

Internet Traffic and Content Consolidation

Craig Labovitz
Chief Scientist, Arbor Networks

S. Iekel-Johnson, D. McPherson
Arbor Networks, Inc.

J. Oberheide, F. Jahanian
University of Michigan



Talk Outline

- **Describe two-year traffic measurement study**
- **The “original” Internet topology**
- **The emerging new Internet**
- **Application transport and the end of end-to-end**
- **A few words on IETF implications**

Two Year Study of Inter-domain Traffic

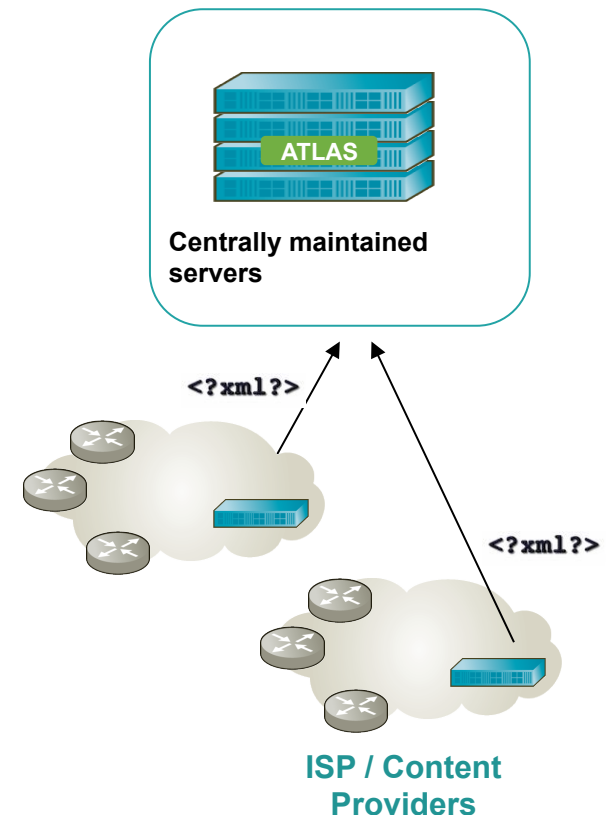


Graphic not an accurate representation of current ATLAS deployments

- **Leverage large, widely deployed commercial Internet monitoring infrastructure**
- **Global deployment across 110+ ISPs / Content Providers**
 - Near real-time traffic and routing statistics (14 Tbps)
 - Participation voluntary and all data sources are anonymous
 - Largest study of its kind

Study Details

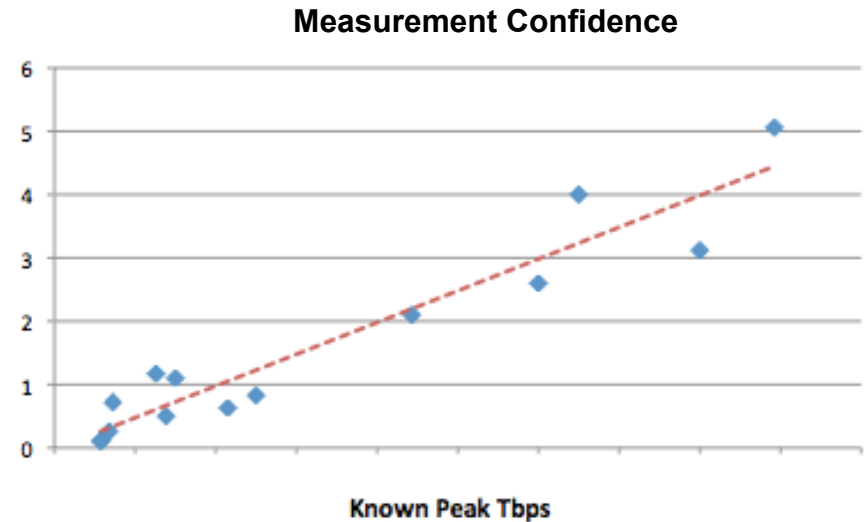
- **Within a given ISP, commercial probe infrastructure**
 - Monitors NetFlow / Jflow / etc and routing across possible hundreds of routers
 - Probes topology aware of ISP, backbone and customer boundaries
 - Routers typically include most of peering / transit edge
 - Some deployments include portspan / inline appliances
- **Deployments send anonymous XML file to central servers**
 - Includes self-categorization of primary geographic region and type
 - Data includes coarse grain anonymized traffic engineering statistics
- **Introduced at NANOG 47 academic paper under review, Arbor blog provides ongoing related bits**



