McCormick
Northwestern Engineering
Department of Electrical Engineering and Computer Science

Internet Inter-Domain Traffic

Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, Farnam Jahanian

Presented by: Mario Sanchez
Instructor: Fabian Bustamante
Date: 01/10/2011

Goals

- Examine changes in Internet inter-domain traffic demands and interconnection policies
- Longitudinal observations of Internet traffic

Methodology

- Focused on inter-domain traffic, not application layer (web hits/tweets/VPN/etc)
- Exported coarse-grain traffic statistics about ASNs, ASPaths, protocols, ports, etc. via anonymous XML forwarded to central servers
- Leverage commercial probes within given ISPs, with limited visibility into payload-based classification
- Incorporated informal and formal discussions with providers, and information about known traffic volumes
- Validated predictions based on a ground-truth based on 12 known ISP traffic demands (Known peak Tbps)

Methodology

- Covered 110 ISPs/content providers
- 3k edge routers
- 100k interfaces
- ~25% of all inter-domain traffic
- Waited for 2 years

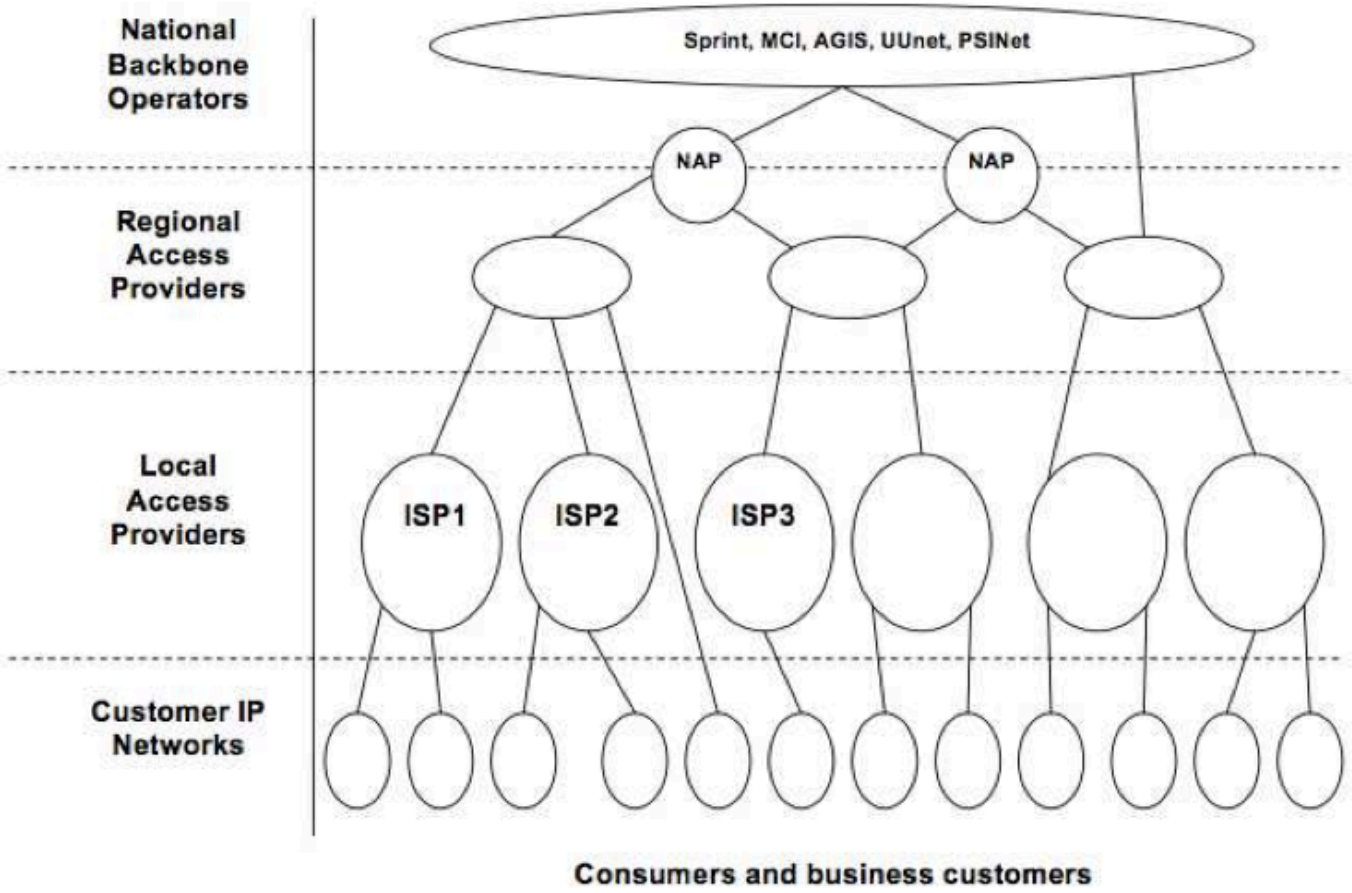
- Calculated percentages per category then weighted averages using number of routers in each deployment

