An Agent-based Model of Interdomain Interconnection in the Internet

Amogh Dhamdhere (CAIDA/UCSD)

With

Constantine Dovrolis (Georgia Tech)

Aemen Lodhi (Georgia Tech)

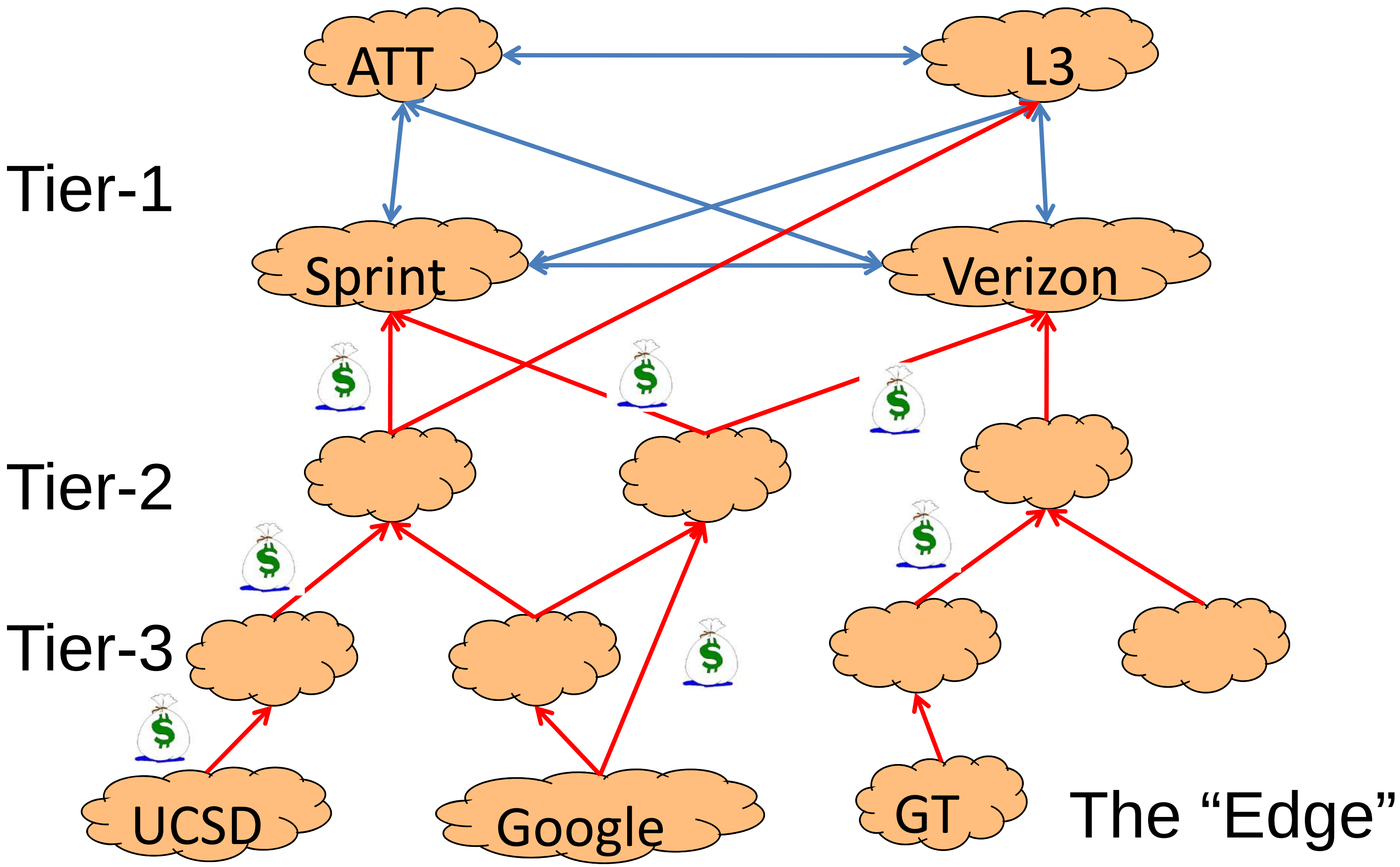
Kc Claffy (CAIDA/UCSD)

Outline

- Motivation
- ITER: A computational model of interdomain interconnection
- Modeling the transition from the "old" to the "new" Internet [CoNEXT 2010]
- Ongoing work: Modeling strategy selection by autonomous networks

[NSF NETSE grant, 2010-2013]

The Interdomain Internet



An “Internet Ecosystem”

- >30,000 autonomous networks independently operated and managed
- The “Internet Ecosystem”
  - Networks differ in their business type
  - Influenced by traffic patterns, application popularity, economics, regulation, policy….
- Network interactions
  - Localized, in the form of bilateral contracts
  - Customer-provider, settlement-free peering, and lots of things in between..

An “Internet Ecosystem”

- >30,000 autonomous networks independently operated and managed
- The “Internet Ecosystem”
  - Networks differ in their business type
  - Influenced by traffic patterns, application popularity, economics, regulation, policy….
- Network interactions
  - Localized, in the form of bilateral contracts
  - Customer-provider, settlement-free peering, and lots of things in between..
- Yes, this is a pretty complex network!

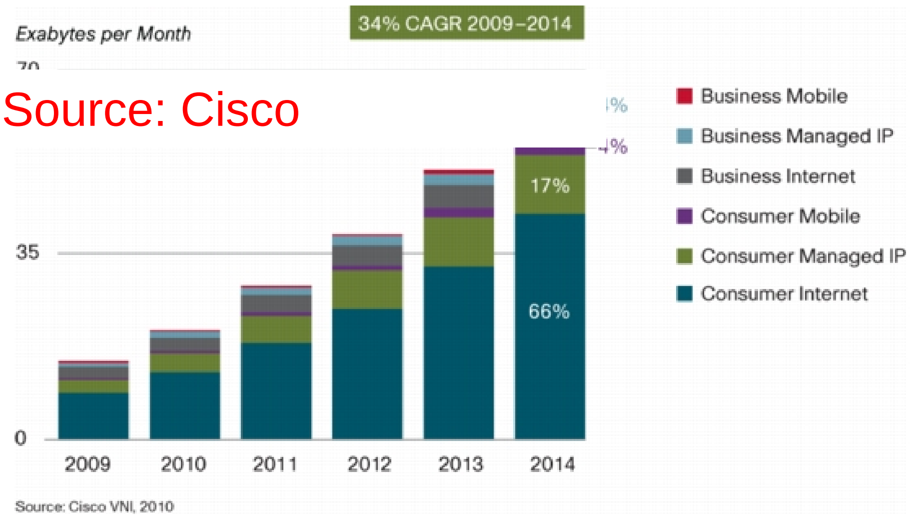
High Level Questions

- How does the Internet ecosystem evolve?
- What is the Internet heading towards?
  - Topology
  - Economics
  - Performance
- Which interconnection strategies of networks optimize their profits, costs and performance?
- How do these strategies affect the global Internet?

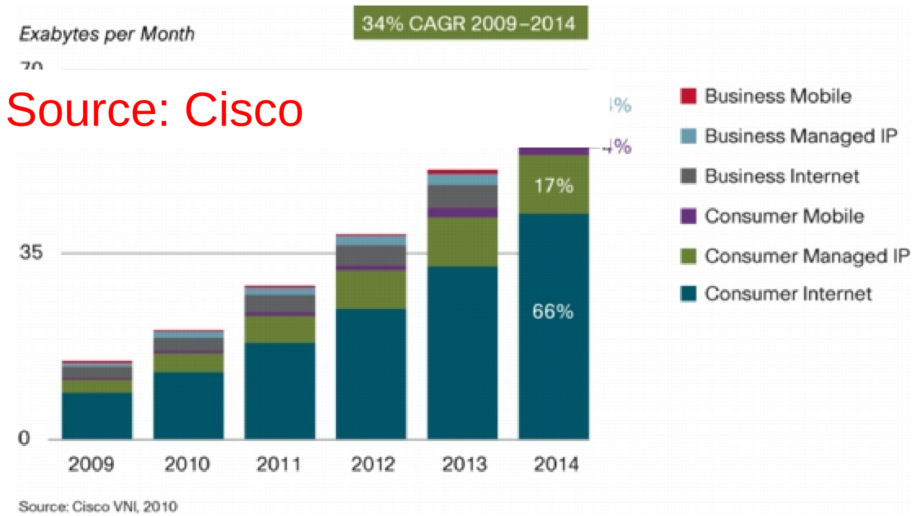
The Dollars Drive Everything!

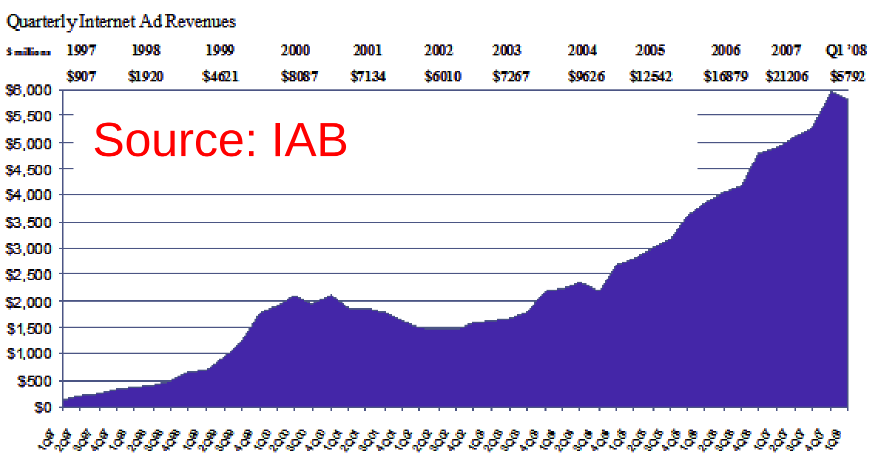
The Dollars Drive Everything!

Source: Cisco

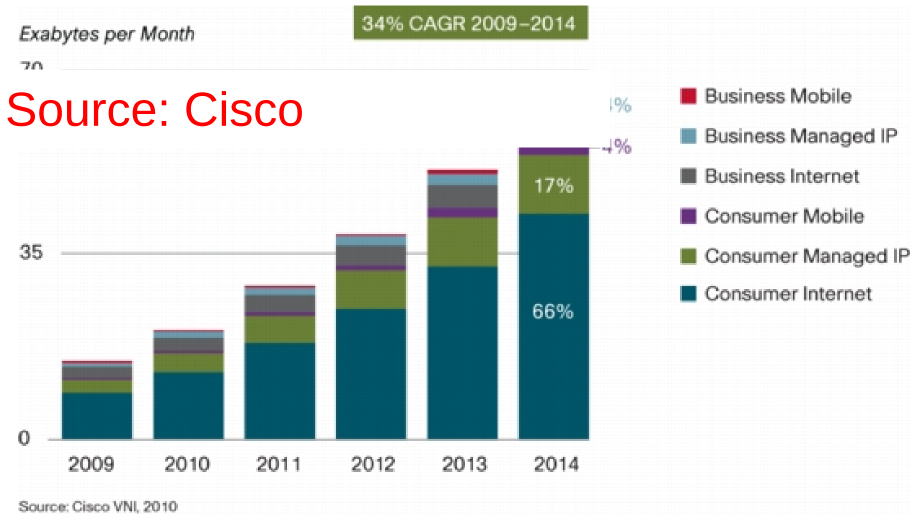


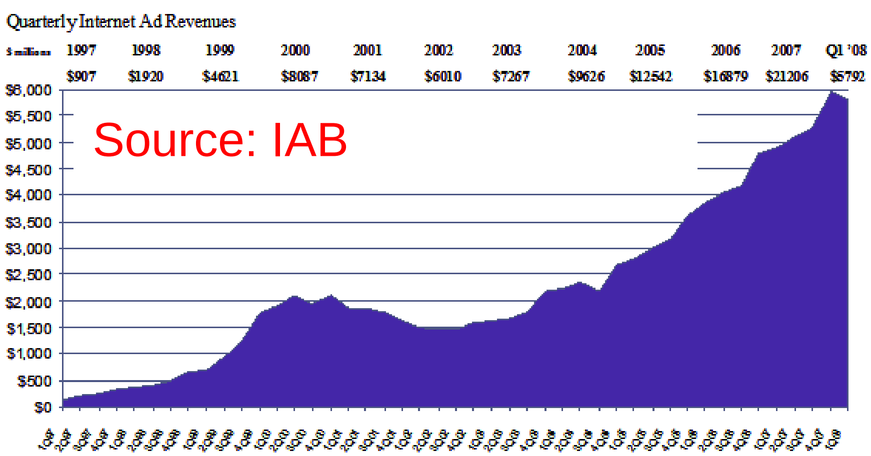
The Dollars Drive Everything!