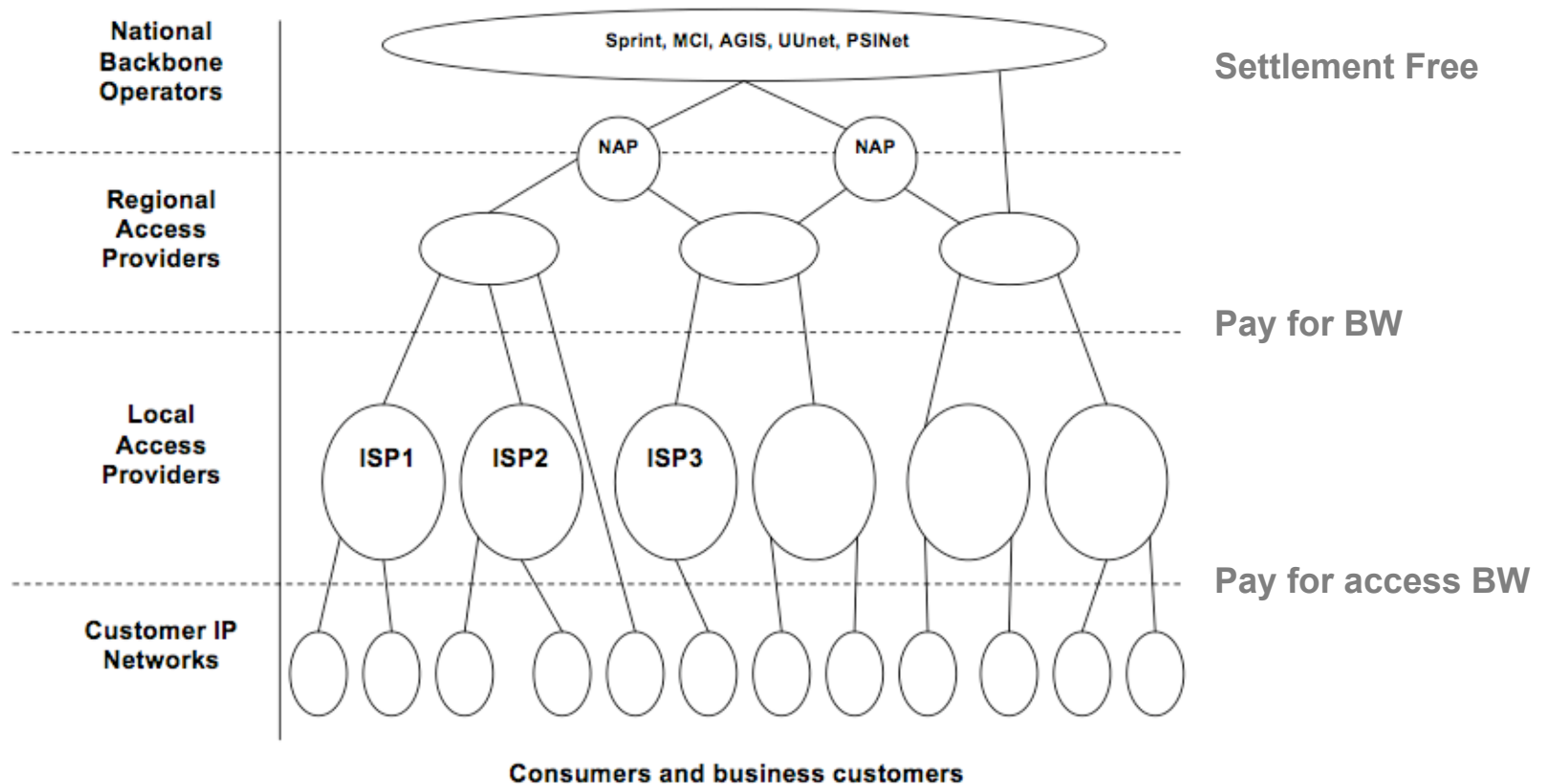
Traffic Measurements

- **Inter-domain traffic volumes**
 - Estimate directly monitoring 25% all inter-domain traffic
 - Believe data representative of global inter-domain traffic
 - Validate predictions based on data (using 12 known ISP traffic demands)

- **Does NOT measure**
 - Number of web hits, tweets, transactions, customers, etc.
 - Internal / private customer traffic (e.g. VPNs, IPTV)
 - ISP success nor profitability

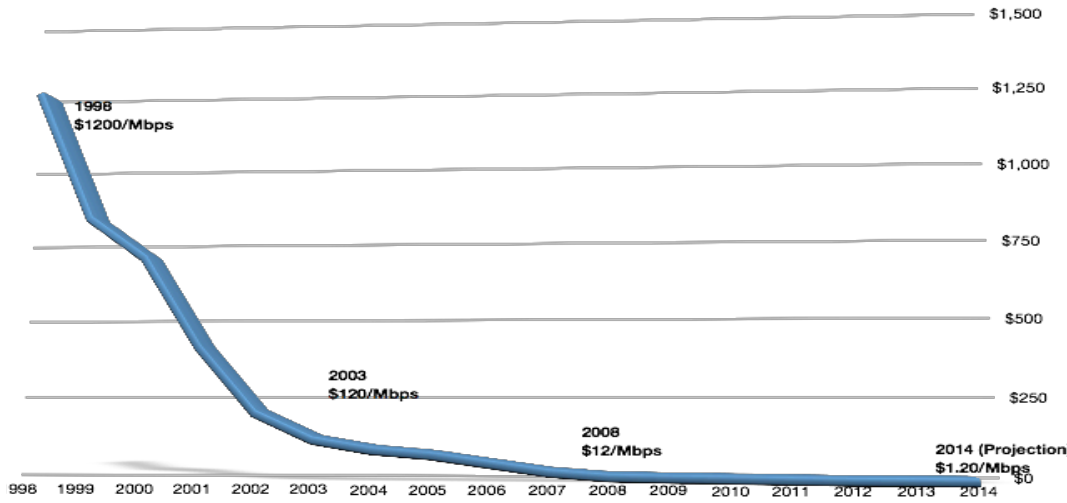


Original Internet (1995 – 2007)



- Textbook diagram (still taught today)
- Hierarchical, relatively sparsely inter-connected Internet
- Mostly accurate until recently (modulo a few name changes over the years)

