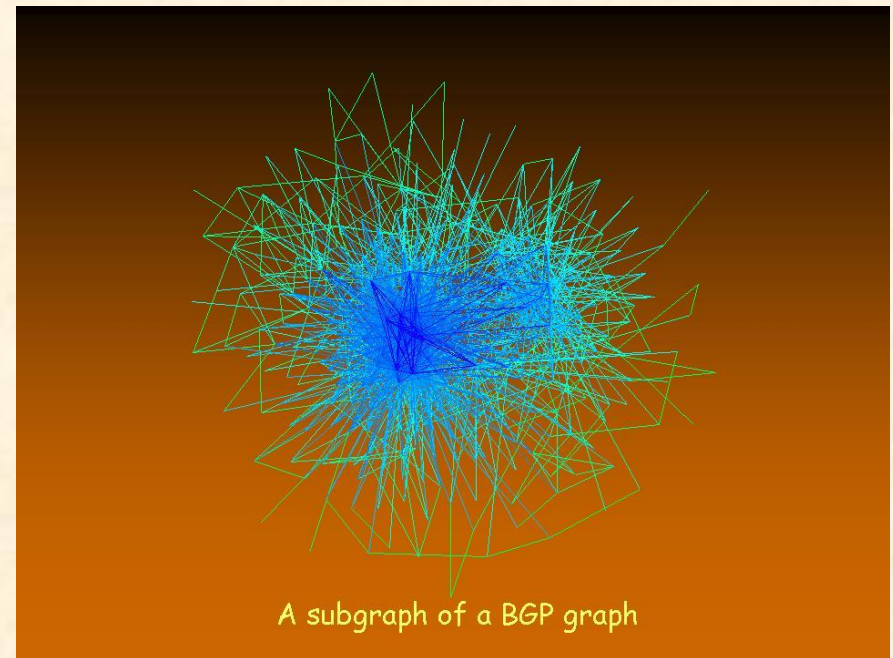
10 minutes

Internet Topology Generators

- ❖ Three stages of development:
 - First generation includes: random topology generators (in 1980s); representative - Waxman generator
 - Second generation includes: structural topology generators (in 1990s); representatives - Tiers and Transit-Stub generators (with hierarchical structures)
 - Third generation (since 2000), based on node degrees; representatives - BRITE and Inet, small-world and especially scale-free network models
- IRL UCLA -- Internet Topology Collection

Internet Topological Properties

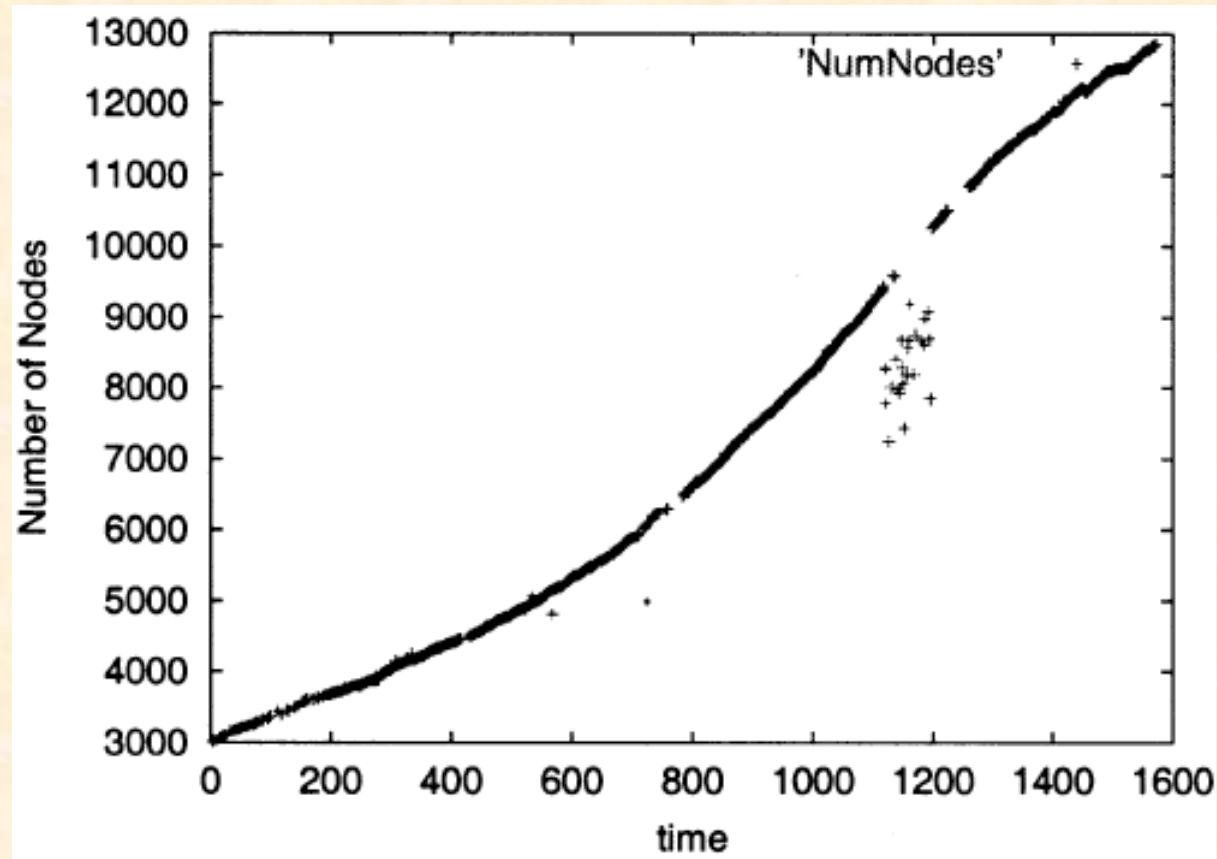
- ❖ Real AS-level Internet data can be obtained from the website of the Oregon Router Views Project, which was managed by the National Laboratory for Applied Network Research (NLNR), expired at the end of 2006, and now managed by the Cooperative Association for Internet Data Analysis (CAIDA)
- ❖ This website is being updated within hours daily by taking snapshots from the routing tables of the Border Gateway Protocol (BGP)



Internet Topological Properties

- ❖ Other useful data about the AS-level Internet:
- Skitter provides Internet topological measures by the CAIDA, using Traceroute (a computer network tool for determining the route taken by packets across an IP network).
- Whois is a domain search tool and data base, identifying the owners and IP addresses of all domains, but it is not automatically managed.
- RIPE NCC supports the infrastructure of the Internet and provides global Internet resources and related services (IPv4, IPv6 and AS Number resources) in Europe, Middle East and parts of Central Asia.

Real AS-Level Internet Data



Numbers of Internet AS (Nov. 1997 – Feb. 2002)

(Siganos et al., 2003)

Power-Law Node-Degree Distributions

Faloutsos M, Faloutsos P, Faloutsos C.

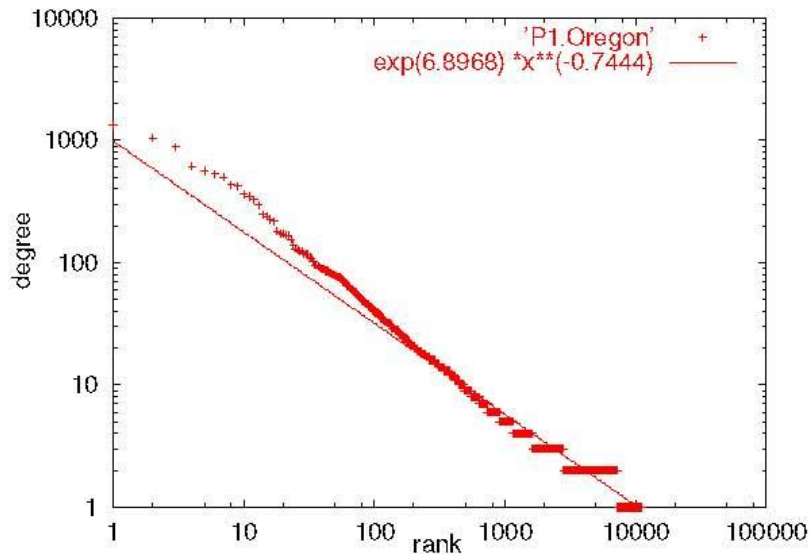
On power-law relationships of the Internet topology.