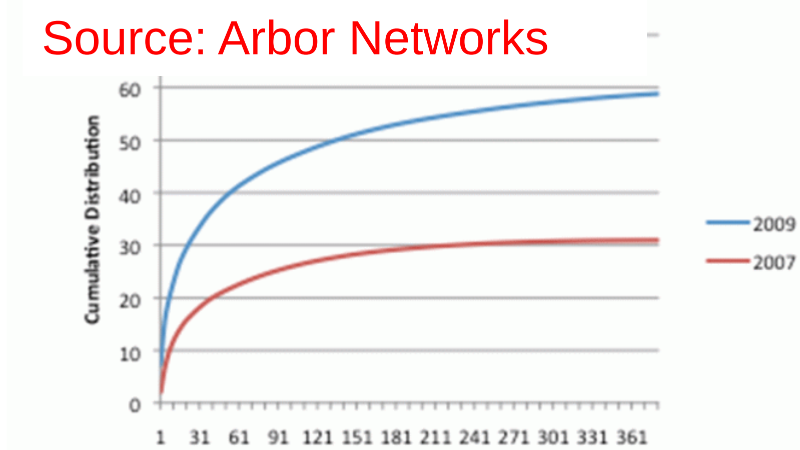


The Dollars Drive Everything!

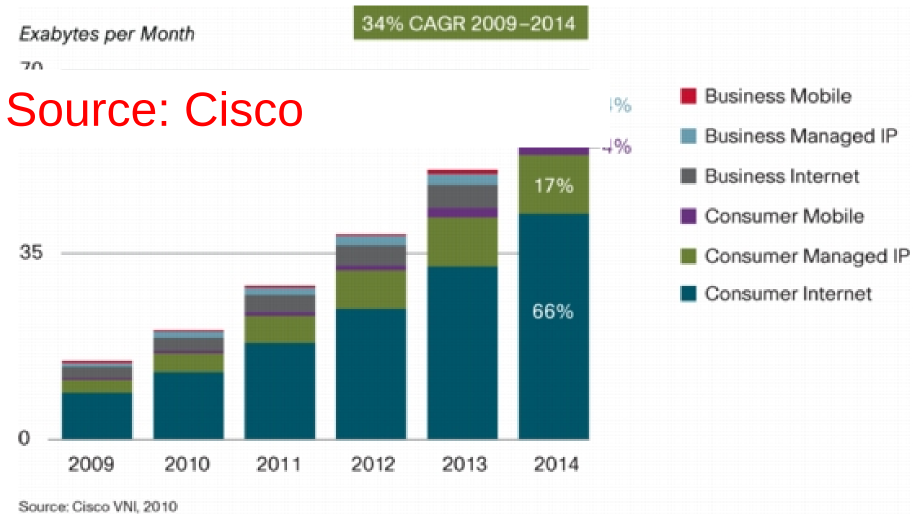


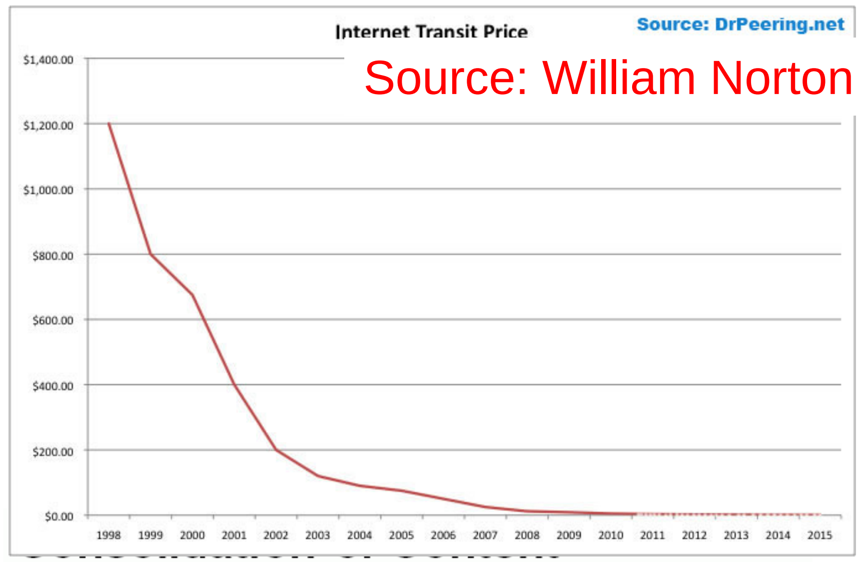


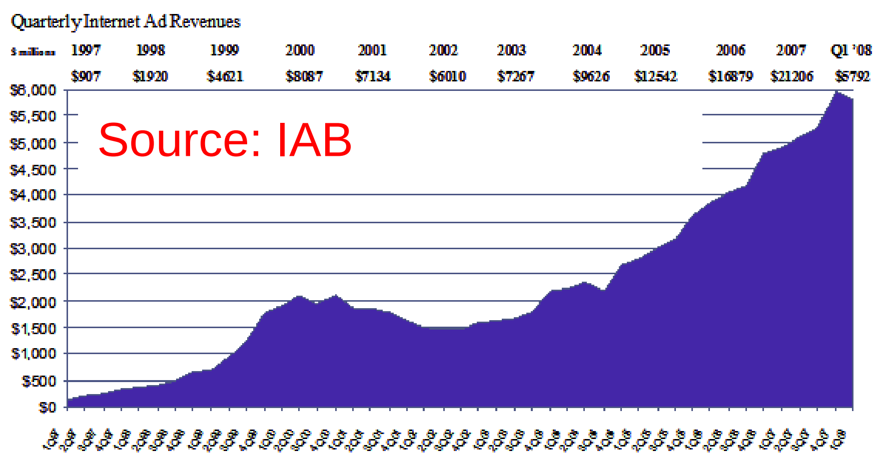
Consolidation of Content

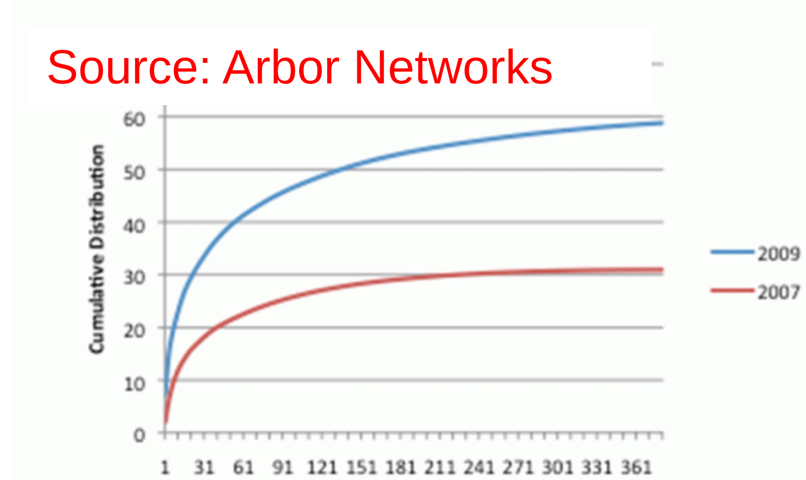


The Dollars Drive Everything!









Economics of the Internet Ecosystem

Share | RSS | Subscriptions | Reprints/License | Print | Email

YouTube May Lose $470 Million In 2009: Analysts

Credit Suisse Report Estimates Video Site Will Generate $240 Million In Revenue

Todd Spangler -- Multichannel News, 4/3/2009 3:43:00 PM

Google's YouTube -- the Internet's most popular video site -- could be on track to lose approximately $470 million in 2009, according to a report Friday by Credit Suisse.

While YouTube remains the leader in online video with 41% share of total domestic video streams, "monetization remains challenging," Credit Suisse analysts Spencer Wang and Kenneth Sena wrote.

According to the firm's analysis of YouTube traffic and ad strategies, the site is on track to generate about $240 million in revenue in 2009, up about 20% year over year.

But the cost of bandwidth, content licensing, ad-revenue shares, hardware storage, sales and marketing and other expenses will total about $711 million, putting YouTube squarely in the red, the Credit Suisse report estimated. Bandwidth accounts for about 51% of expenses -- with a run rate of $1 million per day -- with content licensing accounting for 36%.

"In our view, the issue for YouTube going forward is to increase the percentage of its videos that can be monetized (likely through more deals with content companies) and to drive more advertiser demand through standardization of ad formats and improved ad effectiveness," the analysts wrote.

YouTube, which still derives most of its traffic from user-generated content, has been attempting to increase its lineup of professionally produced content. Earlier this week, for example, YouTube announced a deal with Disney-ABC Television Group and ESPN, which will provide content clips for dedicated channels on the video site.

Credit Suisse projected YouTube will serve 75 billion video streams in 2009, up 38% compared with last year.

To arrive at the estimated $360 million bandwidth tab for YouTube, the analysts assumed the site will receive 375 million unique visitors in 2009 and that a maximum of 20% of those users are on the site at any given time. Credit Suisse's analysis then assumed each user downloads a video at 400 kilobits per second, to yield a peak bit run-rate for YouTube of 30 million megabits per second.

Economics of the Internet Ecosystem

Share | RSS | Subscriptions | Reprints/License | Print | Email

YouTube May Lose $470 Million In 2009: Analysts

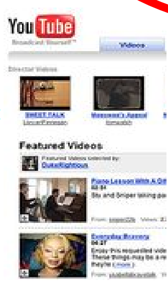
Credit Suisse Report Estimates Video Site Will Generate $240 Million In Revenue

Todd Spangler -- Multichannel News, 4/3/2009 3:43:00 PM

Google's YouTube -- the Internet's most popular video site -- could be on t
million in 2009, according to a report Friday by Credit Suisse.

While YouTube remains the leader in online video with 41% share of total
"monetization remains challenging," Credit Suisse analysts Spencer Wan

According to the firm's analysis of YouTube traffic and ad strategies, the si
million in revenue in 2009, up about 20% year over year.



But the cost of bandwidth, content licensing, ad-revenue shares, hardware storage, sales and marketing and other expenses will total about $711 million, putting YouTube squarely in the red, the Credit Suisse report estimated. Bandwidth accounts for about 51% of expenses -- with a run rate of $1 million per day -- with content licensing accounting for 36%.