$$W_{d,i} = \frac{R_{d,i}}{\sum_{x=1}^{N} R_{d,x}} \qquad P_d(A) = \sum_{x=1}^{N} W_{d,x} * \frac{M_{d,x}(A)}{T_{d,x}} * 100$$

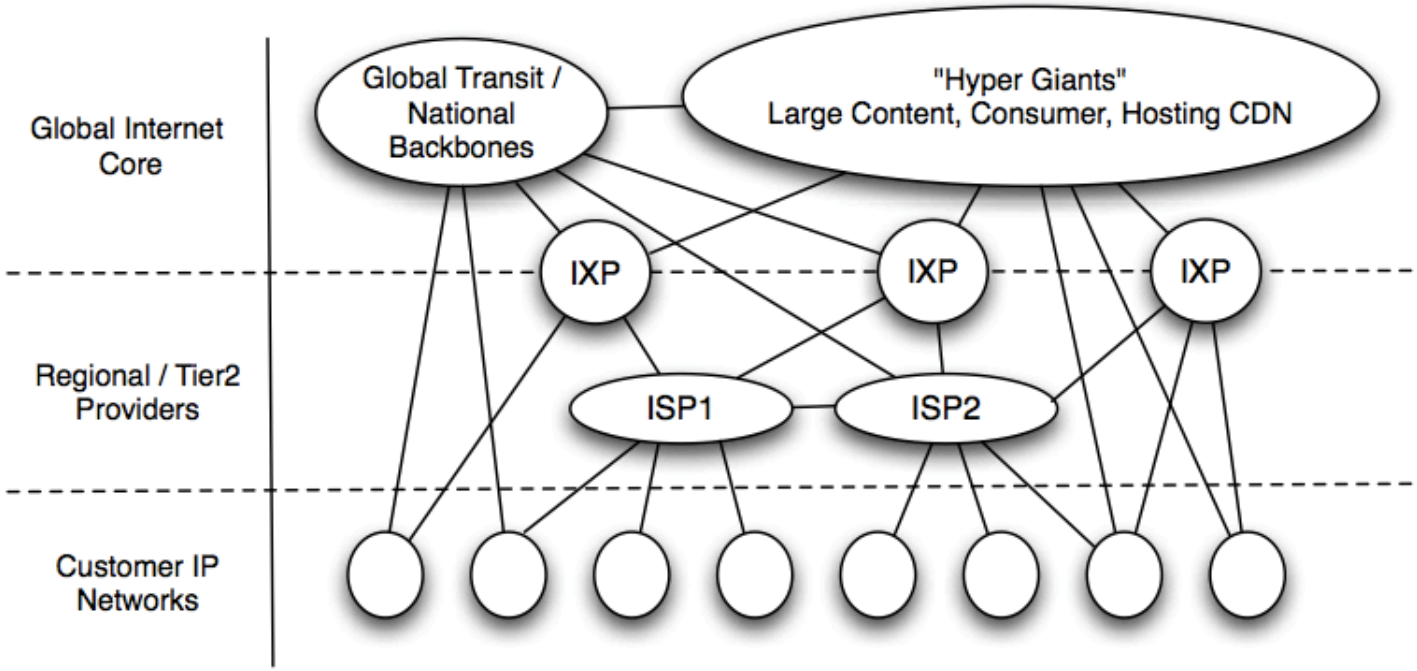
Internet Evolution

- Most of the past 15 years of commercial Internet:
  - 10 to 12 large transit providers, interconnecting:
    - Tier-2,
    - Regional providers,
    - Consumer networks
    - Content/hosting companies
- Las five years saw a shift in Internet inter-domain traffic demands and peering polices
  - Content providers build their own global backbones
  - Cable Internet service providers offer wholesale national transit
  - Transit ISPs offer CDN and cloud/content hosting services

Traditional Internet logical topology



Emerging new Internet logical topology