ACM SIGCOMM Computer Communication Review, 1999, 29(4): 251-262

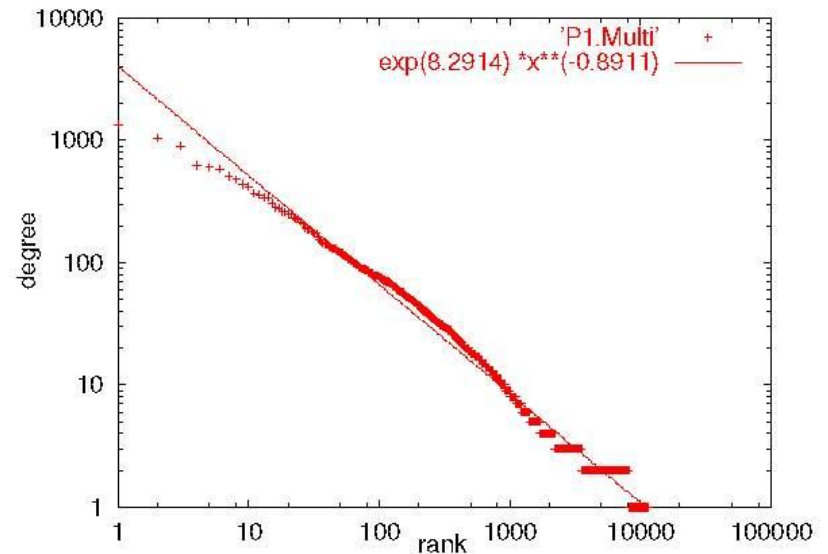


The Faloutsos Brothers

Power-Law Node-Degree Distribution



(a) Oregon



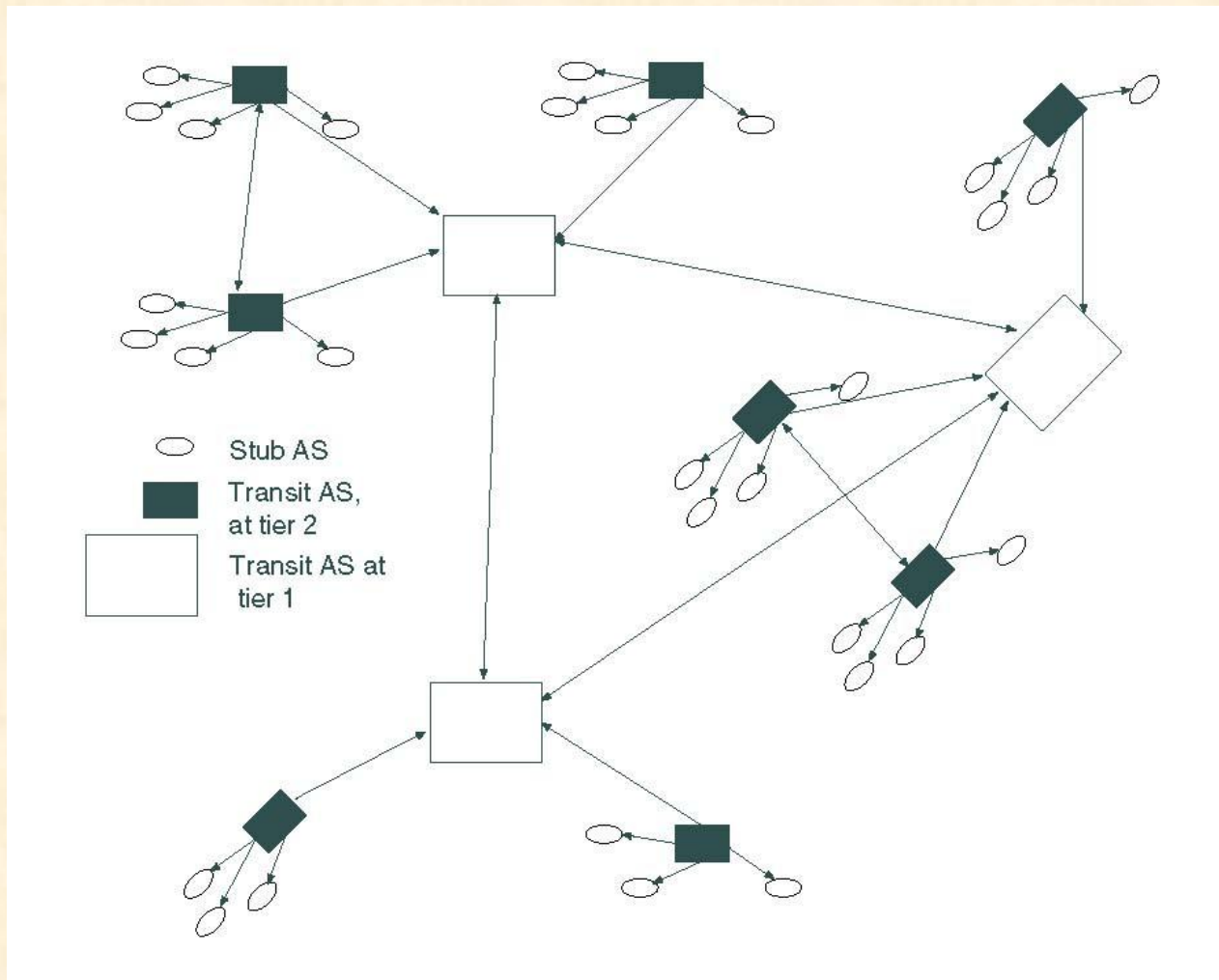
(b) Multi

Power law: $d_v \sim r_v^R$, where d_v is the degree of node v , r_v is the index of node v in the degree-decreasing ordering of all nodes, and $R < 0$ is a rank constant exponent.

Internet Hierarchical Structures

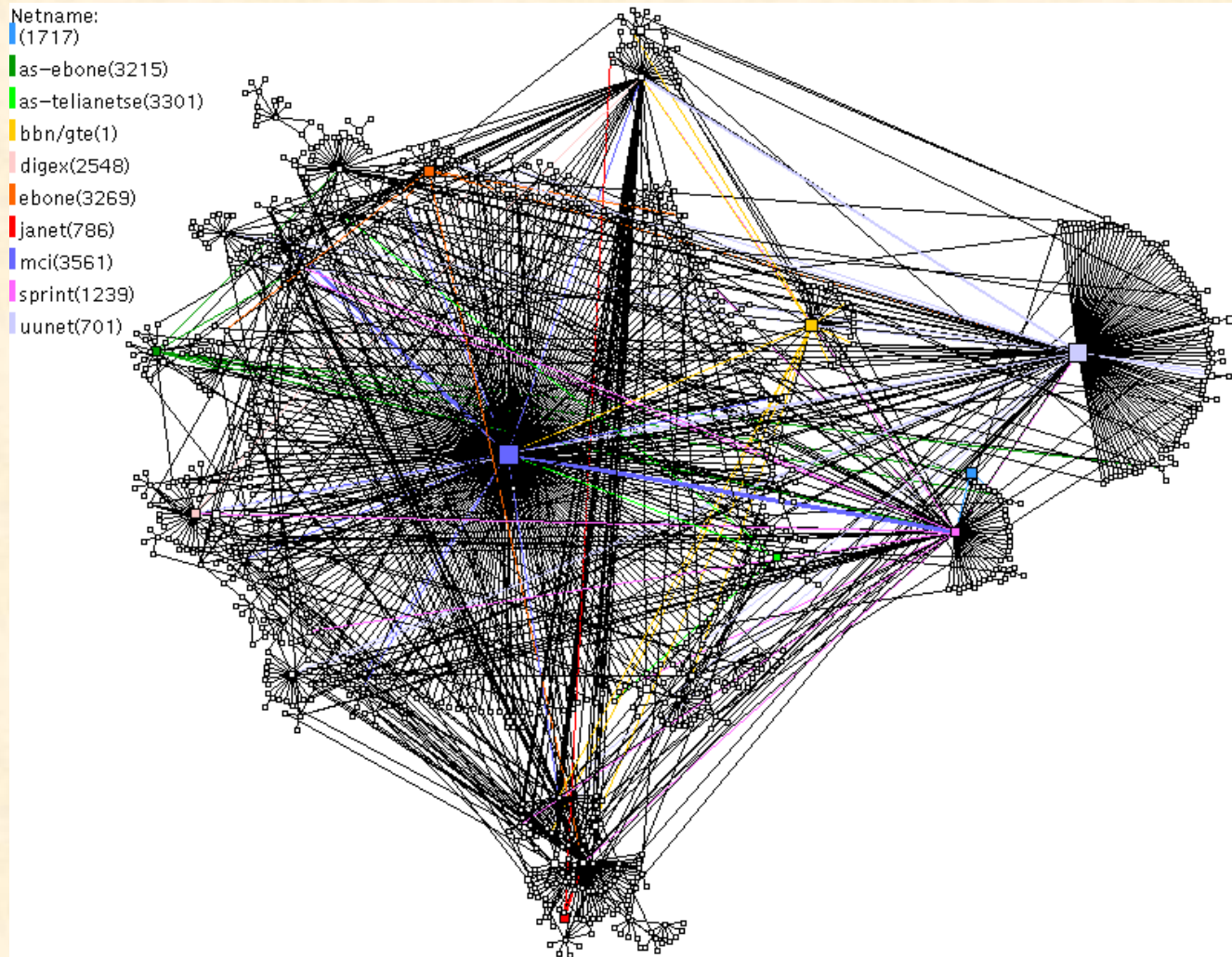
- ❖ Internet consists of a large number of interconnected AS
- ❖ AS may be considered as a *Stub domain* or a *Transit domain*
- ❖ *Transit domain* can be a Metropolitan Area Network (MAN) or a Wide Area Network (WAN), typically a regional or even a national Internet Service Provider (ISP)
- ❖ *Stub domain* consists of campus networks or some other interconnected Local Area Networks (LAN)
- ❖ *Transit domain* is used to link many nearby Stub-domains together
- ❖ *Stub domain* usually only processes the information starting and ending inside the domain, while a *Transit domain* has no such restriction