"In our view, the issue for YouTube going forward is to increase the percentage of its videos that can be monetized (likely through more deals with content companies) and to drive more advertiser demand through standardization of ad formats and improved ad effectiveness," the analysts wrote.

YouTube, which still derives most of its traffic from user-generated content
lineup of professionally produced content. Earlier this week, for example, Y
Disney-ABC Television Group and ESPN, which will provide content clips fo
site.

Credit Suisse projected YouTube will serve 75 billion video streams in 200

To arrive at the estimated $360 million bandwidth tab for YouTube, the ana
375 million unique visitors in 2009 and that a maximum of 20% of those u
time. Credit Suisse's analysis then assumed each user downloads a vide
a peak bit run-rate for YouTube of 30 million megabits per second.

PREVIOUS POST NEXT POST

YouTube's Bandwidth Bill Is Zero. Welcome to the New Net

By Ryan Singel | October 16, 2009 | 2:10 pm | Categories: Broadband, Miscellaneous

YouTube may pay less to be online than you do, a new report on internet connectivity suggests, calling into question a recent analysis arguing Google's popular video service is bleeding money and demonstrating how the internet has continued to morph to fit user's behavior.

In fact, with YouTube's help, Google is now responsible for at least 6 percent of the internet's traffic, and likely more — and may not be paying an ISP at all to serve up all that content and attached ads.

Credit Suisse made headlines this summer when it estimated that YouTube was binging on bandwidth, losing Google a half a billion dollars in 2009 as it streams 75 billion videos. But a new report from Arbor Networks suggests that Google's traffic is approaching 10 percent of the net's traffic, and that it's got so much fiber optic cable, it is simply trading traffic, with no payment involved, with the net's largest ISPs.

"I think Google's transit costs are close to zero," said Craig Labovitz, the chief scientist for Arbor Networks and a longtime internet researcher. Arbor Networks, which sells network monitoring equipment used by about 70 percent of the net's ISPs, likely knows more about the net's ebbs and flows than anyone outside of the National Security Agency.

And the extraordinary fact that a website serving nearly 100 billion videos a year has no bandwidth bill means the net isn't the network it used to be.

Economics of the Internet Ecosystem

Share RSS

YouTube May L

Credit Suisse Repo

Todd Spangler -- Multich

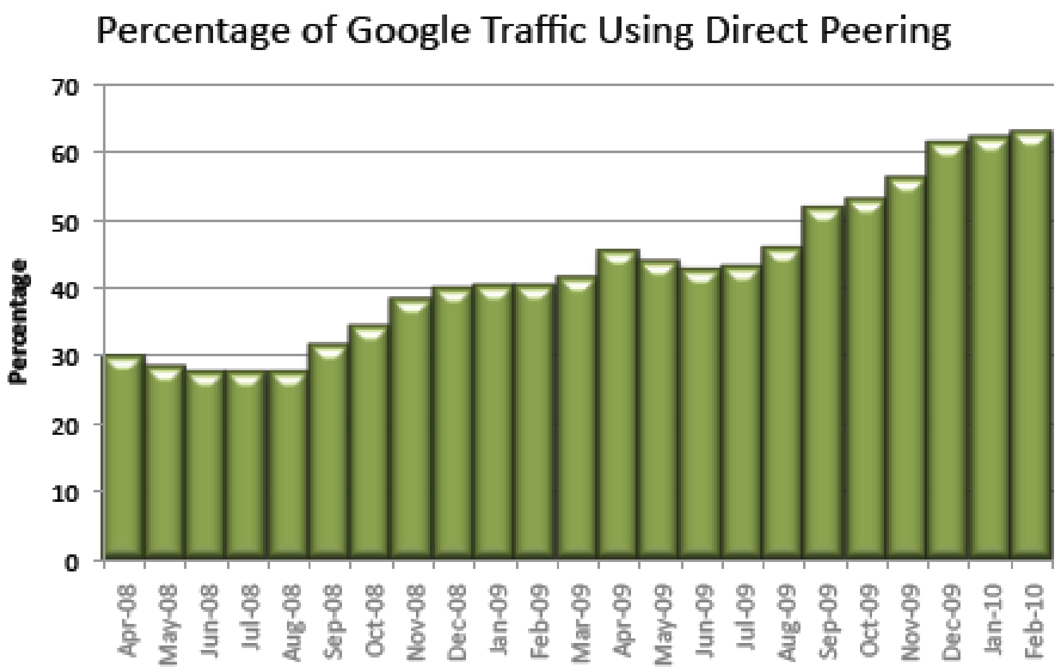
Google's YouTube -- the l
million in 2009, according

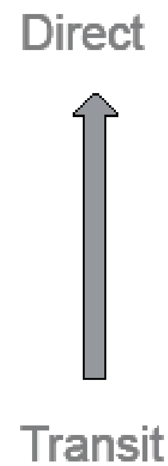
While YouTube remains t
"monetization remains ch

According to the firm's an
million in revenue in 2009

But the cost of bandwidth,
ad-revenue shares, hardw
marketing and other expe
million, putting YouTube s
Suisse report estimated.
51% of expenses -- with a
-- with content licensing a

"In our view, the issue for
increase the percentage o
monetized (likely through
companies) and to drive r
through standardization o
effectiveness," the analys





NEXT POST

to the New

tivity suggests, calling
money and

e internet's traffic, and
attached ads.

YouTube, which still derives most of its traffic from user-generated content
lineup of professionally produced content. Earlier this week, for example, Y
Disney-ABC Television Group and ESPN, which will provide content clips fo
site.

Credit Suisse projected YouTube will serve 75 billion video streams in 20

To arrive at the estimated $360 million bandwidth tab for YouTube, the ana
375 million unique visitors in 2009 and that a maximum of 20% of those u
time. Credit Suisse's analysis then assumed each user downloads a vide
a peak bit run-rate for YouTube of 30 million megabits per second.

inging on bandwidth,
losing Google a half a billion dollars in 2009 as it streams 75 billion videos. But a new report from Arbor Networks suggests that Google's traffic is approaching 10 percent of the net's traffic, and that it's got so much fiber optic cable, it is simply trading traffic, with no payment involved, with the net's largest ISPs.

"I think Google's transit costs are close to zero," said Craig Labovitz, the chief scientist for Arbor Networks and a longtime internet researcher. Arbor Networks, which sells network monitoring equipment used by about 70 percent of the net's ISPs, likely knows more about the net's ebbs and flows than anyone outside of the National Security Agency.

And the extraordinary fact that a website serving nearly 100 billion videos a year has no bandwidth bill means the net isn't the network it used to be.

Economics of the Internet Ecosystem

Share RSS

YouTube May L
Credit Suisse Repo

Todd Spangler -- Multich

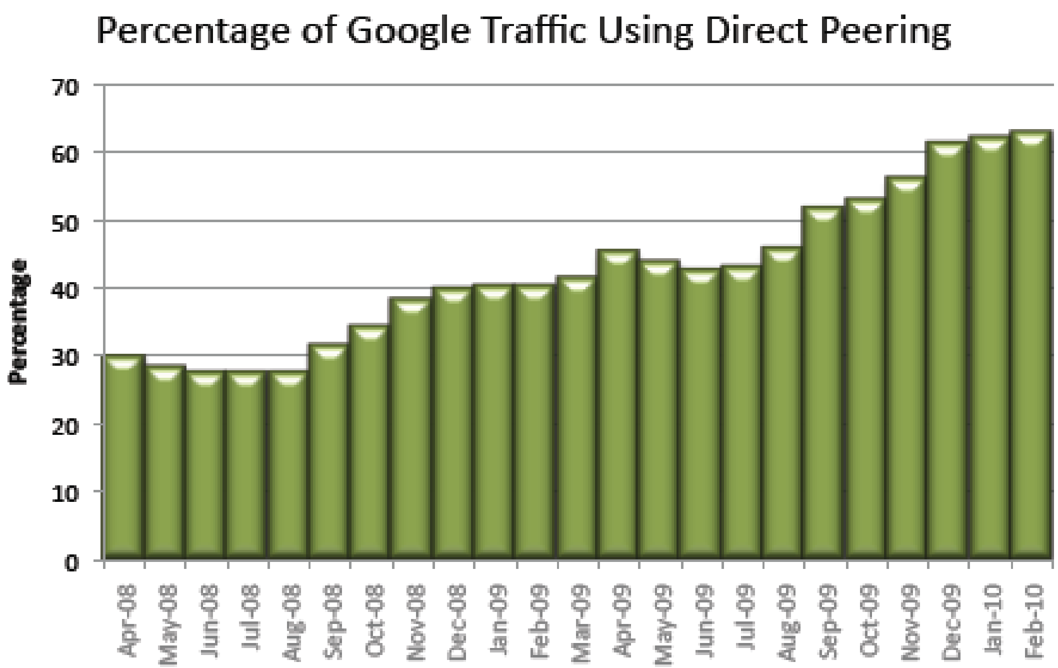
Google's YouTube -- the l
million in 2009, according

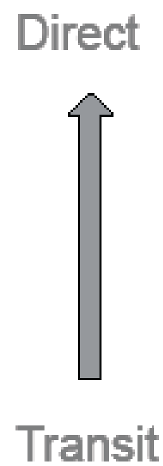
While YouTube remains t
"monetization remains ch

According to the firm's an
million in revenue in 2009

But the cost of bandwidth,
ad-revenue shares, hardw
marketing and other expe
million, putting YouTube s
Suisse report estimated.
51% of expenses -- with a
-- with content licensing a

"In our view, the issue for
increase the percentage o
monetized (likely through
companies) and to drive r
through standardization o
effectiveness," the analys





NEXT POST

to the New

tivity suggests, calling
money and

e internet's traffic, and
attached ads.

YouTube, which still derives most of its traffic from user-generated content
lineup of professional
Disney-ABC Televisio
site.

inging on bandwidth,
om Arbor
it's got so
st ISPs.

Credit Suisse projecte

How do we make sense of all this?

To arrive at the estimated $360 million bandwidth tab for YouTube, the ana
375 million unique visitors in 2009 and that a maximum of 20% of those u
time. Credit Suisse's analysis then assumed each user downloads a vide
a peak bit run-rate for YouTube of 30 million megabits per second.

"I think Google's transit costs are close to zero," said Craig Labovitz, the chief scientist for Arbor Networks and a longtime internet researcher. Arbor Networks, which sells network monitoring equipment used by about 70 percent of the net's ISPs, likely knows more about the net's ebbs and flows than anyone outside of the National Security Agency.

And the extraordinary fact that a website serving nearly 100 billion videos a year has no bandwidth bill means the net isn't the network it used to be.