2009: 65% of study participants use direct links with Google, 52% with Microsoft, 49% with Limelight, 49% with Yahoo

ASN Traffic Analysis

- Calculate 10 larges contributors of inter-domain traffic using weighted average percentage (either originating or transiting each ASN)
- Aggregate all ASNs which are managed by the same Internet commercial entity
- Exclude stub ASNs from the aggregation step which only observed downstream from other corporate ASN

Impact of commercial policy and traffic engineering changes

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

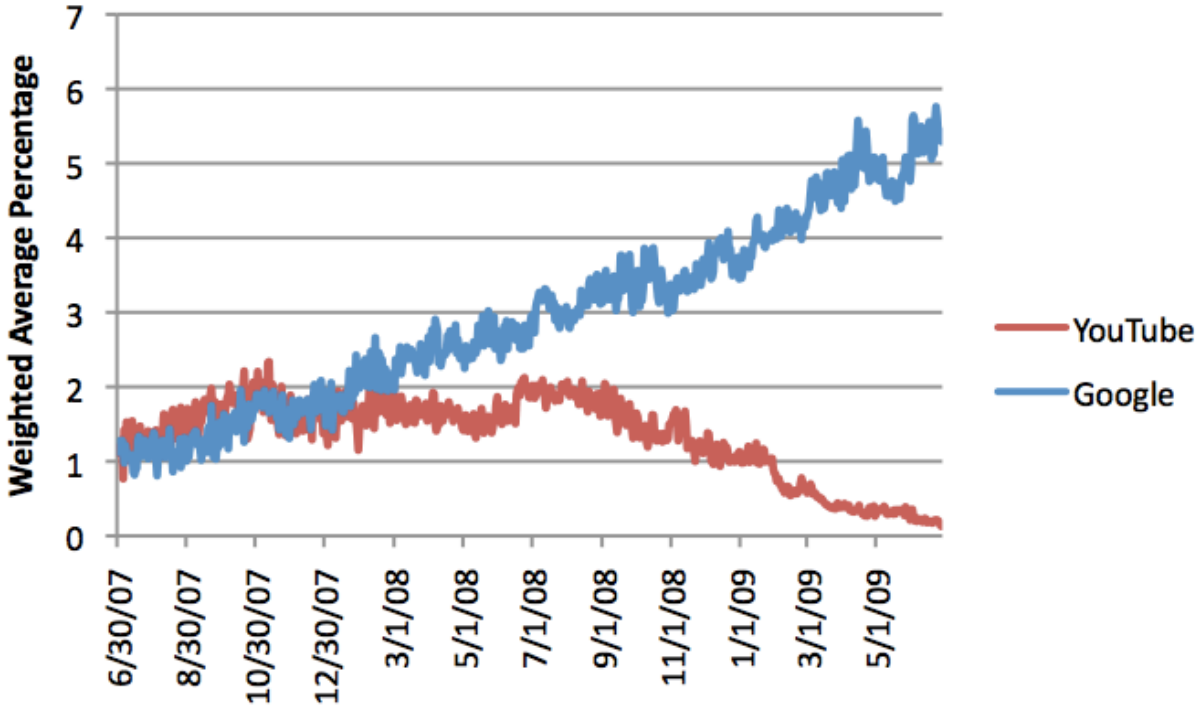
(a) Top Ten 2007

Rank	Provider	Percentage
1	ISP A	9.41
2	ISP B	5.7
3	Google	5.2
4	ISP F	5.0
5	ISP H	3.22
6	Comcast	3.12
7	ISP D	3.08
8	ISP E	2.32
9	ISP C	2.05
10	ISP G	1.89