Internet Hierarchical Structures



AS-level Internet structure (Jaiswal et al., 2004)

Internet Hierarchical Structures

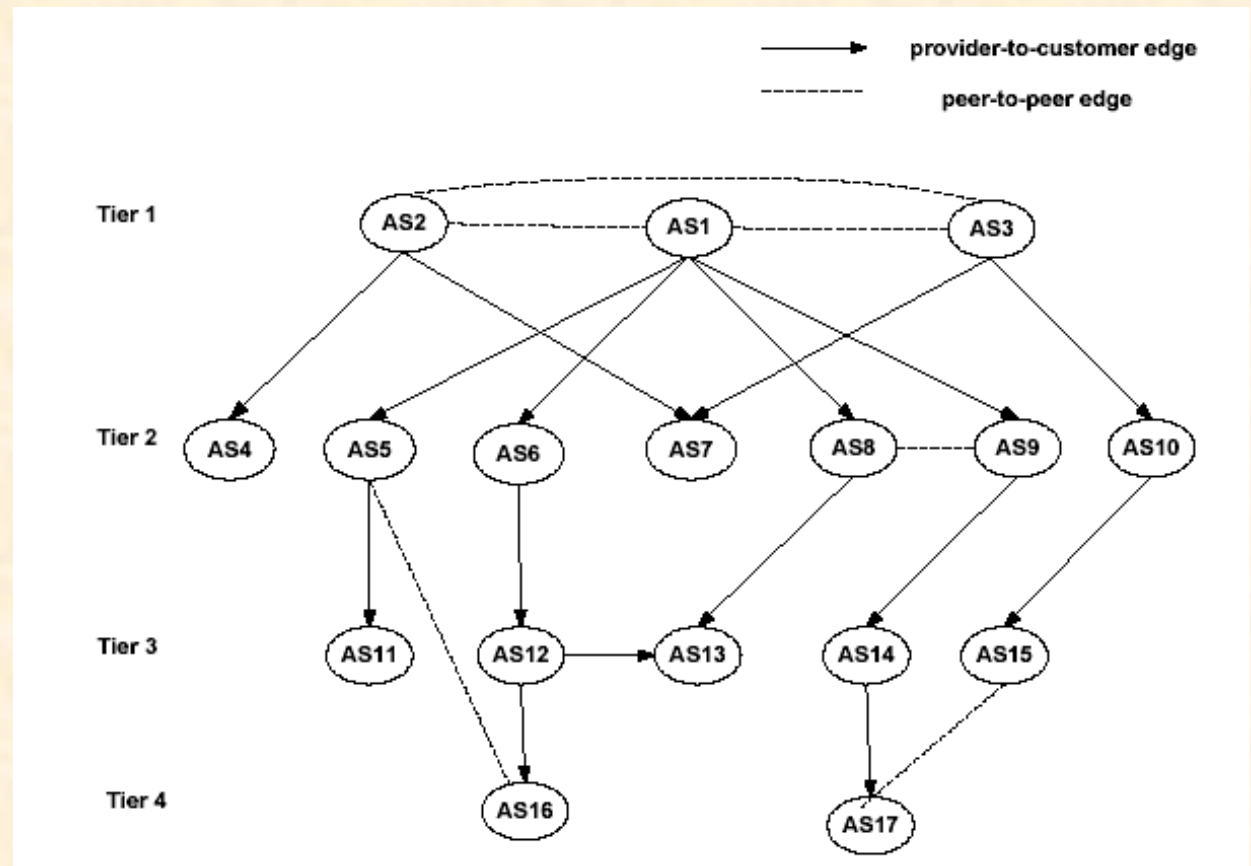


AS-level Internet on 3 December 1998 (generated by Skitter)

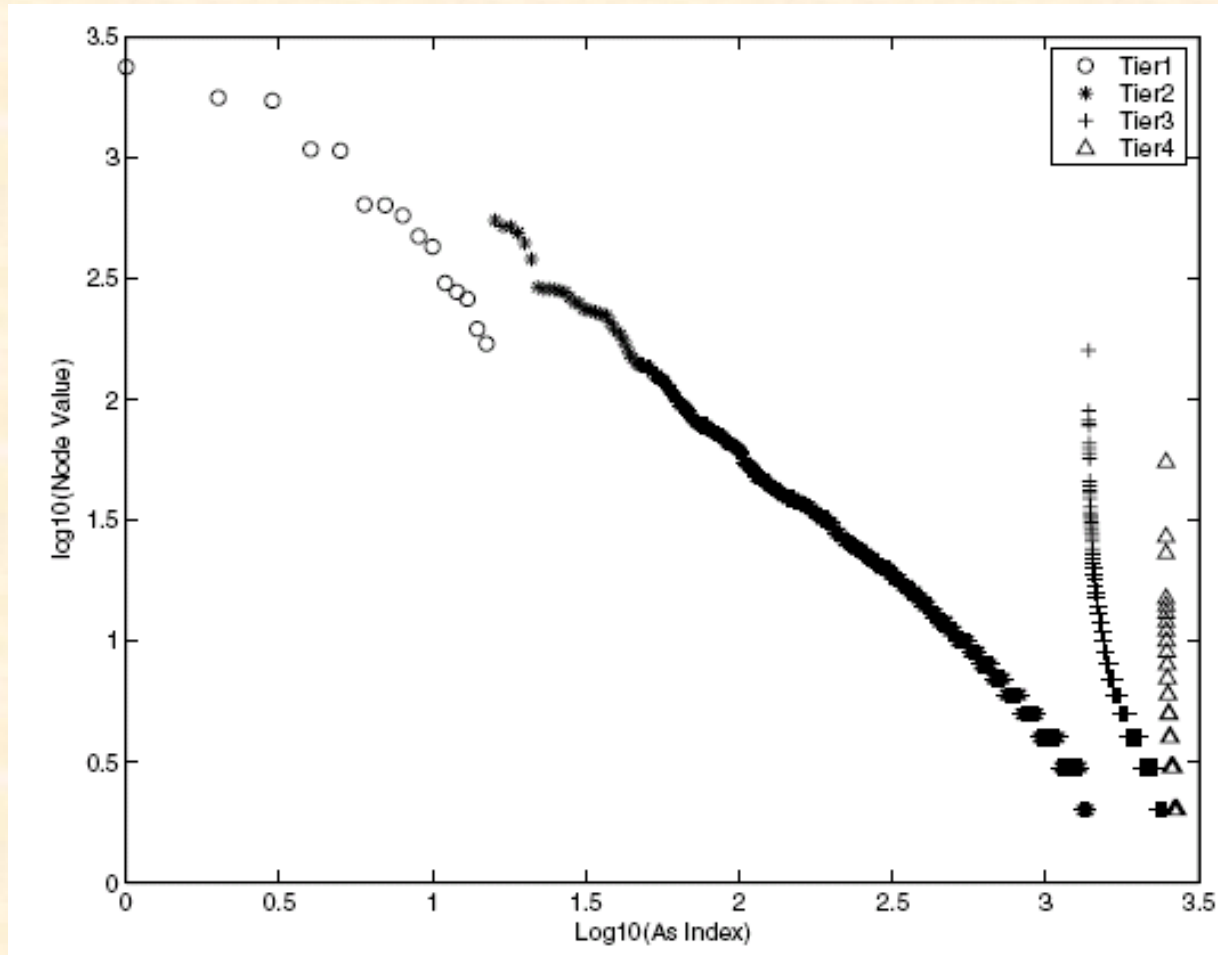
Internet Hierarchical Structures

- ❖ AS on the Internet can be considered as some kind of *Tier*
- ❖ An AS at the highest Tier belongs to the Transit domain, called *Tier-1* provider
- Those Transit and Stub domains at a *lower Tier* depend on the Transit nodes at a *higher Tier* to communicate with the other domains at their same level

(Cai et al., 2004)