Economically-principled models

- Objective: understand the structure and dynamics of the Internet ecosystem from an economic perspective
- Capture interactions between network business relations, internetwork topology, routing policies, and resulting interdomain traffic flow
- Create a scientific basis for modeling Internet interconnection and dynamics based on empirical data

Previous Work

- "Descriptive"
  - Match graph properties e.g. degree distribution
- Homogeneity
  - Nodes and links all the same
- Game theoretic, analytical
  - Restrictive assumptions
- Little relation to real-world data

Previous Work

- “Descriptive”
  - Match graph properties e.g. degree distribution
- Homogeneity
  - Nodes and links all the same
- Game theoretic, analytical
  - Restrictive assumptions
- Little relation to real-world data

Previous Work

- “Descriptive”
  - Match graph properties e.g. degree distribution
- Homogeneity
  - Nodes and links all the same
- Game theoretic, analytical
  - Restrictive assumptions
- Little relation to real-world data

- “Bottom-up”
  - Model the actions of individual networks
- Heterogeneity
  - Networks with different incentives, link types
- Computational, agent-based
  - As much realism as possible
- Parameterize/validate using real data

Outline

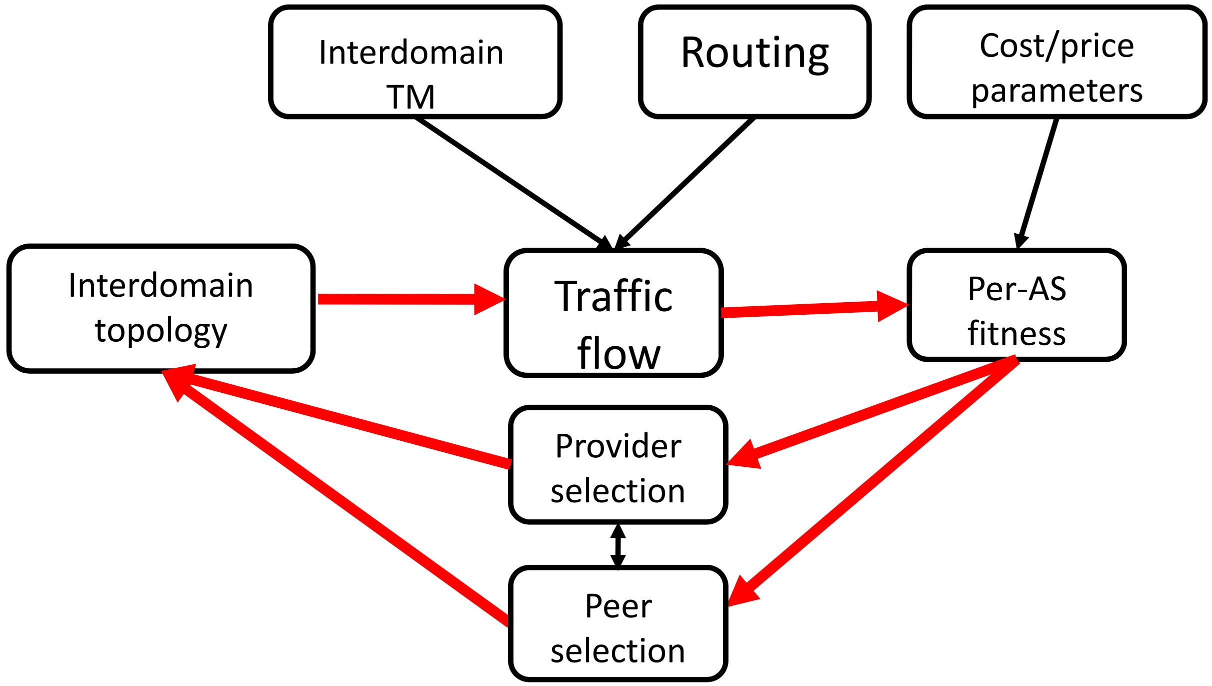
- Motivation
- ITER: A computational model of interdomain interconnection
- Modeling the transition from the “old” to the “new” Internet [CoNEXT 2010]
- Ongoing work: Modeling strategy selection by autonomous networks

[NSF NETSE grant, 2010-2013]

The ITER Model

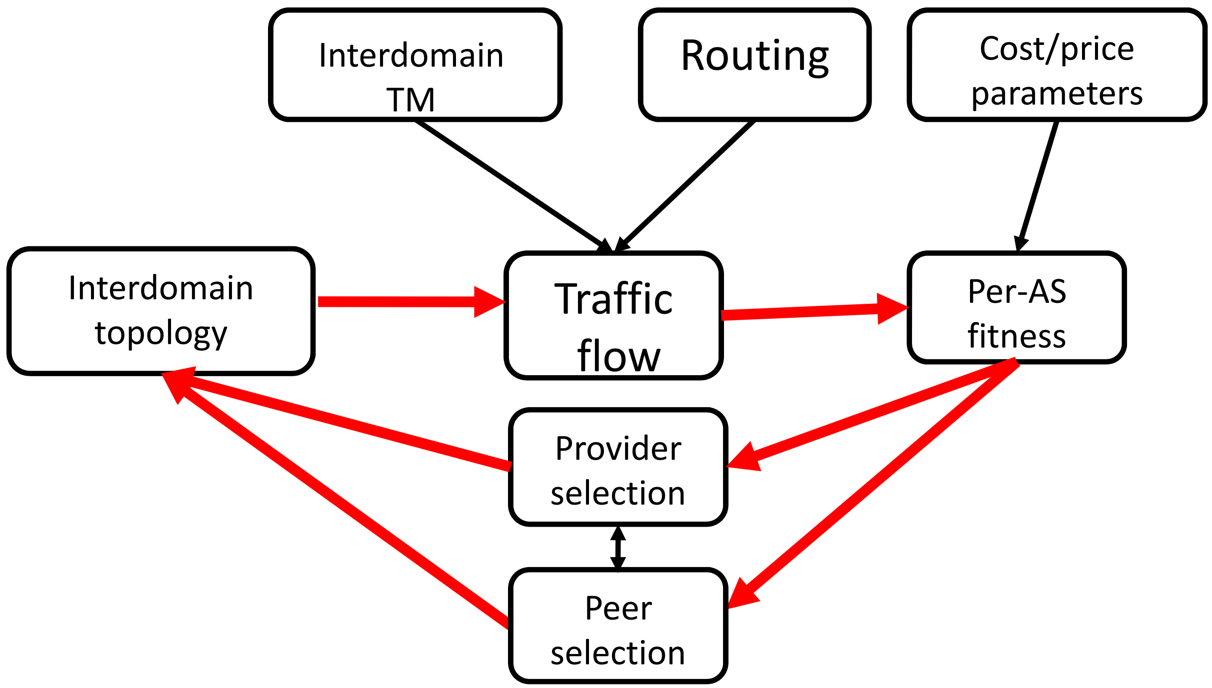
- Agent-based computational model to answer "what-if" questions about Internet evolution
- Inputs: According to the best available data…
  - Network types based on business function
  - Peer/provider selection methods
  - Geographical constraints
  - Pricing/cost parameters
  - Interdomain traffic matrix
- Output: Equilibrium internetwork topology, traffic flow, per-network fitness

The ITER approach



- Compute equilibrium: no network has the incentive to change its providers/peers

The ITER approach



- Measure topological and economic properties of equilibrium e.g., path lengths, which providers are profitable, who peers with whom

Why Study Equilibria?

- The Internet is never at equilibrium, right?
  - Networks come and go, traffic patterns change, pricing/cost structures change, etc.…
- Studying equilibria tells us what's the best that networks could do under certain traffic/economic conditions, and what that means for the Internet as a whole
- If those conditions change, we need to re-compute equilibria

ITER: Network Types

- Enterprise Customers (EC)
  - Stub networks at the edge, e.g. Georgia Tech
- Transit Providers
  - Regional in scope (STP), e.g. Comcast
  - “Tier-1” or global (LTP), e.g., AT&T
- Content Providers (CP)
  - Major sources of content, e.g. Google

ITER: Provider and Peer Selection

- Provider selection
  - Choose providers based on measure of the “size” of a provider
- Peer selection
  - Peer based on total traffic handled; Approximates the “equality” of two ISPs

ITER: Economics, Routing and Traffic Matrix

- Realistic transit, peering and operational costs

- BGP-like routing policies

- Traffic matrix
  - Heavy-tailed content popularity and consumption by sinks

Computing Equilibrium

- Situation where no network has the incentive to change its connectivity
- Too complex to find analytically: Solve using agent-based simulations
- Computation
  - Proceeds iteratively, networks "play" in sequence, adjust their connectivity
  - Compute routing, traffic flow, AS fitness
  - Repeat until no player has incentive to move

Properties of the equilibrium

- Is an equilibrium always found?
  - Yes, in most cases
- Is the equilibrium unique?
  - No, can depend on playing sequence
- Multiple runs with different playing sequence
  - Per-network properties vary widely across runs
  - Macroscopic properties show low variability

Outline

- Motivation
- ITER: A computational model of interdomain interconnection
- Modeling the transition from the "old" to the "new" Internet [CoNEXT 2010]
- Ongoing work: Modeling strategy selection by autonomous networks

[NSF NETSE grant, 2010-2013]

Recent Trends: Arbor Networks Study

- The Old Internet (late 90s – 2007)
- Content providers generated small fraction of total traffic
- Content providers were mostly local
- Peering was restrictive

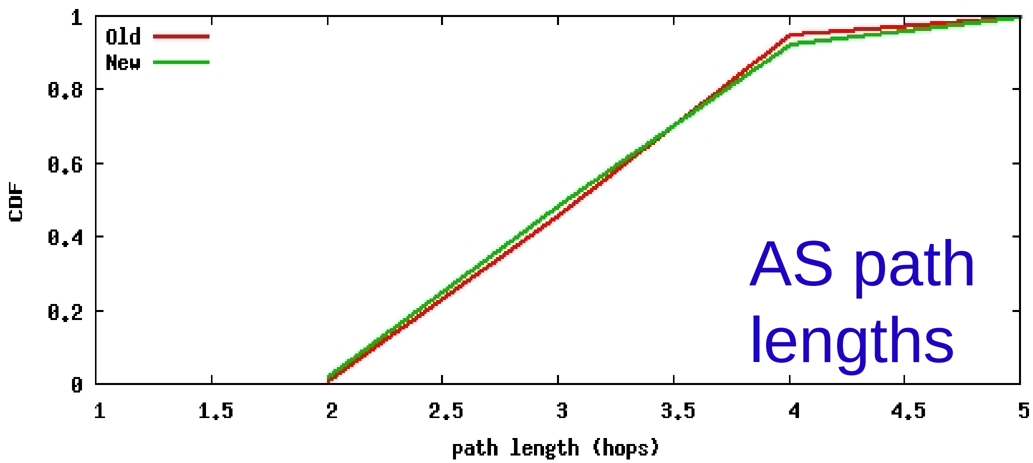
- The New Internet (2007 onwards)
- Largest content providers generate large fraction of total traffic
- Content providers are present everywhere
- Peering is more open

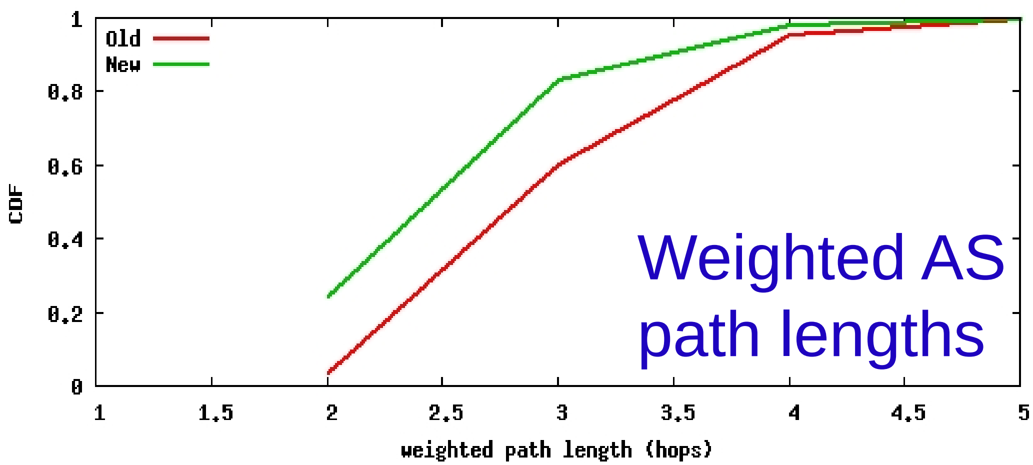
"Internet Interdomain Traffic", Labovitz et al., Sigcomm 2010