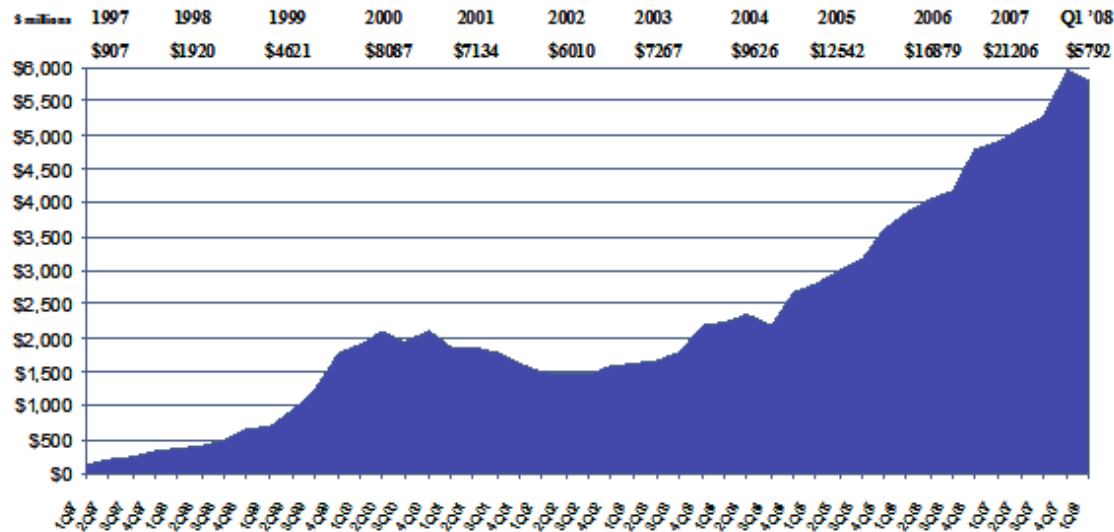
Market Forces Reshape Traffic and Connectivity



Revenue from Internet Transit

Source: Dr. Peering, Bill Norton

Quarterly Internet Ad Revenues



Revenue from Internet Advertisement

Source: Interactive Advertising Bureau

Largest Carriers: Then and Now

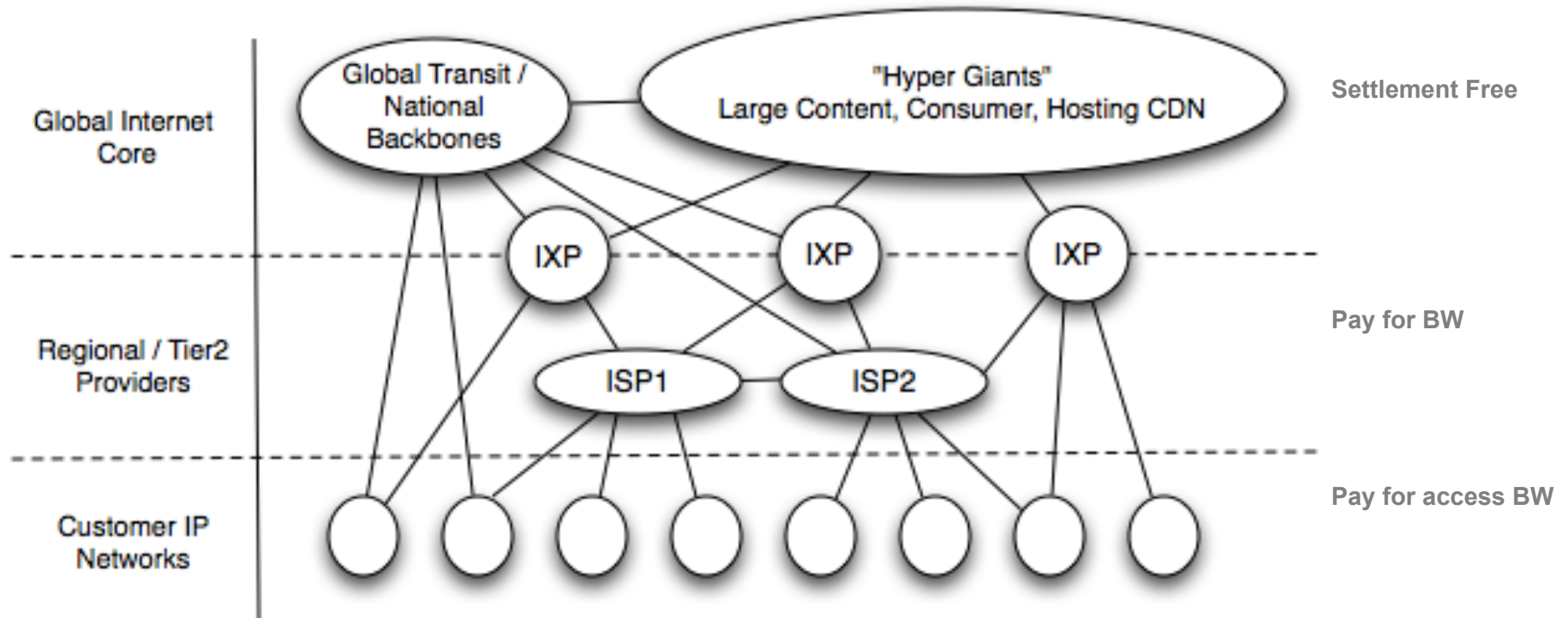
Rank	2007 Top Ten	%
1	ISP A	5.77
2	ISP B	4.55
3	ISP C	3.35
4	ISP D	3.2
5	ISP E	2.77
6	ISP F	2.6
7	ISP G	2.24
8	ISP H	1.82
9	ISP I	1.35
10	ISP J	1.23

Rank	2009 Top Ten	%
1	ISP A	9.41
2	ISP B	5.7
3	Google	5.2
4	-	
5	-	
6	Comcast	3.12
7	-	
8	-	
9	-	
10	-	

Based on analysis of anonymous ASN (origin/transit) data (as a weighted average % of all Internet Traffic). Top ten has NO direct relationship to study participation.