(b) Top Ten 2009

- 2007: Largest Internet providers correlate with 12 largest transit networks (Tier 1)
- 2009: Includes the addition of non-transit companies to the list: Google and Comcast

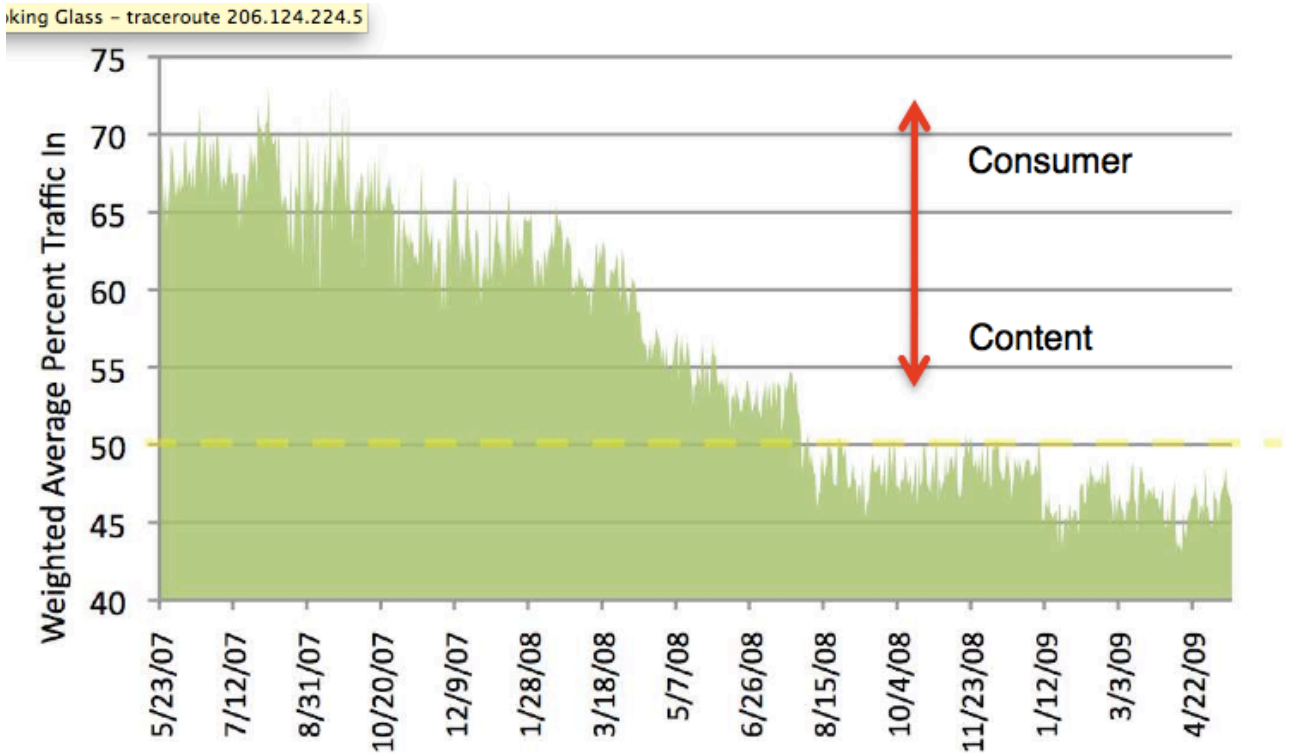
The case of Google



Graph of weighted averaged grouped ASNs

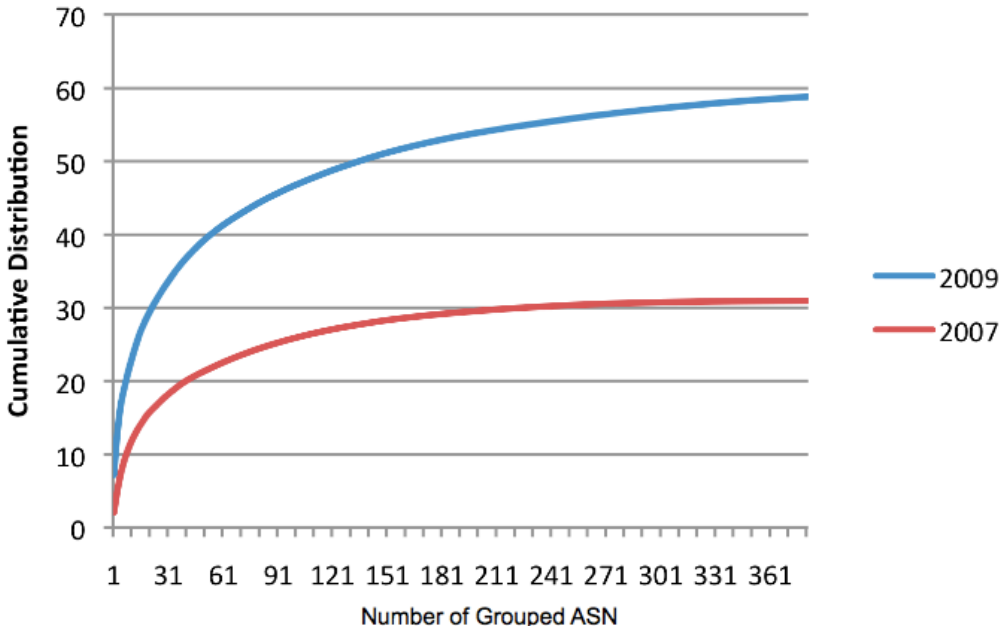
- Google inter-domain traffic enjoyed the largest growth (gaining 4% of all-inter domain traffic)
- Google's traffic share increase came through the post-acquisition migration of YouTube inter-domain traffic to Google's ASNs
- Google the fastest growing ASN group

The case of Comcast – In/Out peering ratio



- Weighted average percentage of inter-domain traffic into all Comcast ASNs vs outbound
- Comcast began offering wholesale transit (GigE and 10GigE IP), cellular backhaul and IP video distribution

Inter-domain traffic consolidation



- 2007: 150 ASNs contribute 30% of all inter-domain traffic
- 2009: 150 ASNs originate more than 50% of all inter-domain traffic
- Majority of traffic by volume flows directly between large content providers, datacenter / CDNs and consumer networks

Application Traffic Analysis

- Methodology
  - Applications are classified by protocol and TCP/UDP port
  - The appliances follow heuristics to select a single probable application (each flow record may contain multiple ports)
- Limited
  - If application uses non-standard ports or ephemeral port numbers
  - Does not identify tunneled applications (video over HTTP) or encrypted traffic (P2P)
  - Port-based heuristics could not identify a probable application in more than 25% of all observed inter-domain traffic
- Augment
  - Using small set of DPI appliances (payload classification)

The winners: HTTP & video

Rank	Application	2007	2009	Change
1	Web	41.68	52.00	+10.31
2	Video	1.58	2.64	+1.05
3	VPN	1.04	1.41	+0.38
4	Email	1.41	1.38	-0.03
5	News	1.75	0.97	-0.78
6	P2P	2.96	0.85	-2.11
7	Games	0.38	0.49	+0.12
8	SSH	0.19	0.28	-0.08
9	DNS	0.20	0.17	-0.04
10	FTP	0.21	0.14	-0.07
	Other	2.56	2.67	+0.11
	Unclassified	46.03	37.00	-9.03