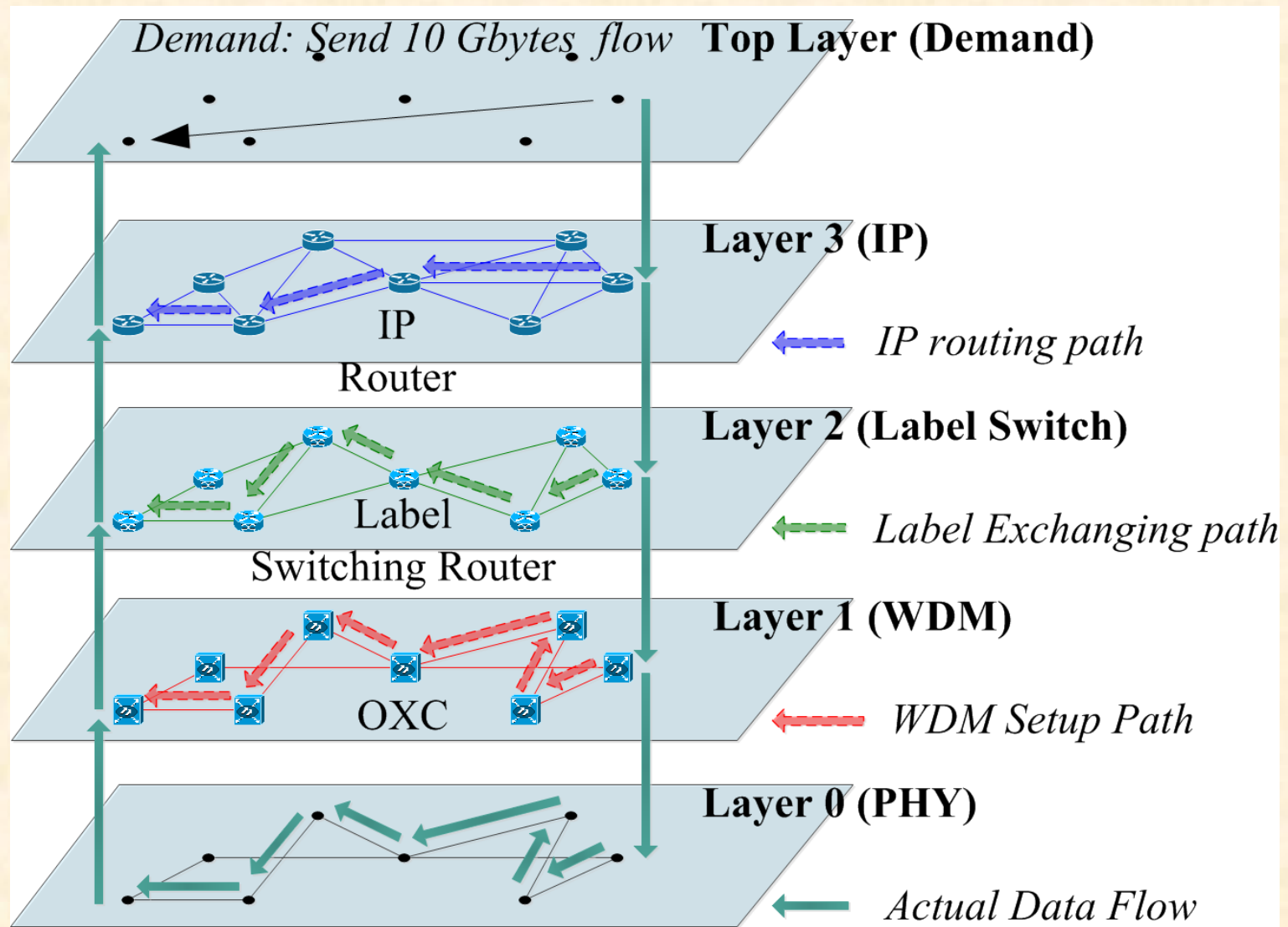
Internet Hierarchical Structures



Degree distribution of the AS-level Internet at different tiers

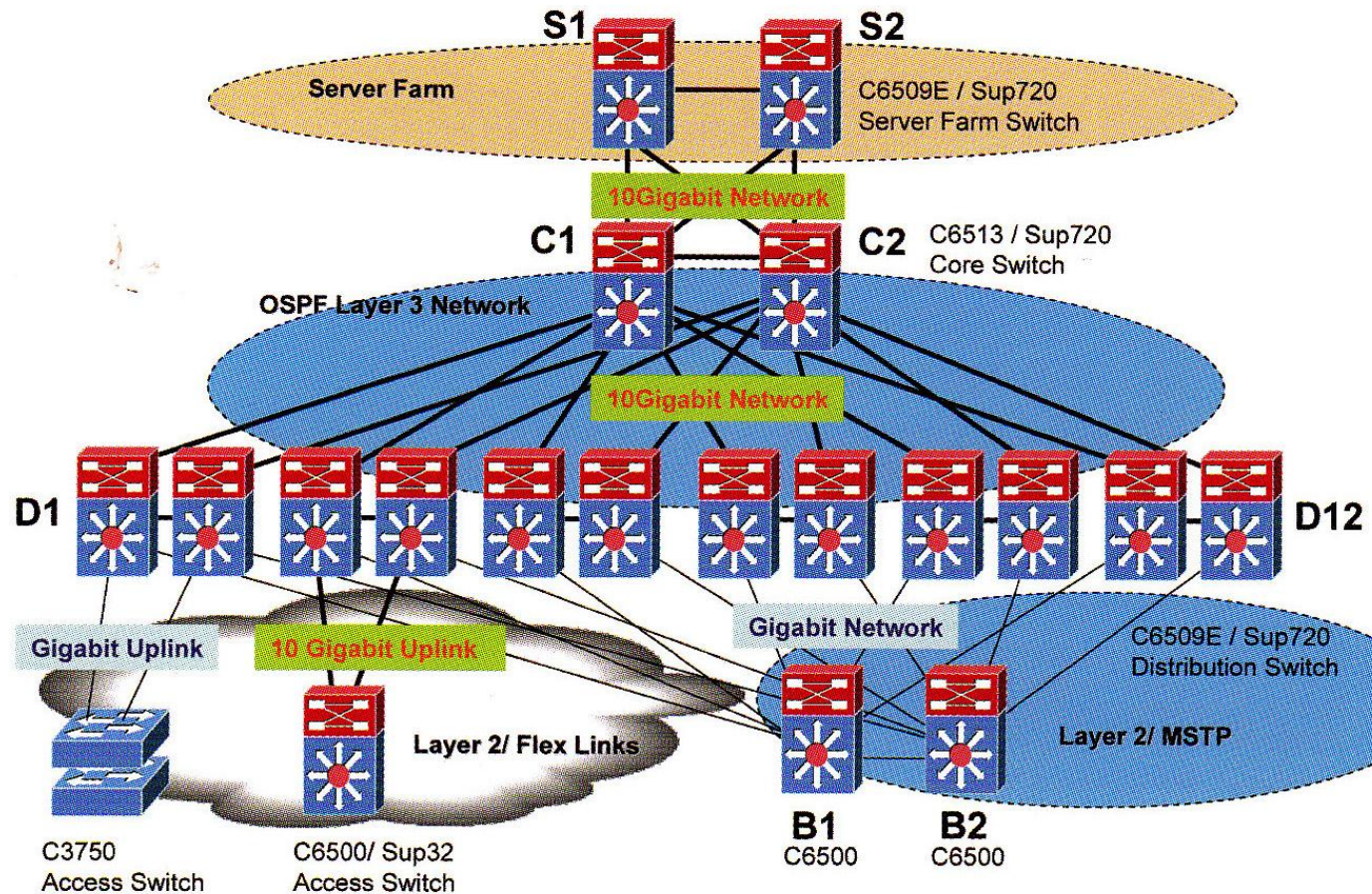
(Jaiswal et al., 2004)

Multi-Layered Network Structure



3-Tier Structure of the CTNET (HK CityU Campus Network, 1985 -)

CityU New Campus Network Diagram



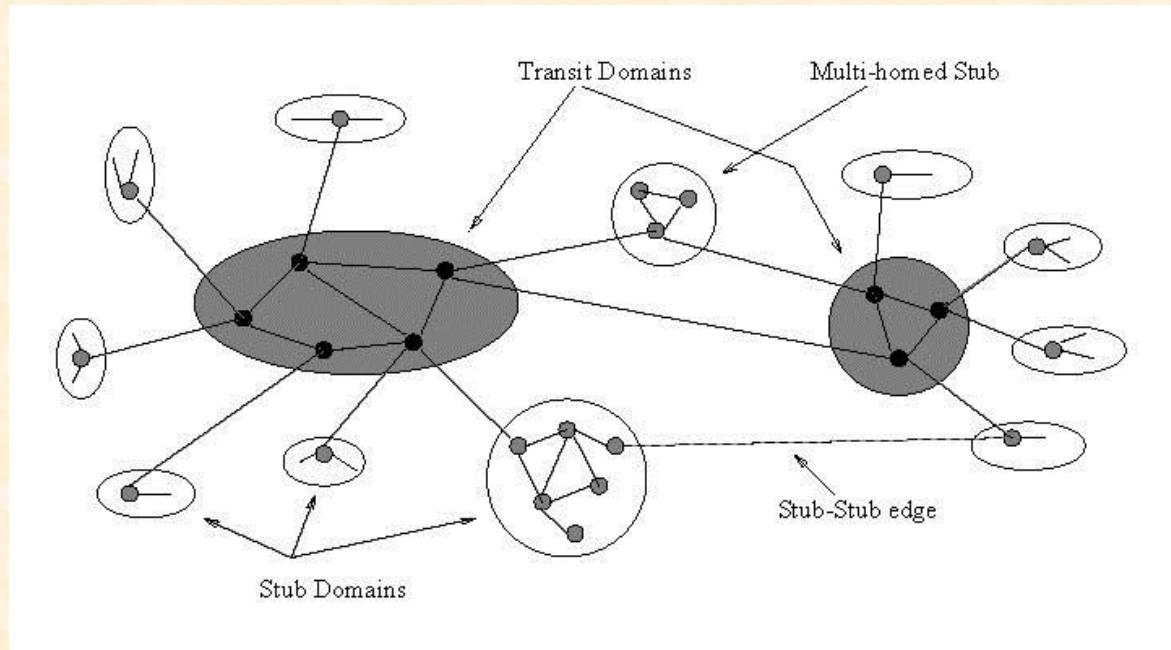
Core
layer

Distribution
layer

Access
layer

Rich-Club Phenomenon

In the Internet, a few nodes have a large number of edges, called **hubs**, and they tend to connect to each other — Rich-club phenomenon