Plugging into ITER

- Simulate two instances of ITER: “Old” and “New” Internet
- Change three parameters
  - Fraction of traffic sourced by CPs
  - Geographical spread of CPs
  - Peering openness
- Compute equilibria for these two instances
  - Compare topological, economic properties

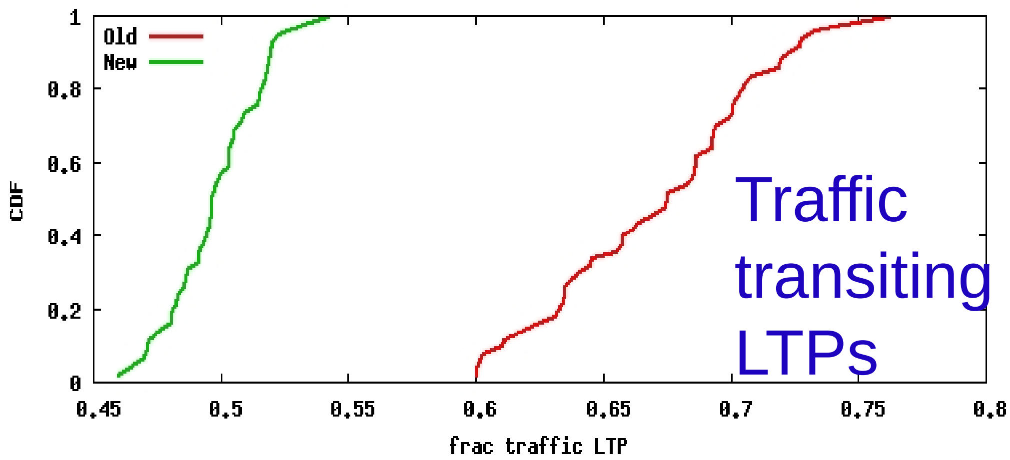
ITER Sims: End-to-end Paths

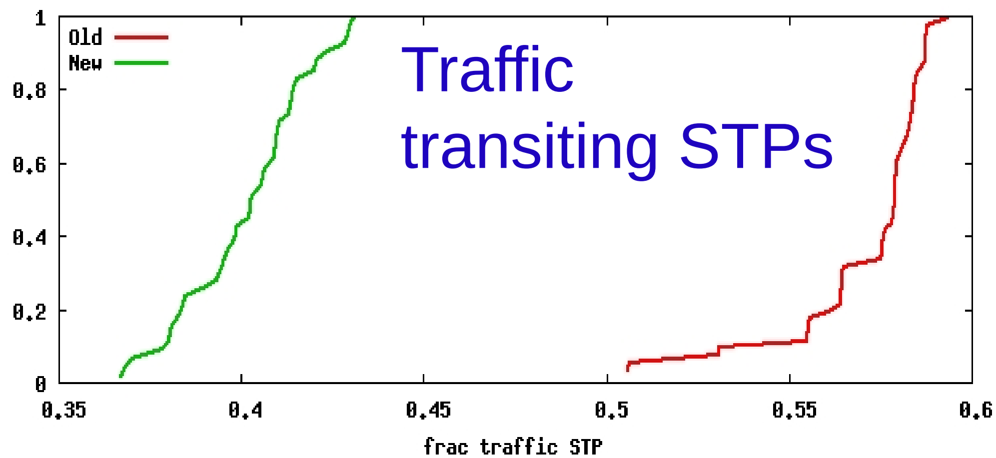




- End-to-end paths weighted by traffic are shorter in the "new" Internet
- Paths carrying the most traffic are shorter

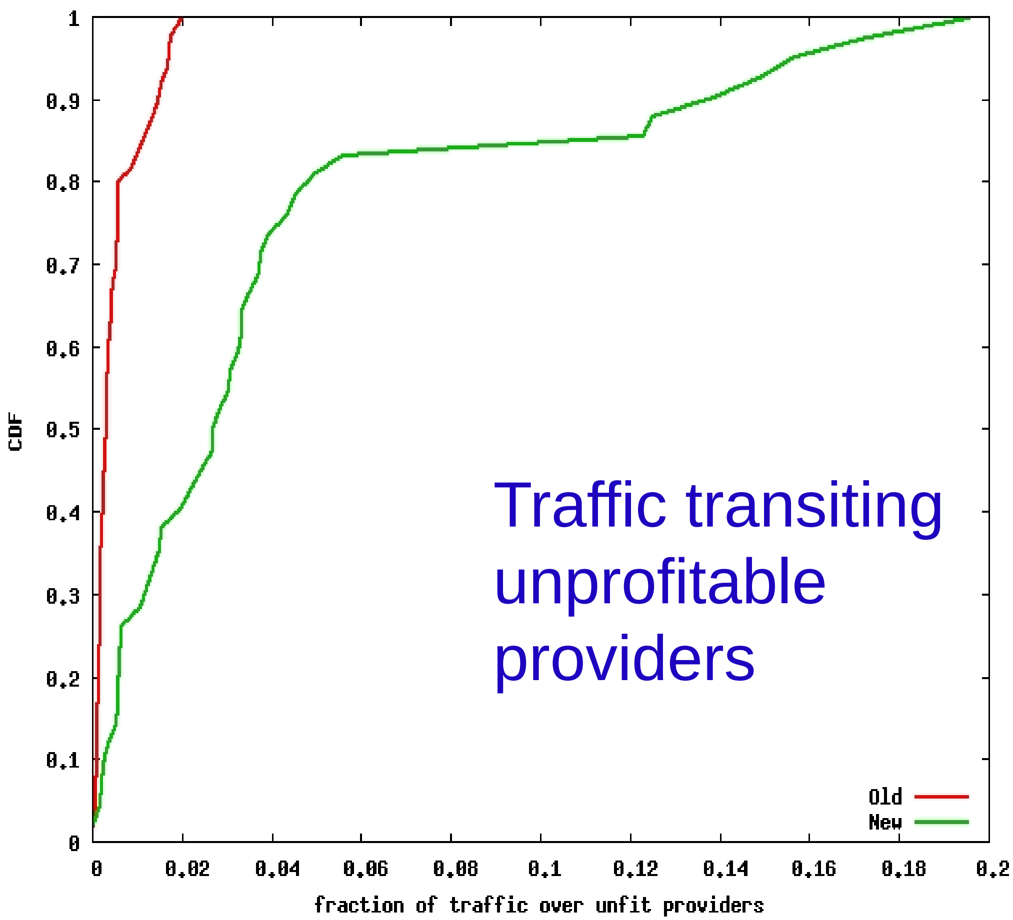
ITER Sims: Traffic Transiting Transit Providers





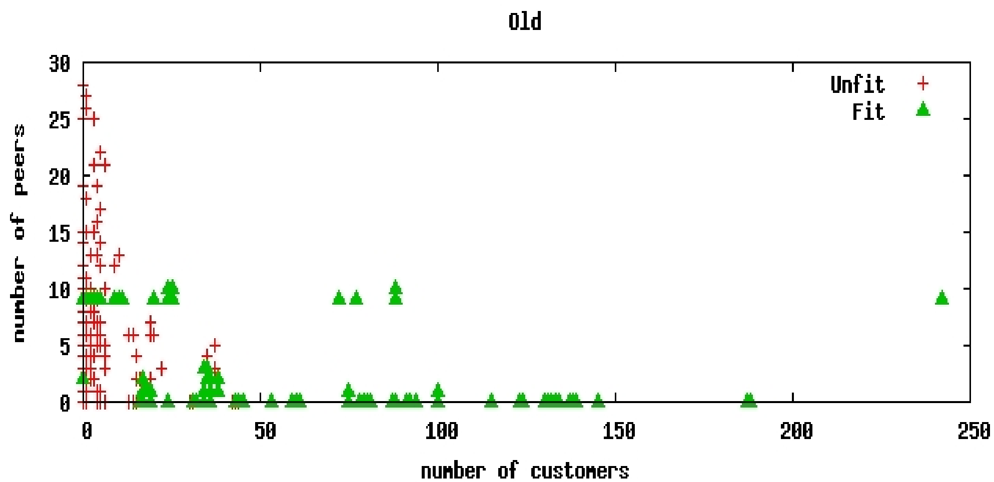
- Traffic bypasses transit providers
- More traffic flows directly on peering links
- Implication: Transit providers lose money!
- Content providers get richer

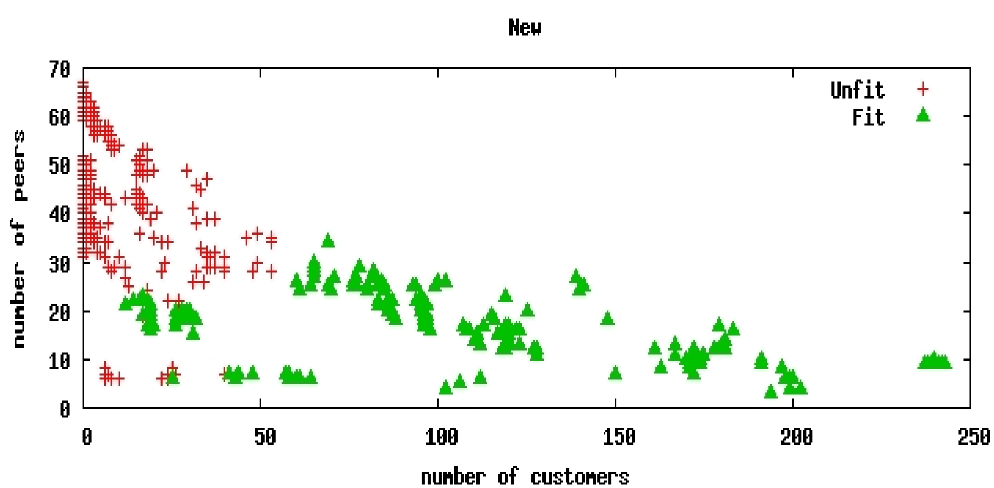
ITER Sims: Traffic Over Unprofitable Providers



- More transit providers are unprofitable in the new Internet
- These unprofitable providers still have to carry traffic!
- Possibility of mergers, bankruptcies or acquisitions