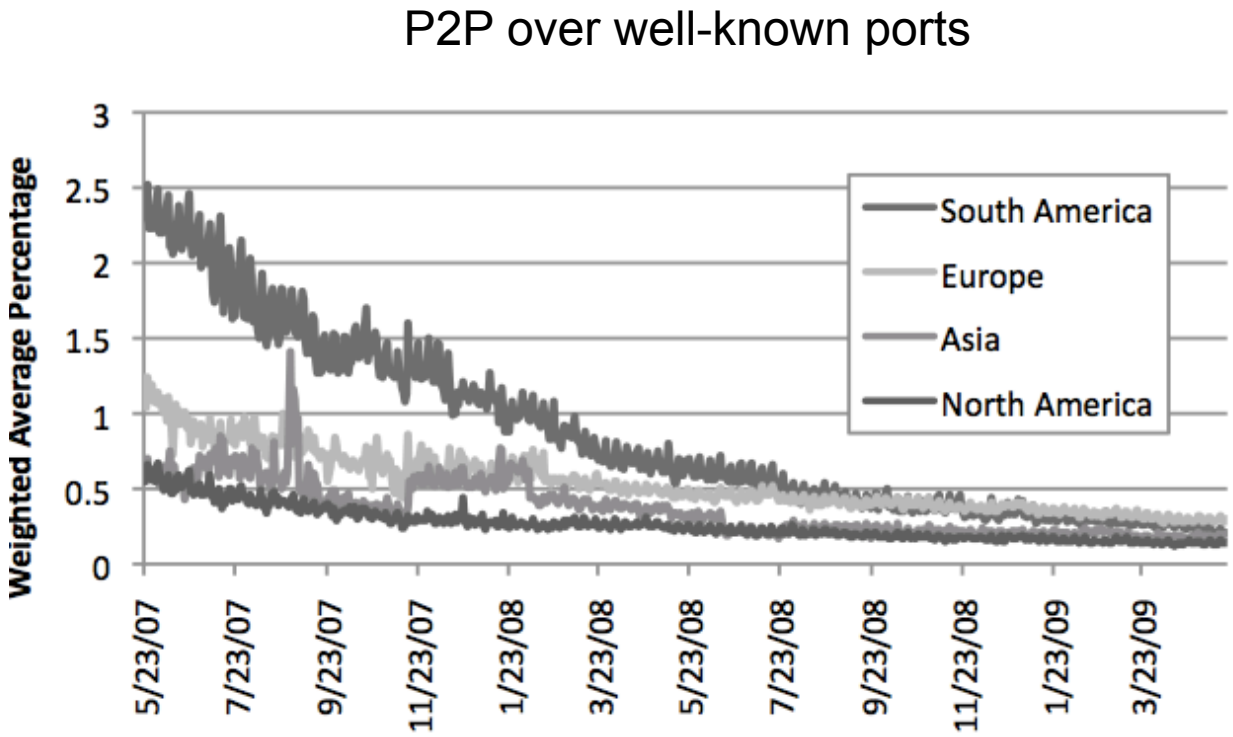
(a) Port / Protocol

Average Percentage	
Web	52.12
Video	0.98
Email	1.54
VPN	0.24
News	0.07
P2P	18.32
Games	0.52
SSH	N/A
DNS	N/A
FTP	0.16
Other	20.54
Unclassified	5.51