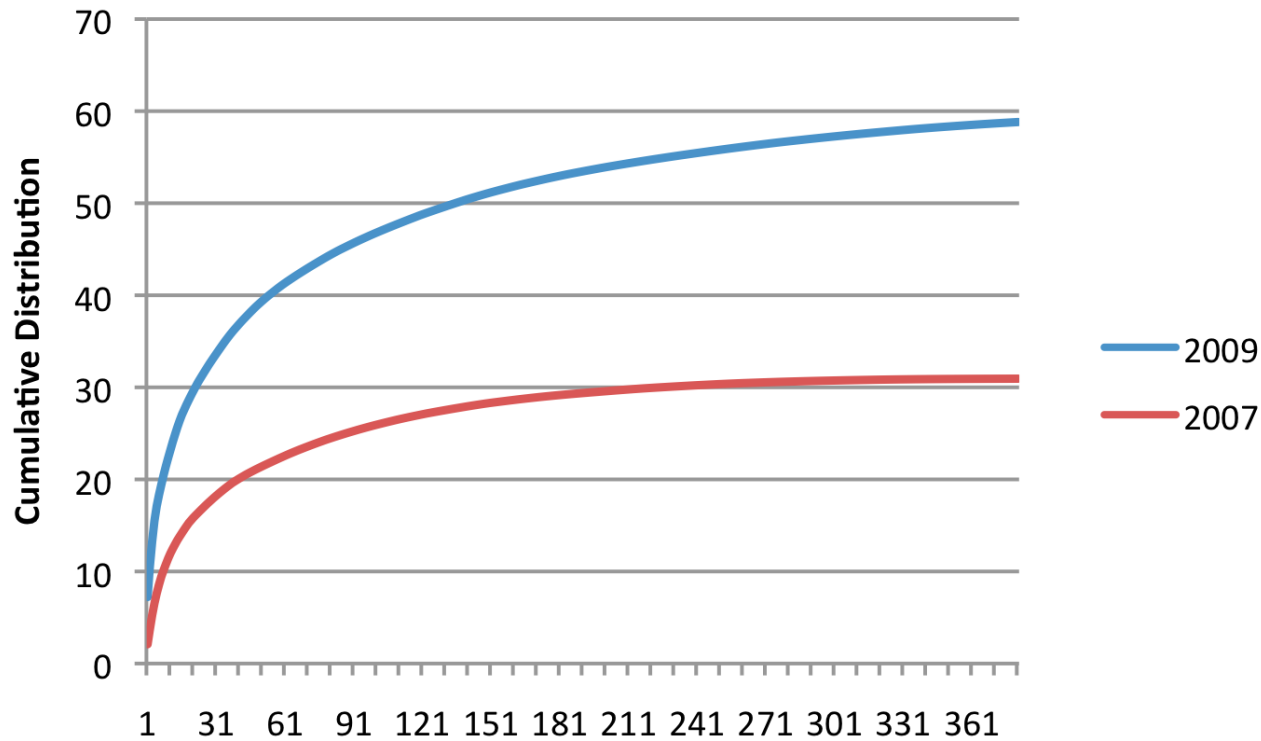
- In 2007, top ten match “tier-1” ISPs (e.g., Wikipedia)
- In 2009, global transit carry significant traffic volumes
 - But Google and Comcast join the list
 - And a significant percentage of ISP A traffic is Google transit

The New Internet



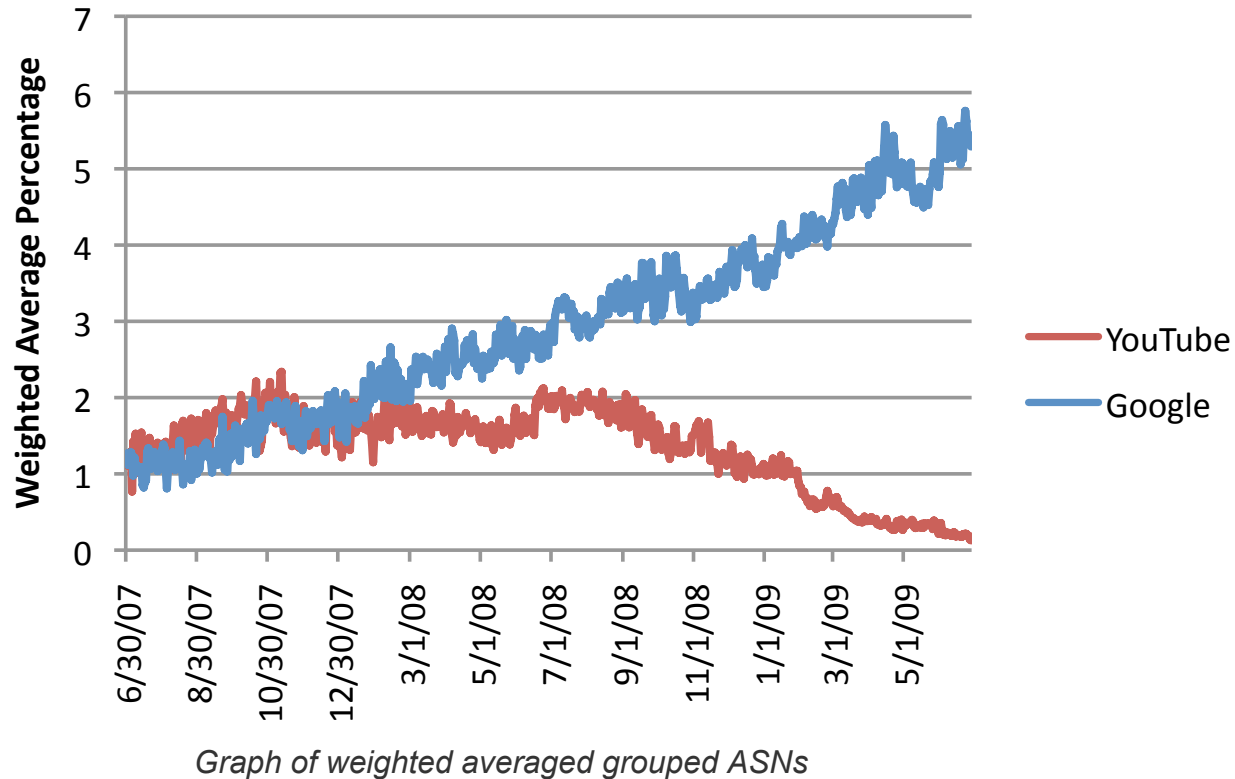
- Flatter and much more densely interconnected Internet
 - Significant routing, traffic, security, economic, implications
- Disintermediation between content and eyeball networks
- New commercial models between content, consumer and transit