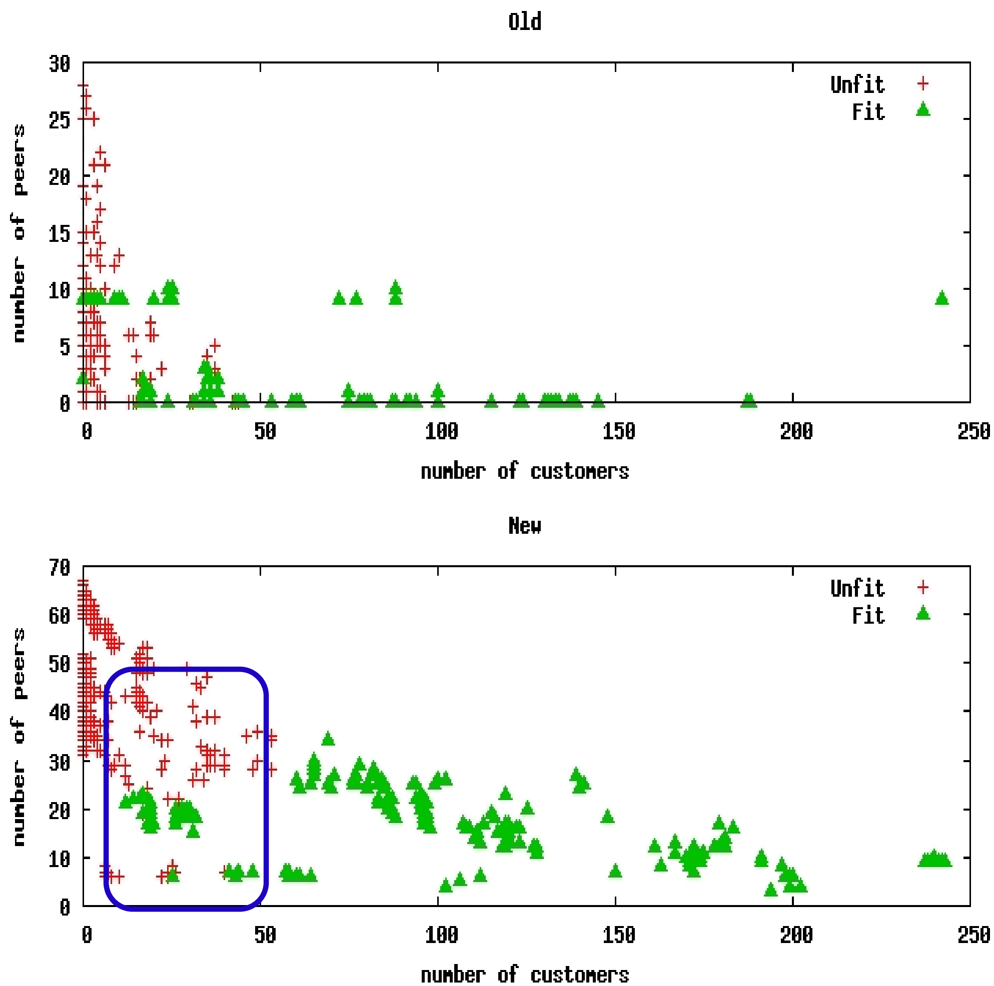
ITER Sims: Peering in the New Internet





- Transit providers need to peer strategically in the “new” Internet
- STPs peering with CPs: saves transit costs
- LTPs peering with CPs: attracts traffic that would have bypassed them

ITER Sims: Peering in the New Internet



- Transit providers need to peer strategically in the “new” Internet
- STPs peering with CPs: saves transit costs
- LTPs peering with CPs: attracts traffic that would have bypassed them

Outline

- Motivation
- ITER: A computational model of interdomain interconnection
- Modeling the transition from the “old” to the “new” Internet [CoNEXT 2010]
- Ongoing work: Modeling strategy selection by autonomous networks

[NSF NETSE grant, 2010-2013]

Strategy selection by Autonomous Networks

- So far, every network used a fixed strategy
- But network strategies can evolve over time
- Can we model how networks dynamically change their peer selection strategies?
  - What is the best strategy for different network types?

Myopic Strategy Selection

- Networks still “play” in sequence
- In each move, a network
  - Tries to interconnect using each available peering strategy, assuming it knows the peering strategies of other networks
  - Computes fitness for each possible strategy
  - Chooses strategy that results in best fitness
- Compute a “strategy equilibrium” where each network settles on a peering strategy

Early (surprising?) Results

- Studied three strategies: Open peering, selective peering, restrictive peering
- With myopic strategy selection, every network ends up wanting to peer openly
- ISPs that peer openly do worse than if they peered selectively or restrictively
- Is this because of
  - Myopic strategy selection?
  - No co-ordination between ISPs?
  - Non-economic considerations?

In the Real World

- There is a trend towards more open peering (measured in real data from peeringDB)
- But we do not see all ISPs peering openly
- So what prevents the “open peering epidemic” in the real world?
- Currently studying: co-ordination (coalitions) between ISPs
- But perhaps it is non-economic factors that prevent the system from collapsing!

Summary

- We need realistic, economically-principled models to make sense of the economics behind interdomain interconnection
- We developed ITER, a computational model of interdomain interconnection
- Currently working on modeling strategy selection by autonomous networks
- Your feedback is welcome!

Thanks!

amogh@caida.org

www.caida.org/~amogh

Backup slides

Avoiding “garbage-in, garbage-out”

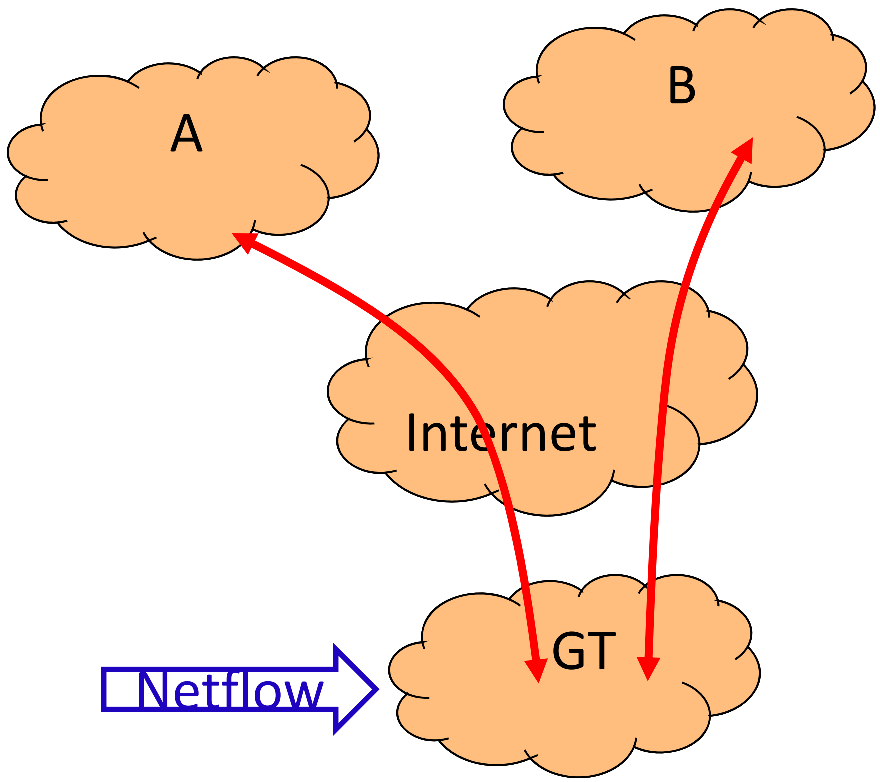
- Models are only as good as the data you provide as input
- How do we get the best possible data to parameterize ITER-like models?
- What data do we need?
  - Interdomain traffic patterns
  - Peering policies
  - Geographical presence of networks
  - Cost/pricing structures

Avoiding “garbage-in, garbage-out”

- Models are only as good as the data you provide as input
- How do we get the best possible data to parameterize ITER-like models?
- What data do we need?
  - Interdomain traffic patterns
  - Peering policies
  - Geographical presence of networks
  - Cost/pricing structures

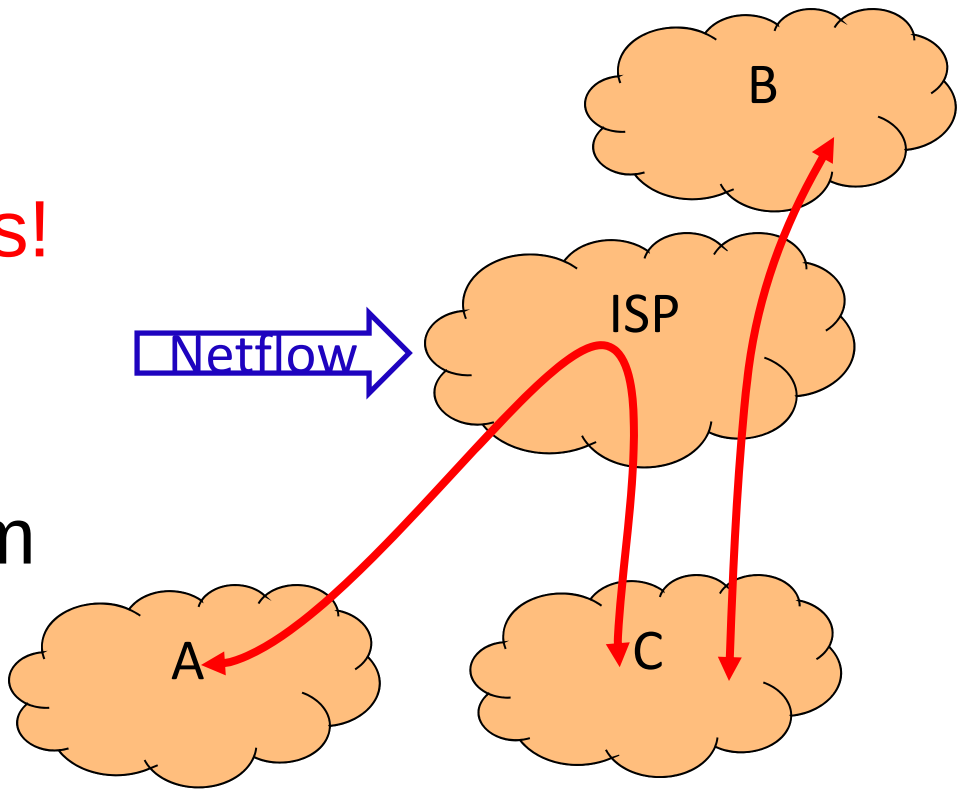
Measuring Interdomain Traffic

- We don't really know how much traffic each pair of networks exchanges!
- Measure qualitative properties of the interdomain TM from different vantage points



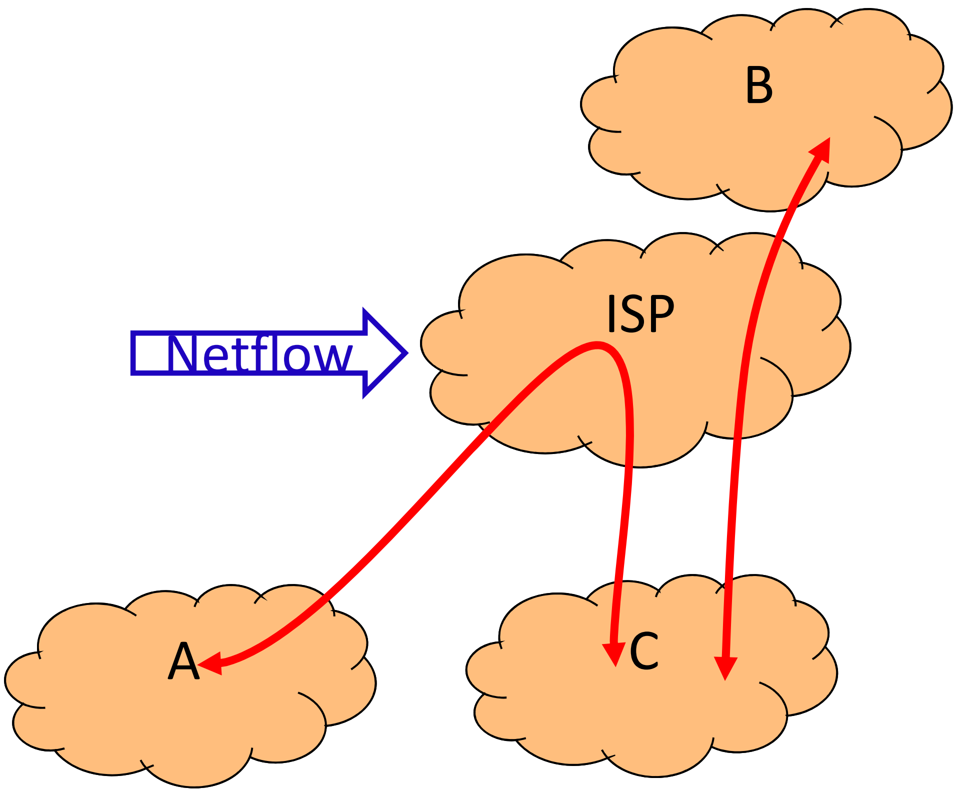
Measuring Interdomain Traffic

- We don't really know how much traffic each pair of networks exchanges!
- Measure qualitative properties of the interdomain TM from different vantage points