Rich-club phenomenon (Zegura et al., 1997)

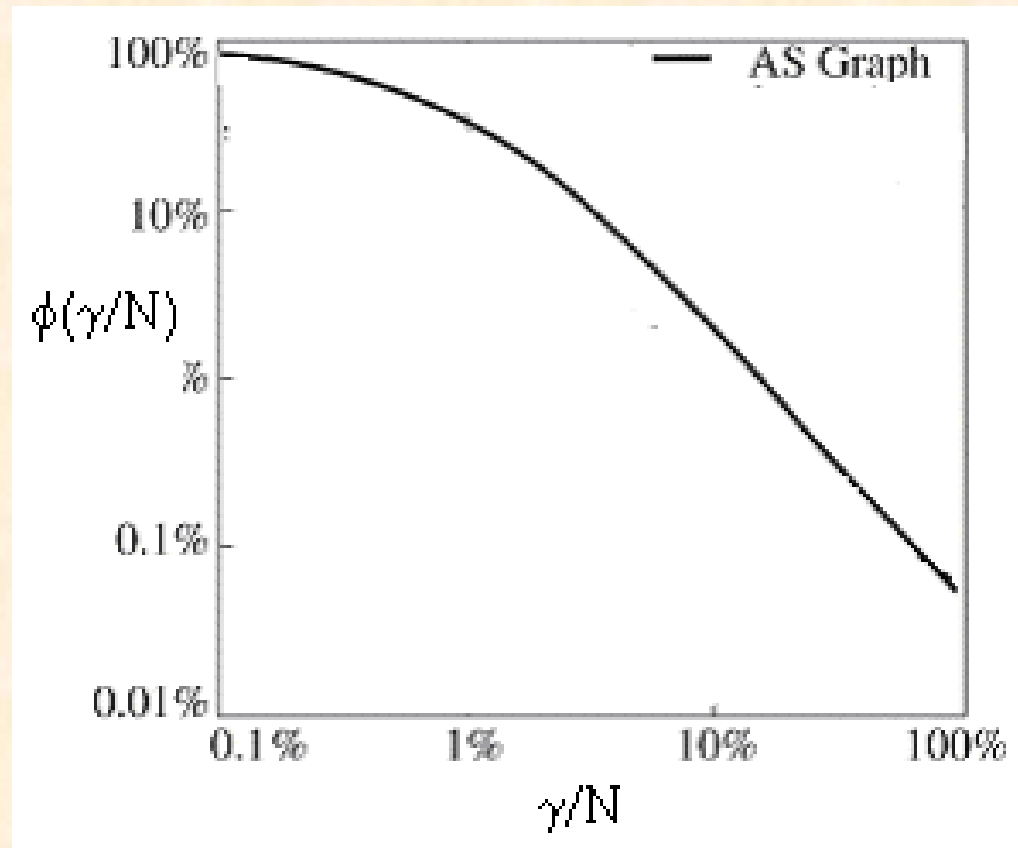
Rich-Club Phenomenon

- ❖ The rich-club phenomenon in an AS layer of size N can be described by the connectivity index $\Phi(r / N)$ of its first r biggest nodes defined by the ratio of the number L of their existing edges versus the number $r(r-1)/2$ of all possible edges among them:

$$\Phi(r / N) = \frac{L}{r(r-1)/2} = \frac{2L}{r(r-1)}$$

- ❖ If $\Phi(r / N) = 1$, then the first r biggest nodes compose a fully-connected sub-network.
- ❖ Some real data have verified that the connectivity index follows nearly a power-law form.

Rich-Club Phenomenon



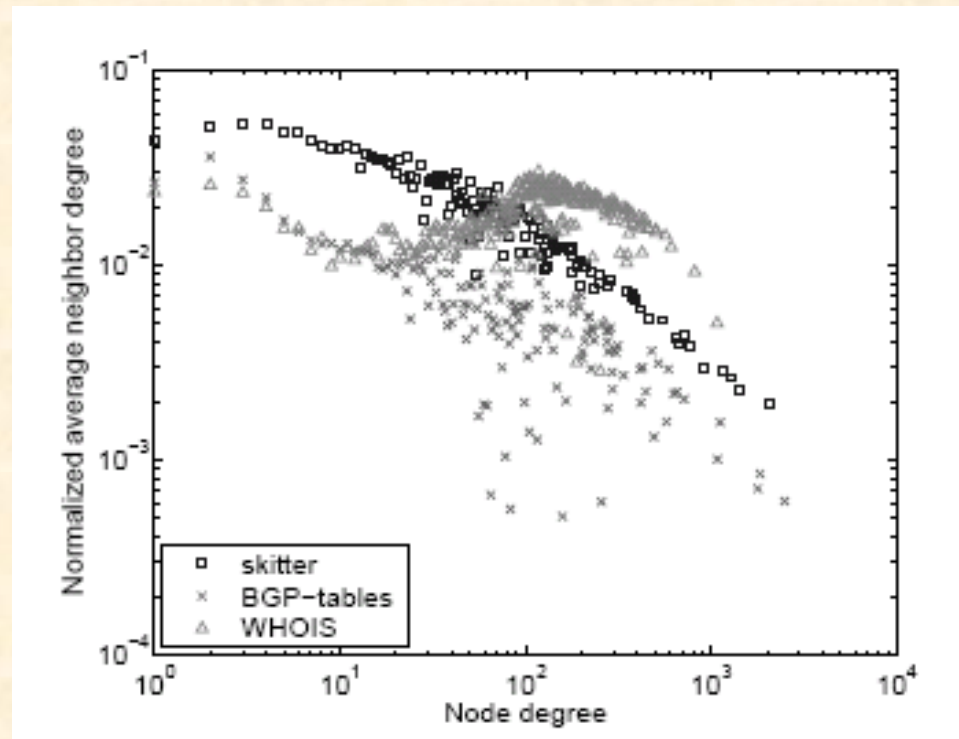
Rich-club phenomenon in the AS-level Internet (Zegura et al., 1997)

Yet, most nodes are small in the Internet

- ❖ In the Internet, hubs are well interconnected (rich clubs).
- ❖ Most neighbors of a hub typically have small degrees.
- ❖ Analysis on the Internet data of April 2002 shows that nodes with degrees of 1, 2, and 3 were 26%, 38% and 14%, respectively, which sums up to about 80% of the whole network.

(Zhou and Mondragon, 2004)

(Mahadevan et al., 2005)



(Dis)Assortativity

Definition. The *Assortativity Coefficient* of a network is defined by

$$r = \frac{M^{-1} \sum_i j_i k_i - \left[M^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2}{M^{-1} \sum_i \frac{1}{2} (j_i^2 + k_i^2) - \left[M^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2}$$