Consolidation of Content *(Grouped Origin ASN)*



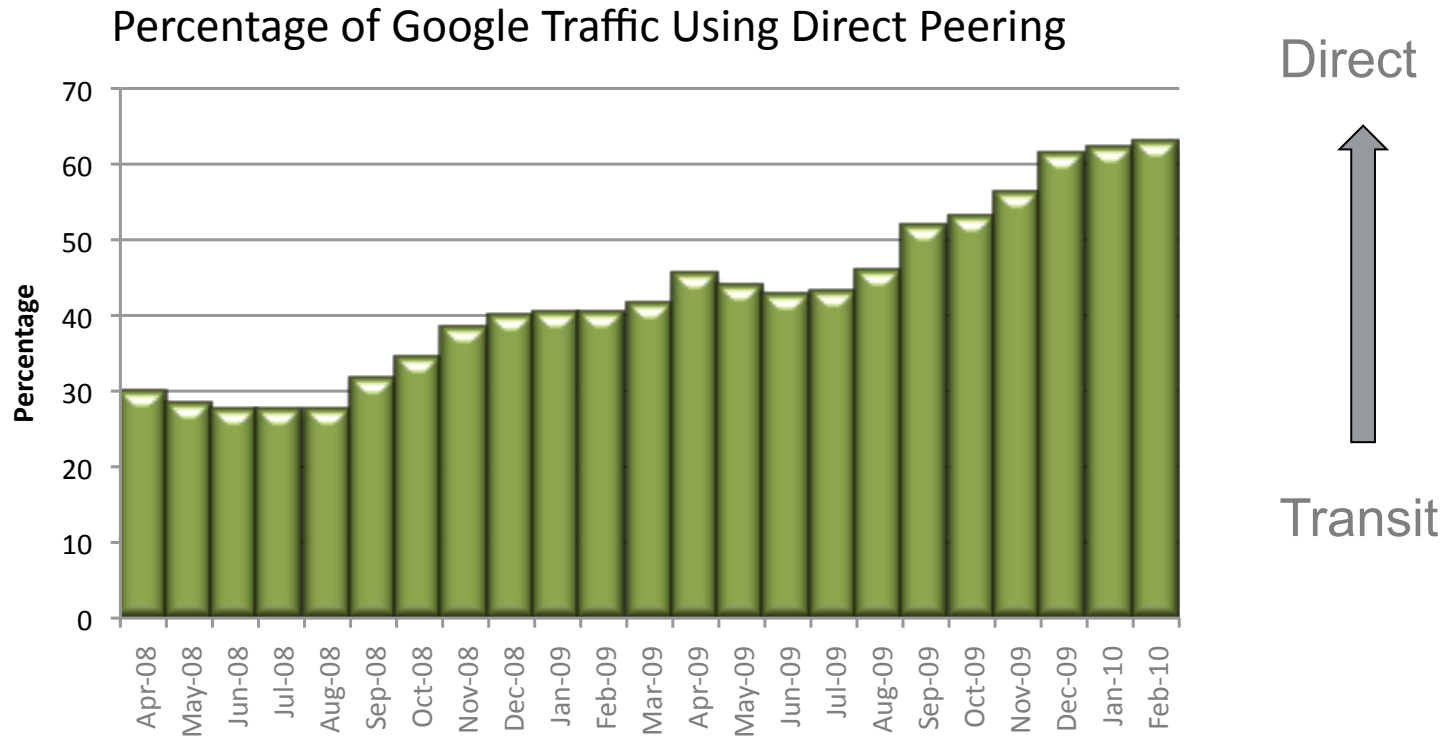
- In 2007, thousands of ASNs contributed 50% of content
- In 2009, 150 ASNs contribute 50% of all Internet traffic
- Approximates a power law distribution

Case Study: Google



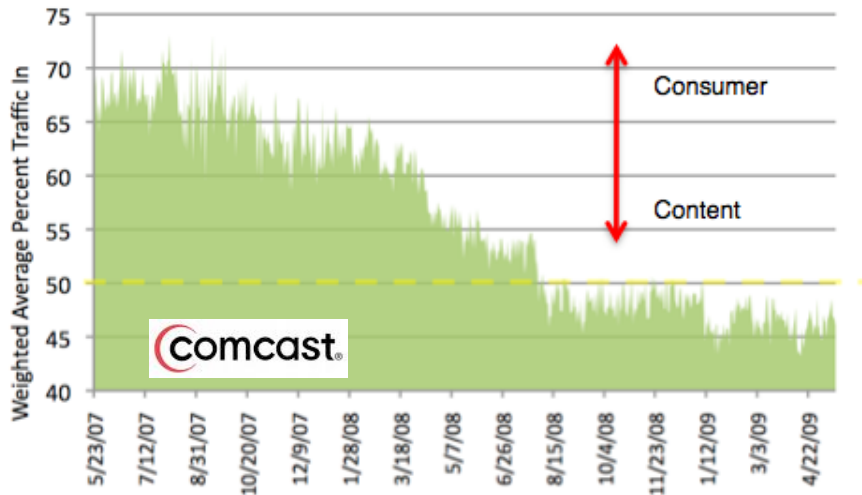
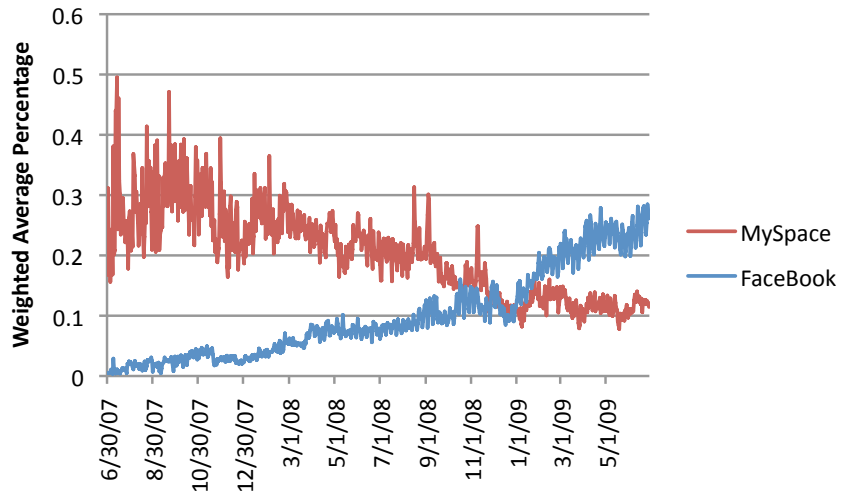
- Over time Google absorbs YouTube traffic
- As of July 2009, Google accounts for 6% of all Internet inter-domain traffic
- Google the fastest growing ASN group