Measuring Interdomain Traffic

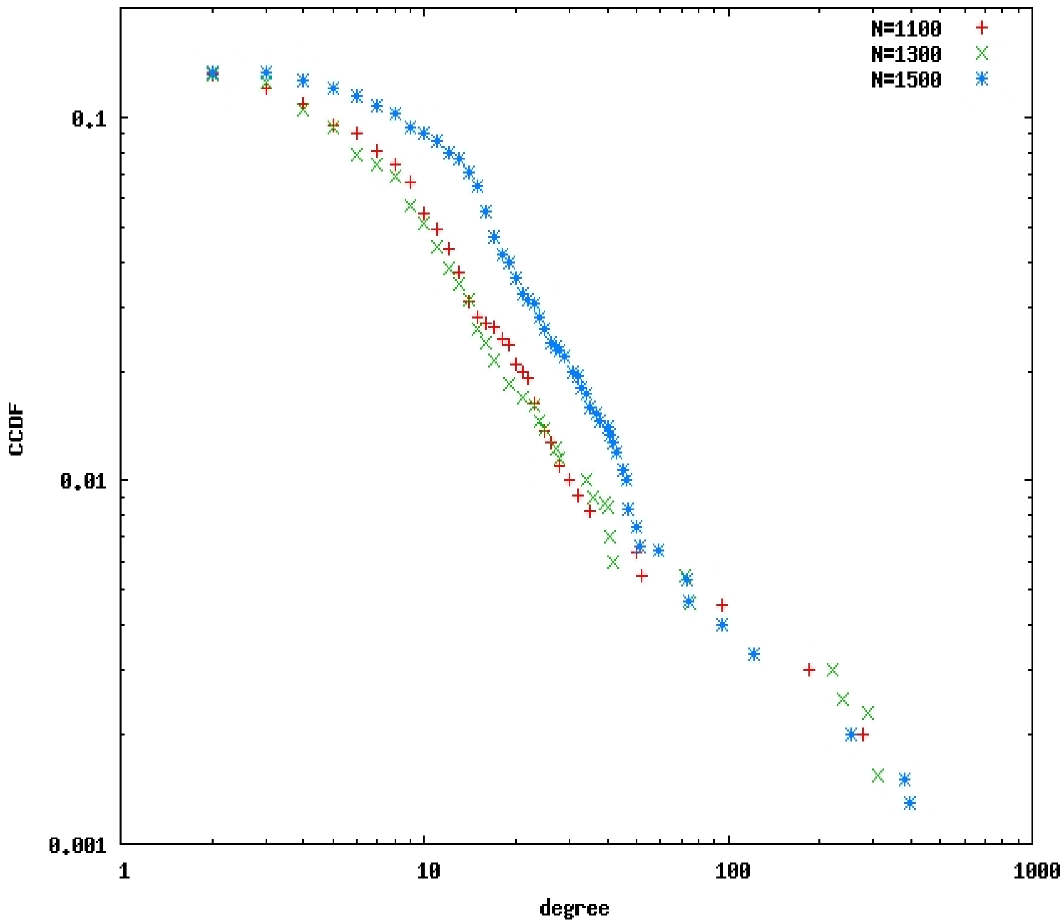
- We need data from as many vantage points as possible!
- Currently working with GEANT, SWITCH, Georgia Tech
- Let us know if you can help!



Validation

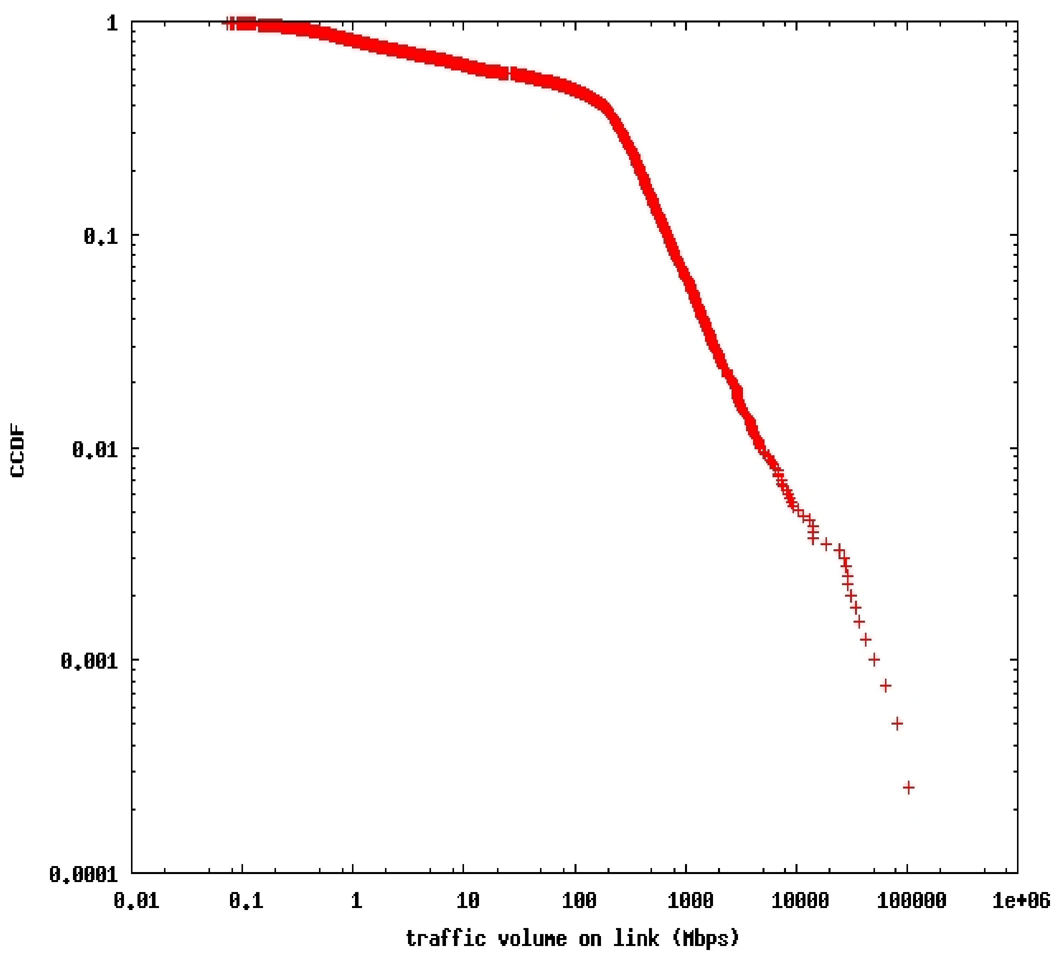
- Validation of a model that involves traffic, topology, economics and network actions is hard!
- “Best-effort” parameterization and validation
- Parameterized transit, peering and operational costs, traffic matrix properties, geographical spread using best available data

Validation



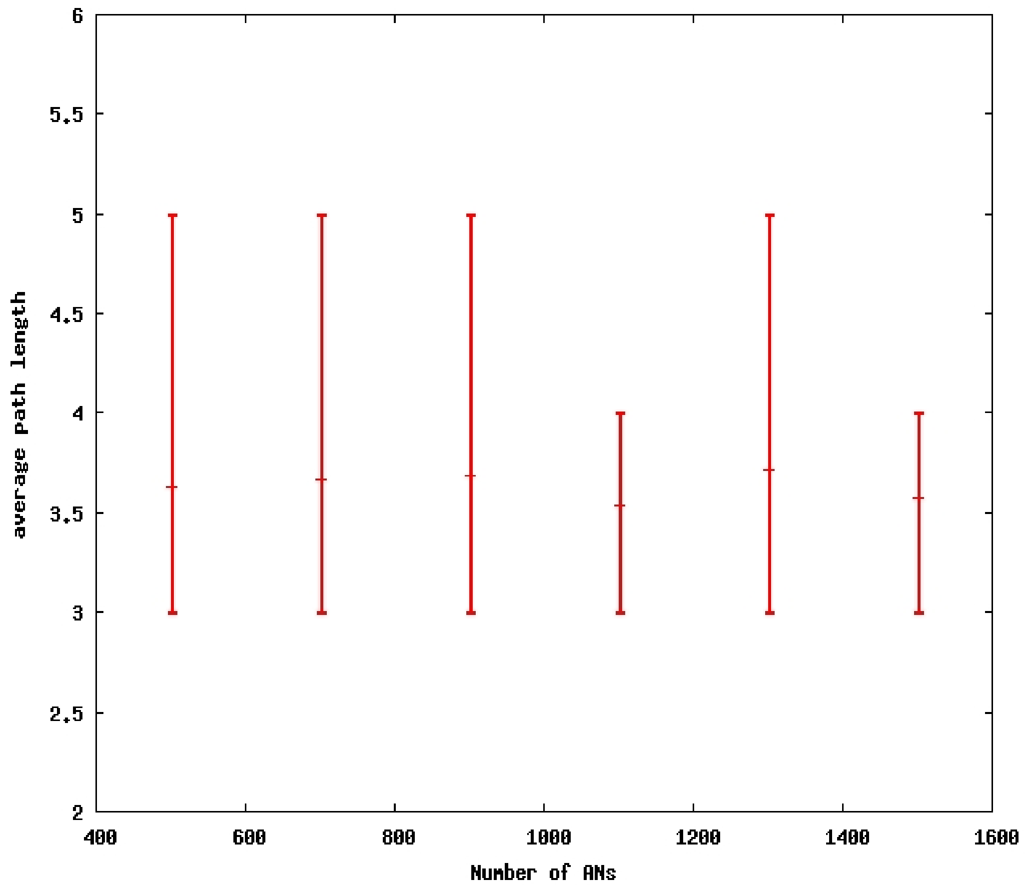
- ITER produces networks with heavy-tailed degree distribution