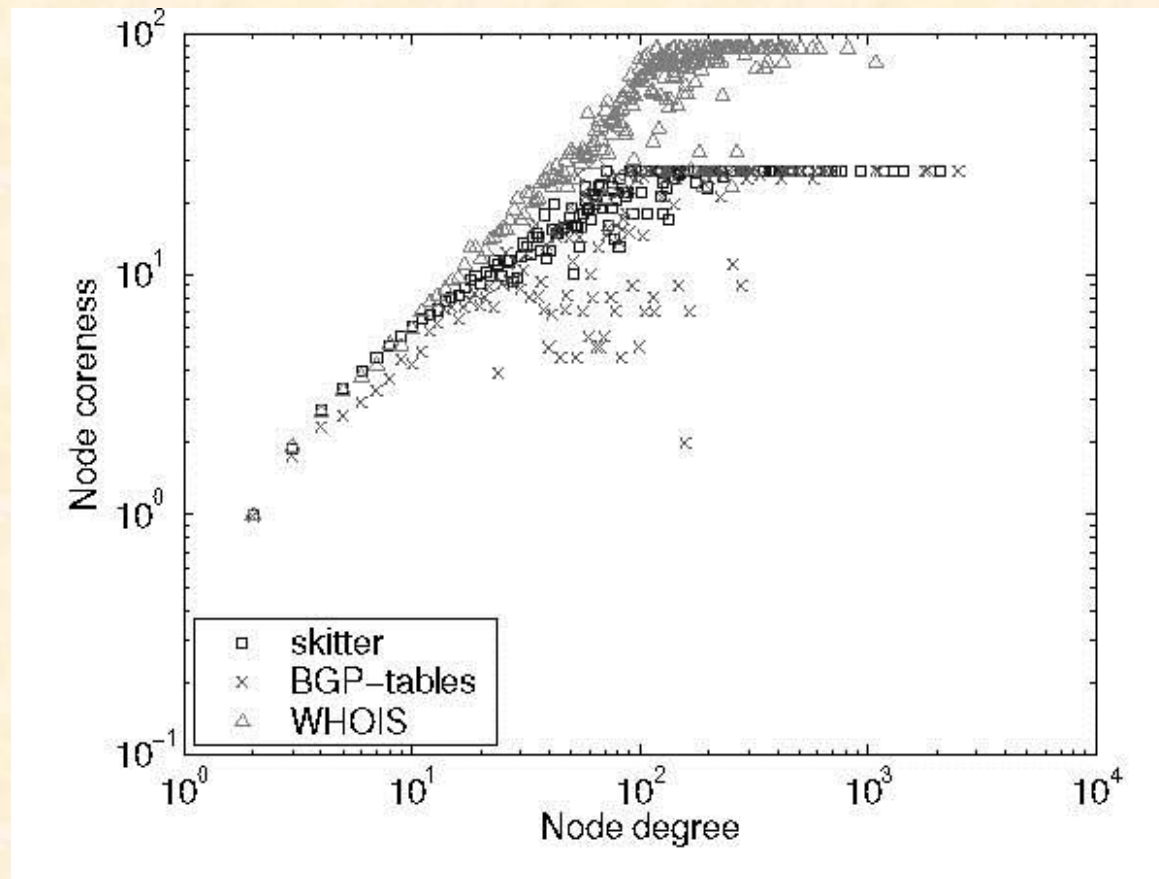
where k_i and j_i are the degrees of the end nodes of edge i , and M is the total number of edges in the network.

If $r > 0$ then the network is *assortative* (big-big nodes);

if $r < 0$ then the network is *disassortative* (big-small nodes).

Analysis on real data using BGP (Border Gateway Protocol, Skitter and Whois about the Internet at some AS levels shows that their assortativity coefficients of the Internet are -0.19 , -0.24 and -0.04 , respectively, implying that Internet is overall *disassortative*.

Coreness

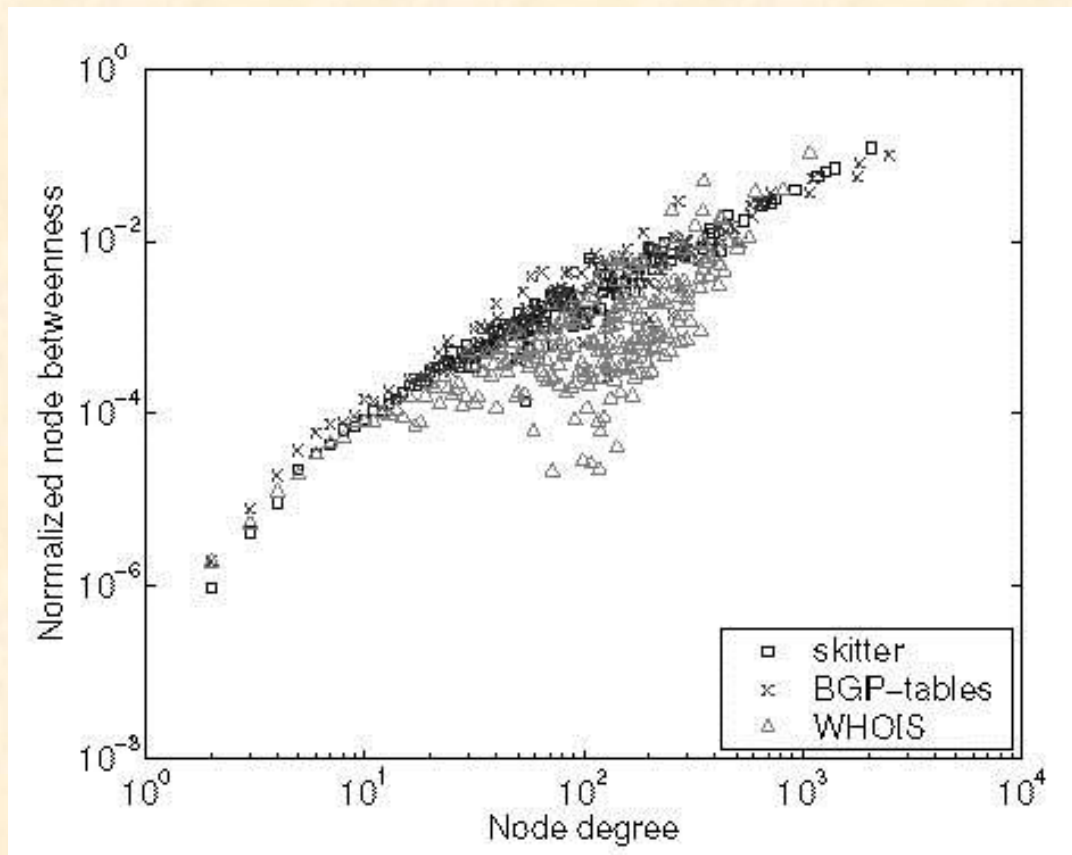


Internet data from Skitter, BGP and Whois:

(Mahadevan et al., 2005)

- when the node degree is relatively small, they have a power-law relation, with exponents 0.58, 0.68 and 1.07, respectively;
- when the node-degree is larger than 100, their coreness values become saturated.

Betweenness



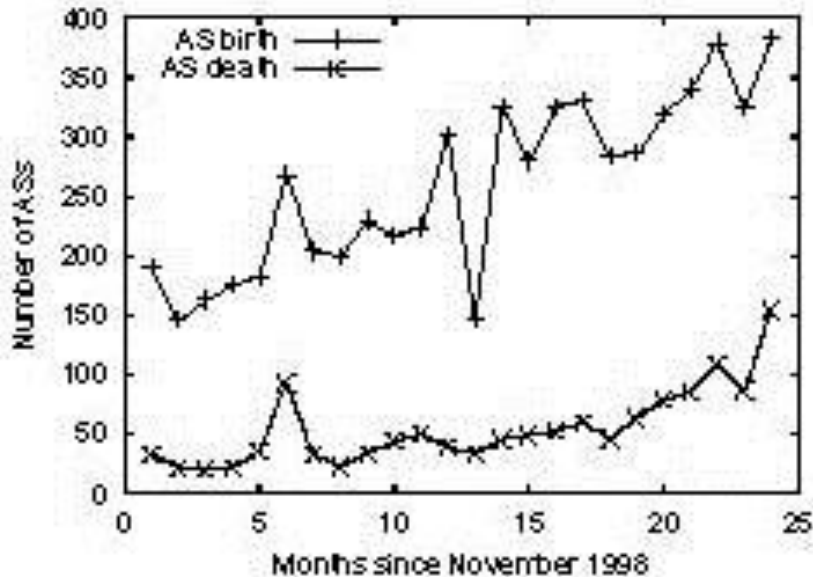
Internet data from Skitter, BGP and Whois: (Mahadevan et al., 2005)

For those from Skitter and BGP, their relations follow prominent power-law distributions with components 1.35 and 1.17, respectively

Growth of the Internet

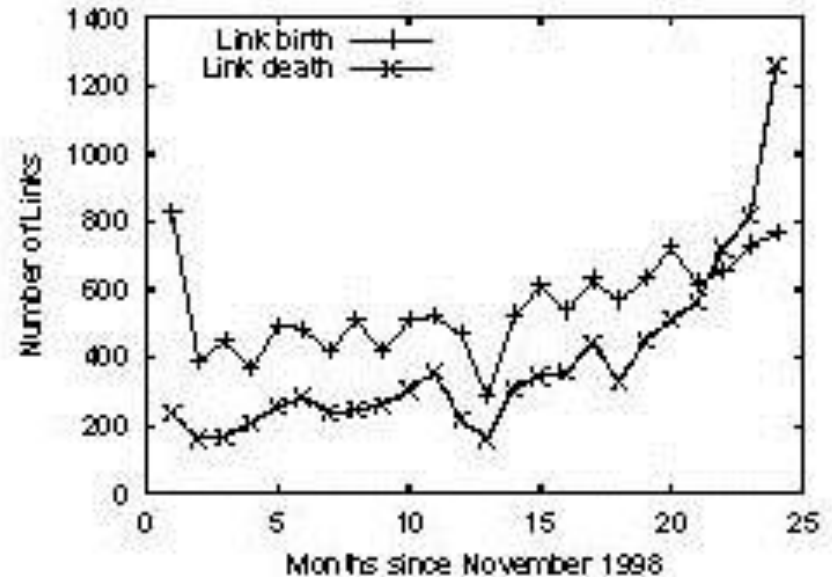
- ❖ The Internet is actually a dynamically evolving complex network, which is rapidly and continuously growing and restructuring
- ❖ Some historical data of the Internet will be examined, which were provided by
 - the SCAN Project with the software named Mercator
 - the Oregon router services (at the AS level)
 - the Topology Project of the Computer Science Department of the Michigan University (at the Extended AS level)