Google Dense Interconnection



- Over time, Google increasingly using direct peering with tier2/3 and eyeball networks
- As of February 2010, more than 60% of Google traffic does not use transit
 - Remainder largely global transit carriers
- These numbers do not include GGC

Other Case Studies



- **Rapid rise of new content players, e.g.**
 - CDNs
 - Facebook
 - Baidu
 - Apple / MSFT
- **Change in traffic patterns and business strategies of consumer networks**

What's Happening?

- **Commoditization of IP and Hosting / CDN**
 - Drop price of wholesale transit
 - Drop price of video / CDN
 - Economics and scale drive enterprise to “cloud”
- **Consolidation**
 - Bigger get bigger (economies of scale)
 - e.g., Google, Yahoo, MSFT acquisitions
- **Success of bundling / Higher Value Services**
 - Triple and quad play, etc.
- **New economic models**
 - Paid content (ESPN 360), paid peering, etc.
 - Difficult to quantify due to NDA / commercial privacy
- **Disintermediation**
 - Direct interconnection of content and consumer
 - Driven by both cost and increasingly performance

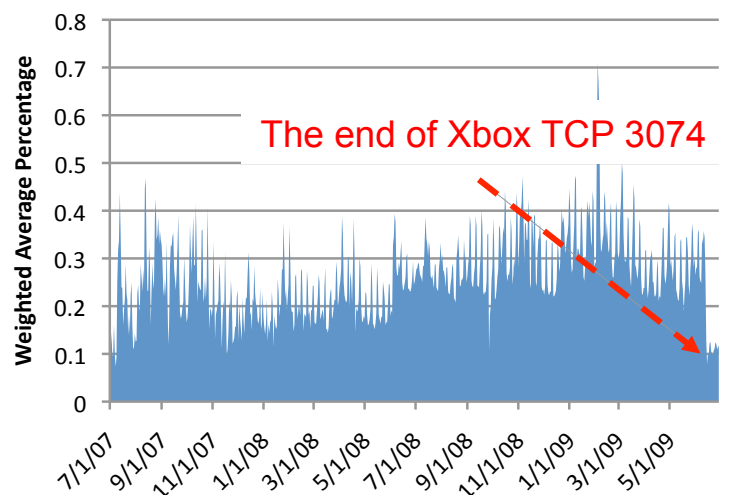
Applications

Rank	Application	2007	2009	Change
1	Web	41.68%	52.00%	24.76%
2	Video	1.58%	2.64%	67.09%
3	VPN	1.04%	1.41%	35.58%
4	Email	1.41%	1.38%	-2.13%
5	News	1.75%	0.97%	-44.57%
6	P2P (*)	2.96%	0.85%	-71.28%
7	Games	0.38%	0.49%	28.95%
8	SSH	0.19%	0.28%	47.37%
9	DNS	0.20%	0.17%	-15.00%
10	FTP	0.21%	0.14%	-33.33%
	Other	2.56%	2.67%	4.30%
	Unclassified	46.03%	37.00%	-19.62%

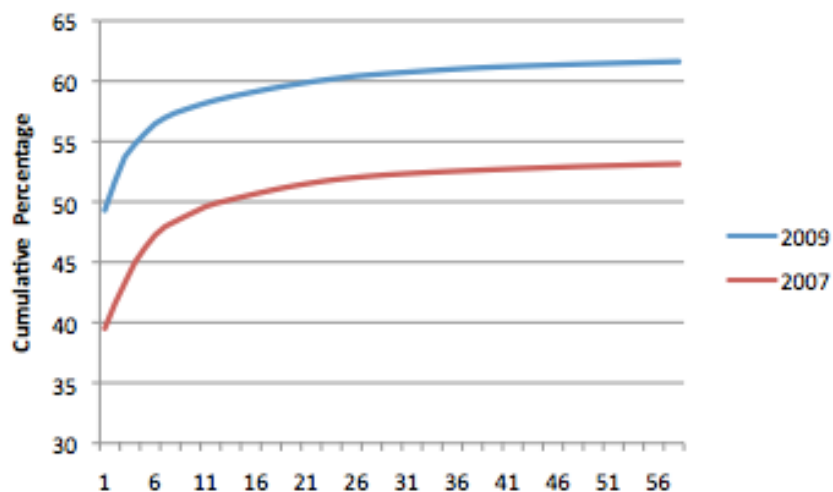
(*) 2009 P2P Value based on 18% Payload Inspection
Weighted average percentage of all Internet traffic using well-known ports

- **Growing volume of Internet traffic uses port 80 / 443**
 - Includes significant video component and source of most growth
- **Unclassified includes P2P and video**
 - Payload matching suggests P2P at 18%
 - P2P is fastest declining

The End of End-to-End?