(b) Payload Matching

Video represents the second largest and second fastest growing application class

The losers: P2P



P2P saw the largest decline with a drop of 2.8% percentage between 2007 and 2009

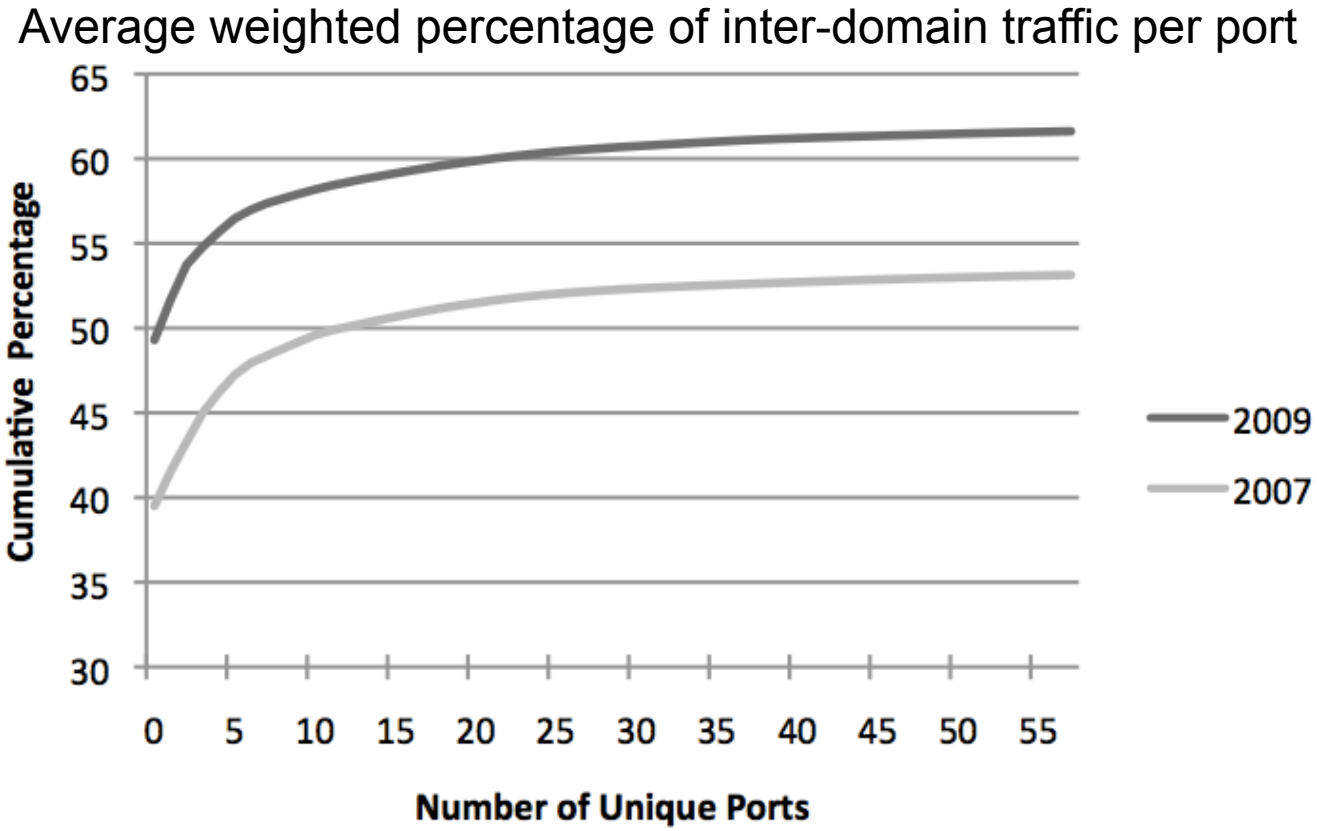
DPI: in 2007 shows P2P at 40% of all traffic, in 2009 shows P2P at 20%

P2P decline

- Reasons?
  - Improvements in P2P client and algorithm efficiency
  - Stealthier P2P clients and algorithm
  - Migration to Tunneled overlays (IPV6)
  - Provider traffic management policies
  - Increased use of P2P encryption
  - Migration to other distribution alternatives: direct download and streaming video
  - Payload analysis also suggests encrypted P2P / other ports represent another 10-15% of uncategorized traffic

  - So we don’t really know whether 20% decline is really that high?

Consolidation



2007: 52 ports contributed 60% of the traffic
2009: 25 ports contributed 60% of inter-domain traffic!

Discussion

- Significant evolution of provider interconnection strategies and resultant inter-domain traffic demands
- Rapid transition to a more densely interconnected and less hierarchical inter-domain Internet topology