Growth of the Internet



Nodes

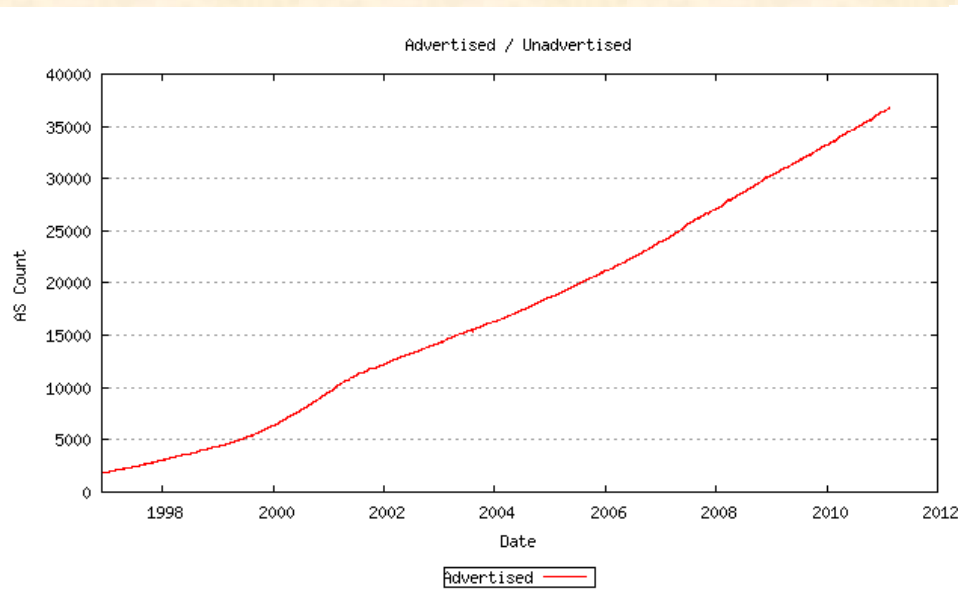
(Qin et al., 2002)



Edges

- ❖ Monthly numbers of birth and death of the Internet at the AS level
- ❖ Real data collected from Oregon servers, the Looking Glass site, and the Internet Routing Registry (IRR)

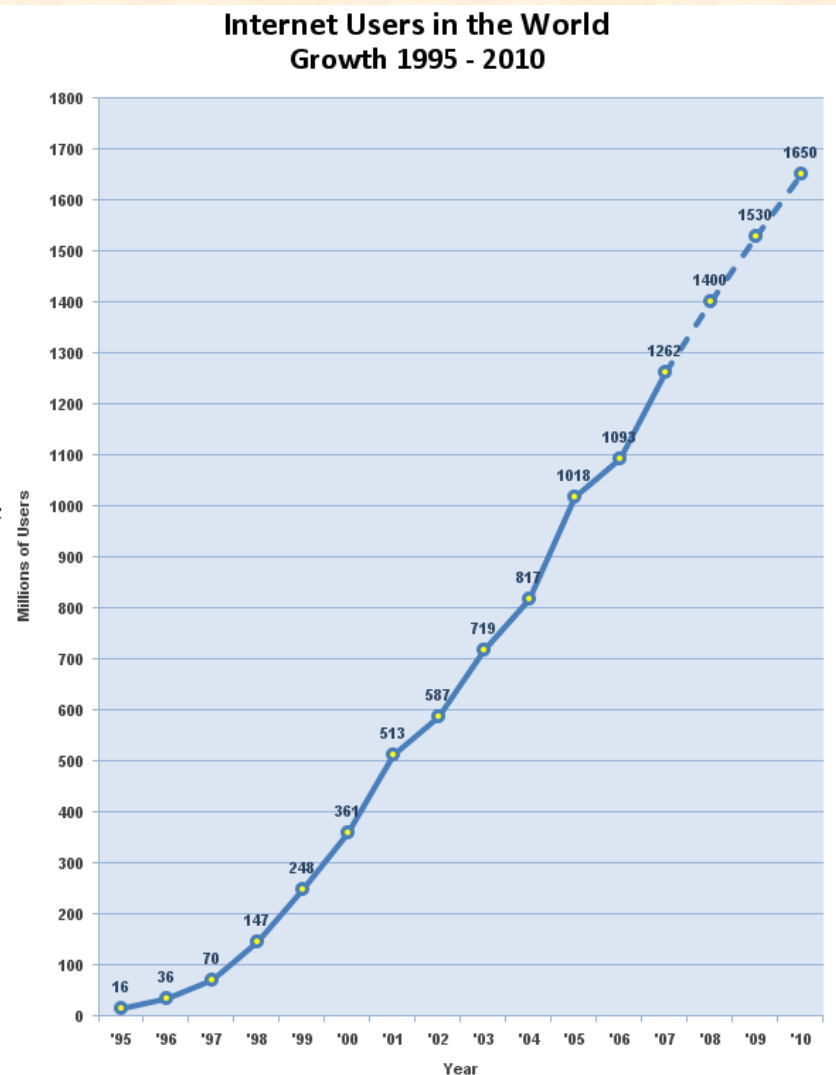
Growth of the Internet



AS Numbers

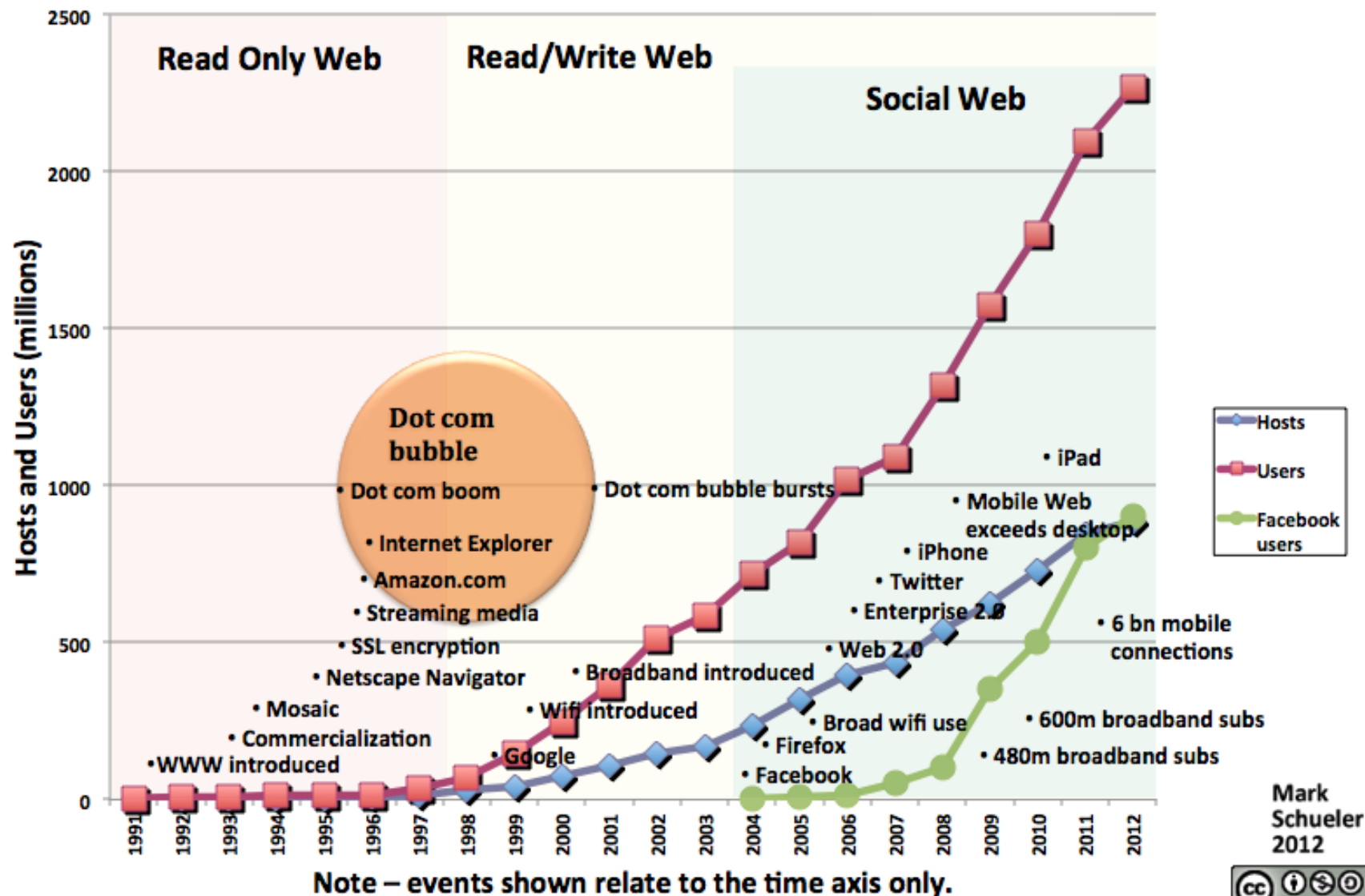
<http://www.potaroo.net/tools/asn32/>

Users Numbers →



Source: www.internetworldstats.com - January, 2008
Copyright © 2008, Miniwatts Marketing Group