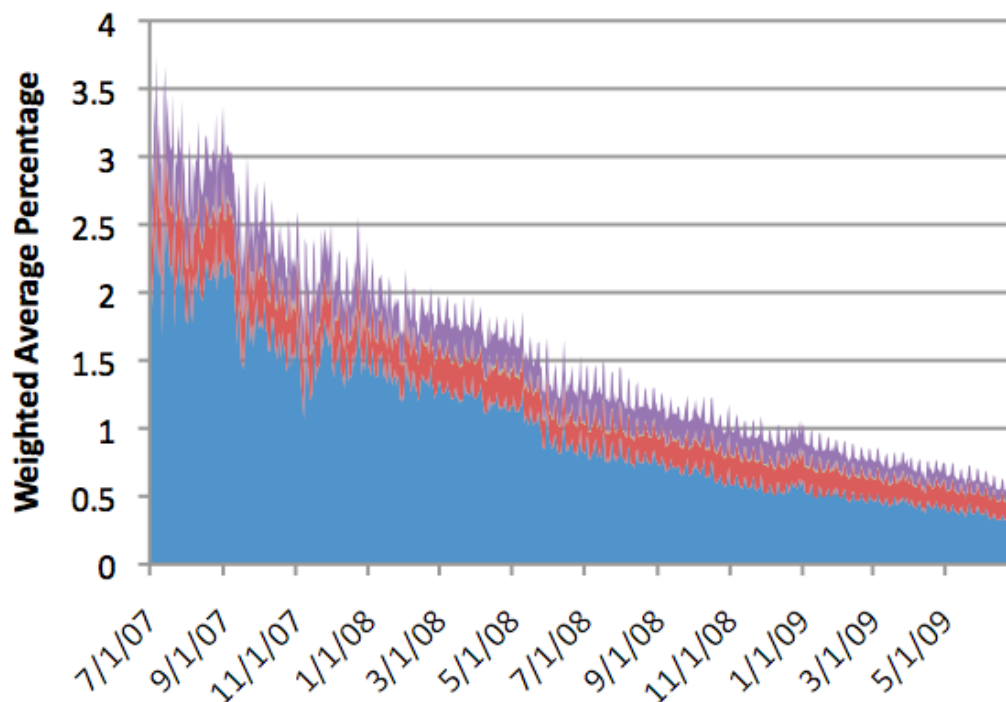
Weighted average percentage of Xbox Internet traffic



Cumulative Distribution of Traffic to TCP / UDP Ports

- Growing dominance of web as application front-end
- Plus burden of ubiquitous network layer security policies
- Results in growing concentration of application traffic over a decreasing number of TCP / UDP ports
 - Especially port 80
 - Especially video

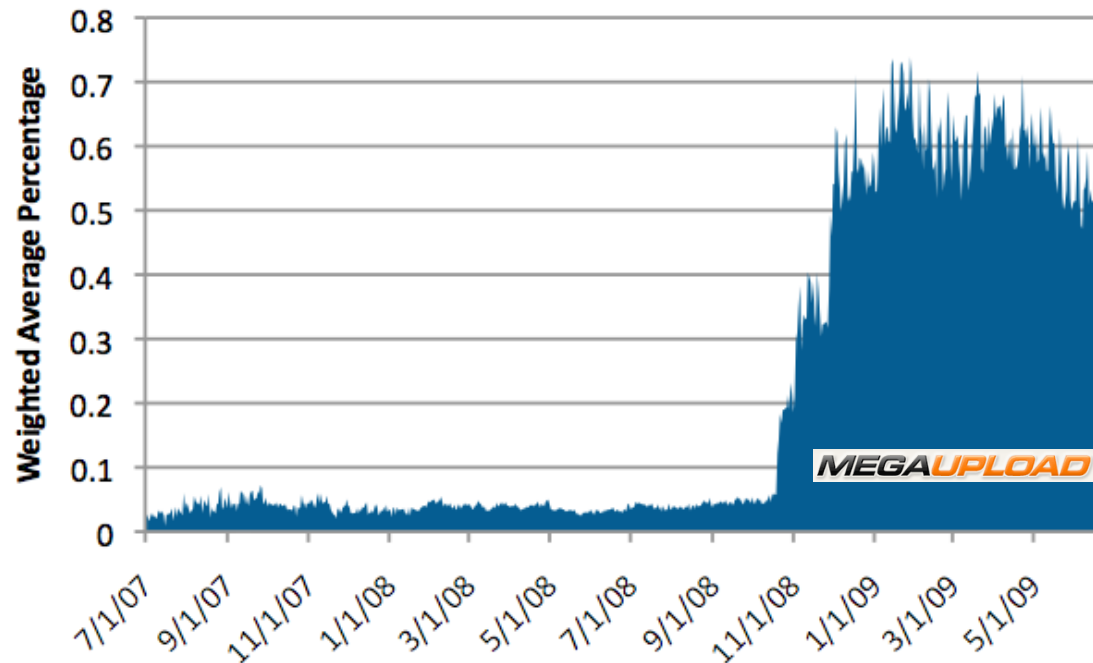
P2P



Graph of weighted average traffic using well-known P2P ports

- **In 2006, P2P one of largest threats facing carriers**
 - Significant protocol, engineering and regulatory effort / debate
- **In 2010, P2P fastest declining application group**
 - Trend in both well-known ports and payload based analysis
 - Still significant volumes
 - Slight differences in rate of decline by region (i.e. Asia is slower)