Internet Growth - Usage Phases - Tech Events

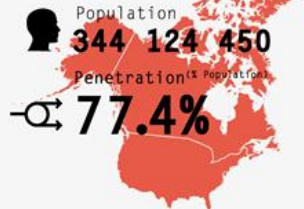


Mark
Schueler
2012



← THE INTERNET GROWTH →

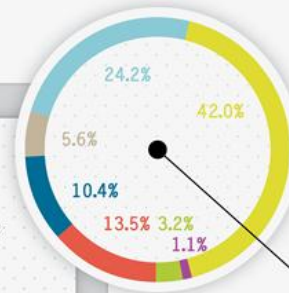
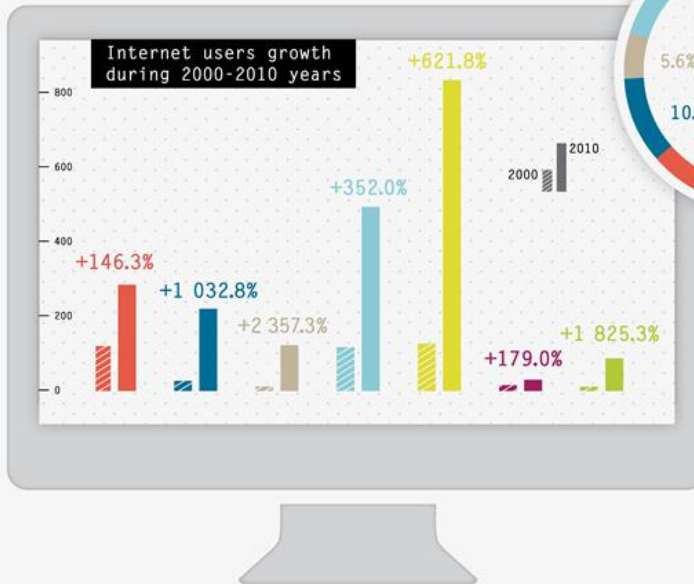
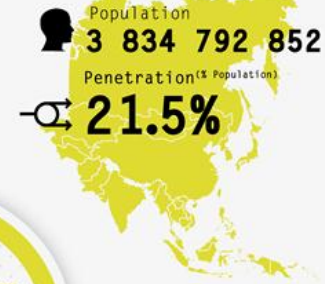
North America



Europe



Asia



Middle East



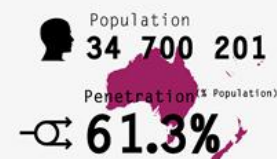
Latin America/Caribbean



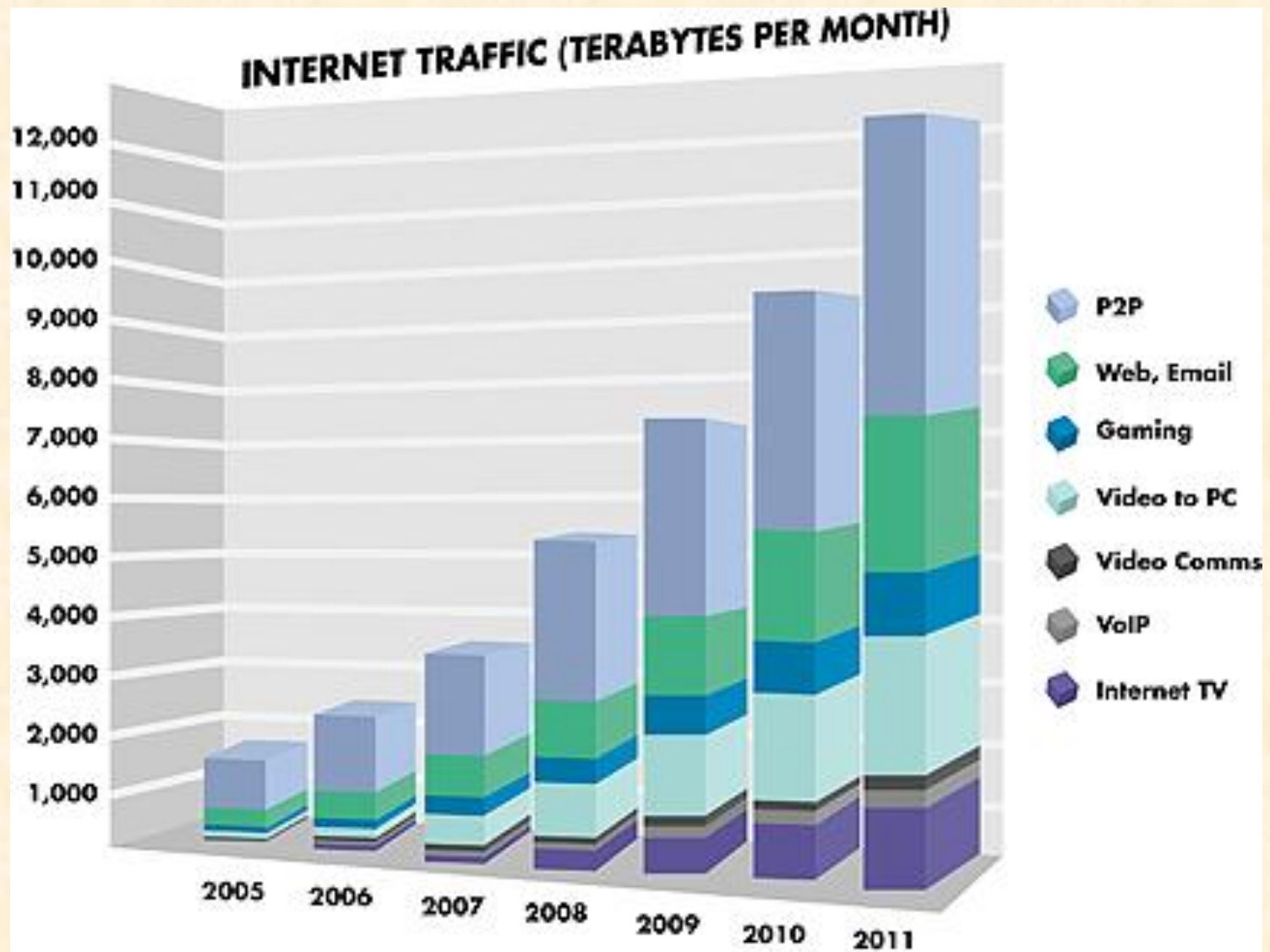
Africa



Oceania/Australia



Internet Traffic Growth



Features of the Internet

Table 4-1 Some average measures of the Internet at the IR, AS, and EAS levels

Level	N	E	$\langle k \rangle$	C	L	B
IR	228,263	320,149	2.8	0.03	9.5	5.3
AS	11,174	23,409	4.2	0.30	3.6	2.3
EAS	11,461	32,730	5.7	0.35	3.6	2.3

IR -- Internet Router

AS -- Autonomous System

EAS -- Extended Autonomous System

N -- number of nodes

$\langle k \rangle$ -- average node degree

L -- average shortest path lengths

E -- number of edges

C -- average clustering coefficient

B -- average node-betweenness

(Pastor-Satorras and Vespignani, 2004)

Features of the Internet

Reproducing the IPv6 Internet AS topology

	AS by Dolphin	BA Model	EBA Model
Number of nodes	419	419	419
Number of edges	1812	1661	1694
Avg degree	8.64	7.92	8.09
Exponent of $P(k)$	1.27	2.82	2.13
Max degree	119	55	57
Avg distance	2.78	3.00	3.12
Normalized avg node betweenness	4.2E-3	4.7E-3	5.0E-3
Normalized avg edge betweenness	1.5E-3	1.8E-3	1.5E-3
Avg clustering	0.692	0.041	0.293
Exponent of rich club	-1.26	-1.10	-1.19

B. Xiao, et al. "Modeling the IPv6 Internet AS-level topology", Physica A (2009)

Node Evolution of the Internet

Table 4-2 Total numbers of additions and deletions of the Internet at the AS level

Degree	Number of Additions	Number of Deletions
1	5,591	1,184
2	816	204
3	23	22
4	4	6
5	1	4
6	1	1
7	1	1
9	0	1
10	1	0
11	1	0
12	0	1
14	1	0
48	0	1

(Qin et al., 2002)

Edge Evolution of the Internet

Table 4-3 Monthly rate of new edges connecting old nodes to new and old ones

Year	1998	1999
$E_{n,o}$	170	231
$E_{o,o}$	350	450
$E_{n,o} / E_{o,o}$	0.48	0.53

$E_{n,o}$ -- numbers of new edge additions between incoming nodes and existing nodes

$E_{o,o}$ -- numbers of new edge additions between two existing nodes

(Vazquez et al., 2002)

End