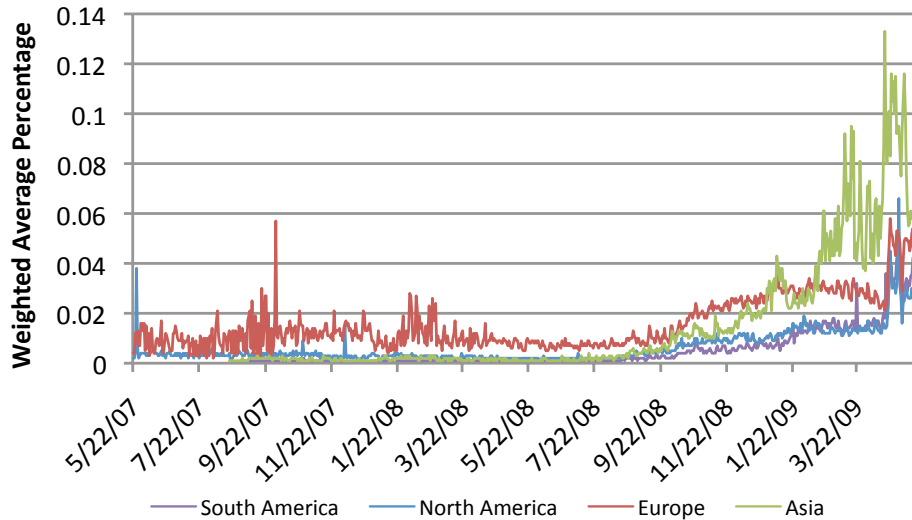
P2P Surpassed by Direct Download



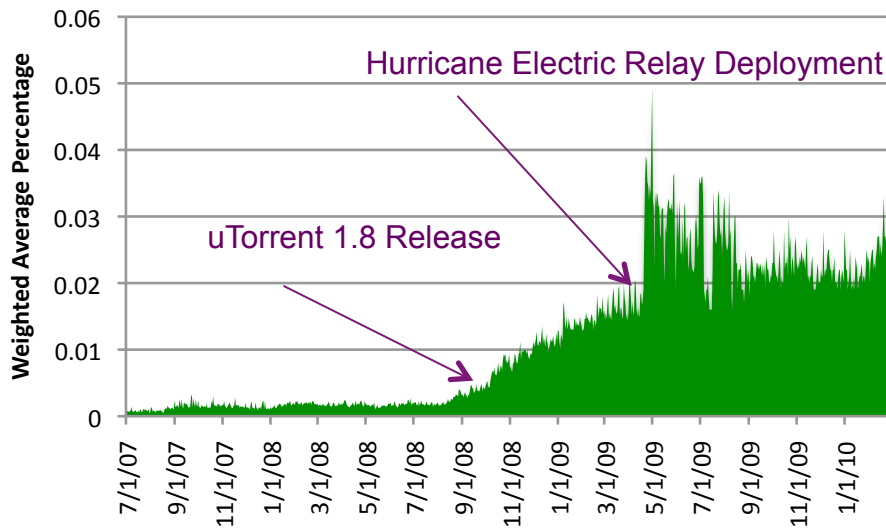
Weighted average percentage of Internet traffic contributed by Carpathia ASNs

- **Normally study lacks visibility into hosting customers**
- **Mega [Upload|Video|Erotic] is an exception**
 - Carpathia small hosting company by traffic volume in Fall 2008
 - Mega becomes Carpathia customer in November 2008
 - Carpathia Hosting grows overnight to more than 0.5% of all traffic

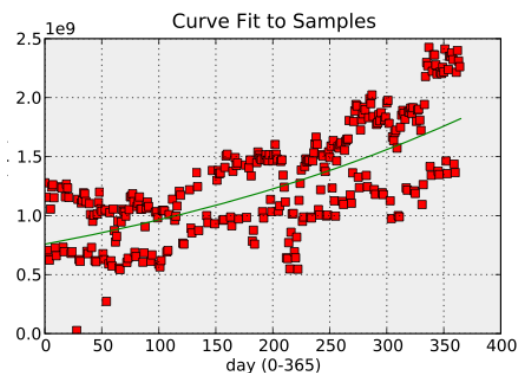
IPv6



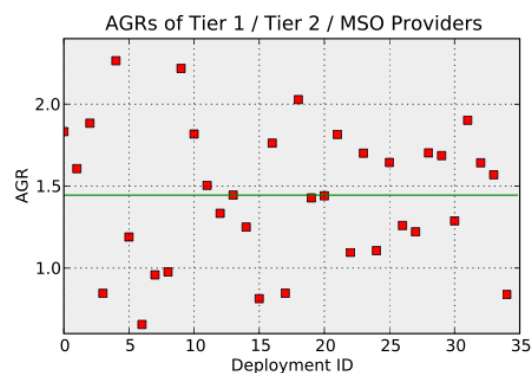
- IPv6 miniscule percentage of Internet traffic (.04 %)
- Still relatively little native IPv6 peering between large carriers
- Few carriers with v6 traffic visibility (i.e. flow)
- Tunneled IPv6 shows growth since IPv6
 - Due to uTorrent
 - And Hurricane Electric global Teredo deployment (see blog)



Internet Size / Growth



(a) Example AGR Calculation



(b) Per-deployment AGRs

Estimate	Observatory	ISP Survey	Cisco	MINTS
Traffic Volume Per Month	9 exabytes	N/A	9 exabytes	5-8 exabytes
Traffic Annual Growth Rate	44.5%	35-45%	50%	50-60%

- **In 2009, Internet (inter-domain) roughly ~45 Tbs**
 - And growing at 45% per year
- **Significant, but no “Exaflood”**
 - Followed MINTS methodology for AGR
 - Used 10 known ISP totals (MRTG / Flow based) to extrapolate Internet total

IETF Implications

- **Increasingly dense Internet and impact on routing scalability and convergence**
- **Slow IPv6 deployment highlights need for alternative transition mechanisms**
- **The “end” of end-to-end**
 - Increasing impact of firewall, NAT
 - Silo’ed ecosystems

Conclusion

- **Internet is at an inflection point**
- **Focus shifting from transmission to content**
 - Battle for access to eyeballs (and control of content)
 - Transit is commoditized and devalued
 - New focus on datacenters and co-location (caches)
- **New technologies reshaping definition of Internet**
 - “Web” / Desktop Applications, Cloud computing, CDN
- **Changes mean significant new commercial, security and engineering challenges**
- **This is just the beginning...**