Validation



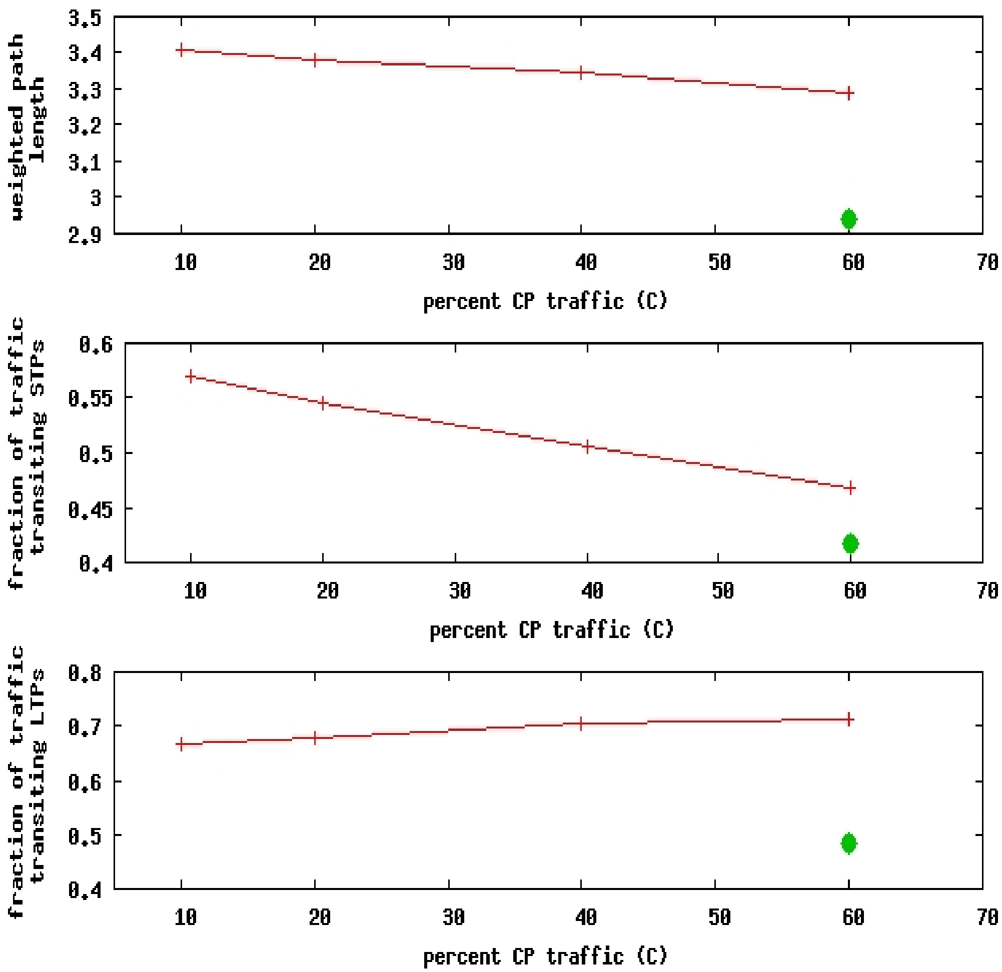
- ITER produces networks with a heavy-tailed distribution of link loads

Validation



- Average path lengths stay almost constant as the network size is increased

Three Factors

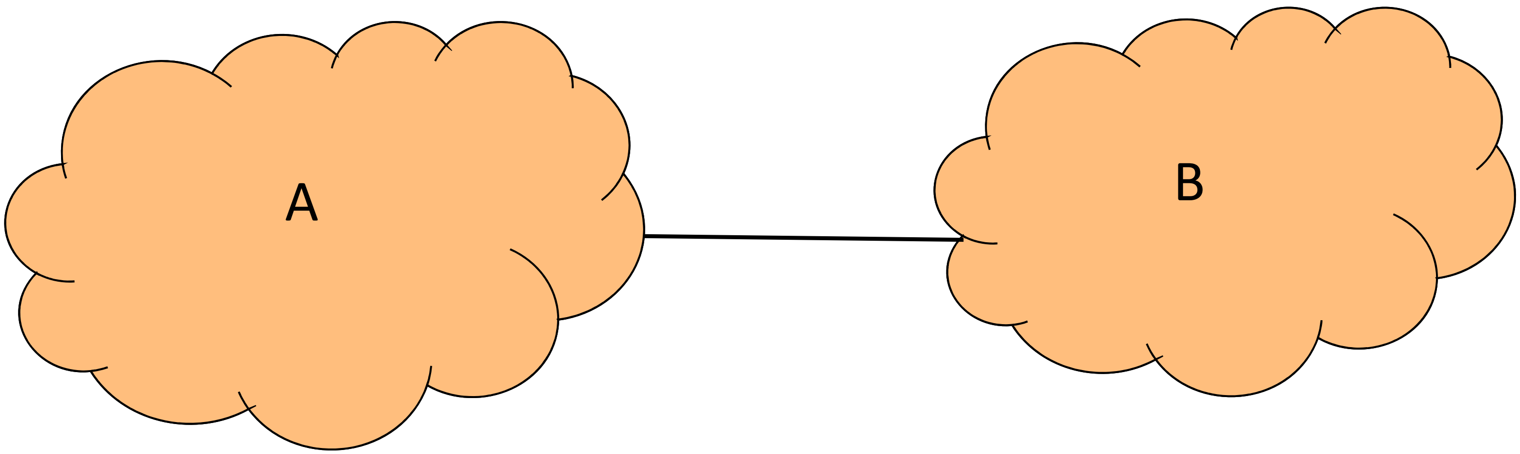


- Fraction of traffic sourced by CPs
- Geographical presence of CPs
- Peering openness
- All three factors need to change to see the differences between the “old” and “new” Internet

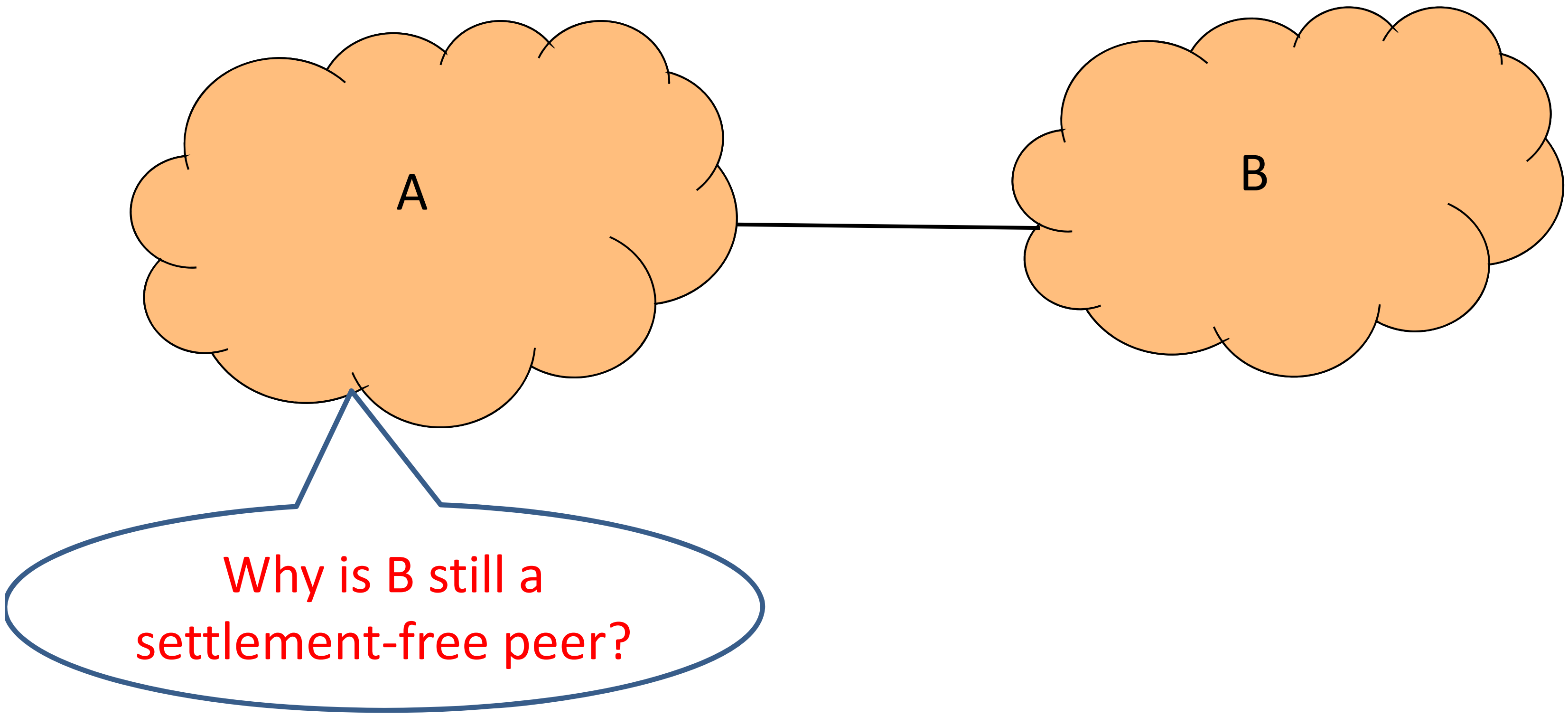
Peering Requirements

- Laundry list of conditions that networks specify as requirements for (settlement-free) peering
  - Traffic ratios, minimum traffic, backbone capacity, geographical spread …
- Heuristics to find networks for which it makes sense to exchange traffic for “free”
  - But when it comes to paid peering..
  - What is the right price? Who should pay whom?
- Are these heuristics always applicable?
  - Mutually beneficial peering links may not be formed

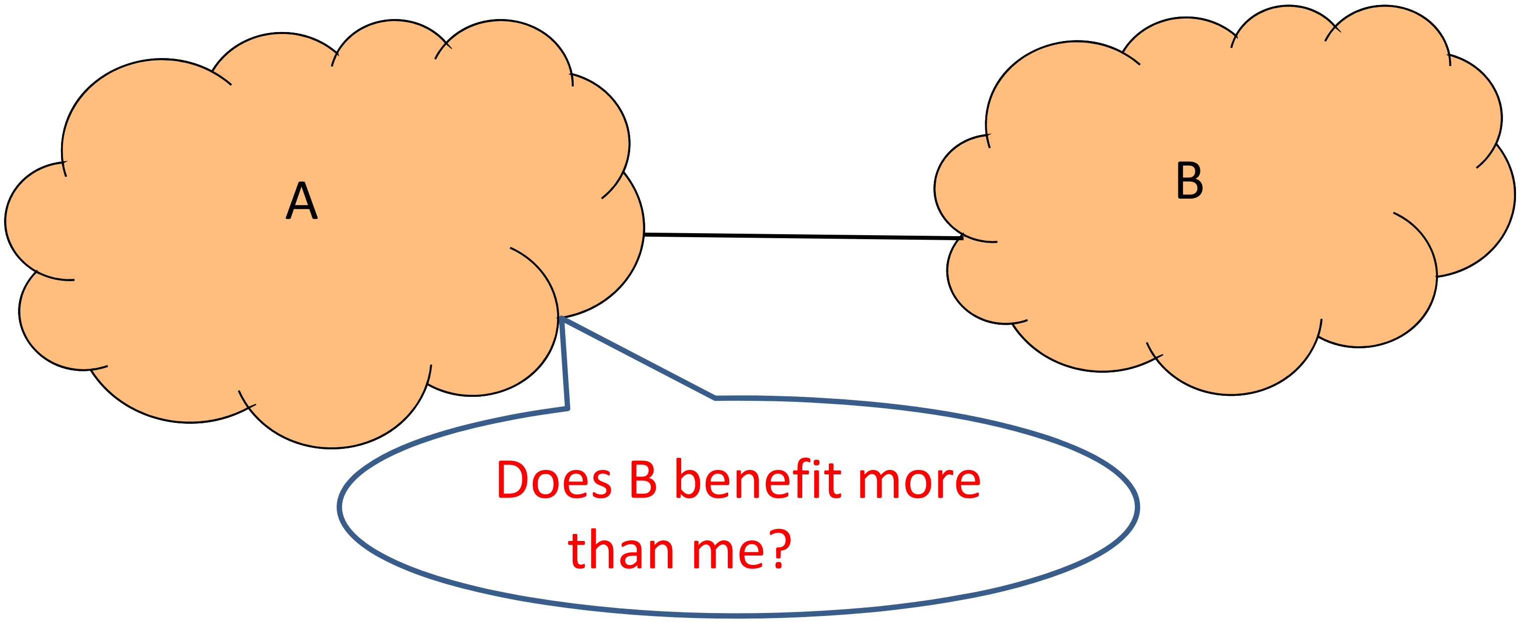
Peering Uncertainty – Current Peers



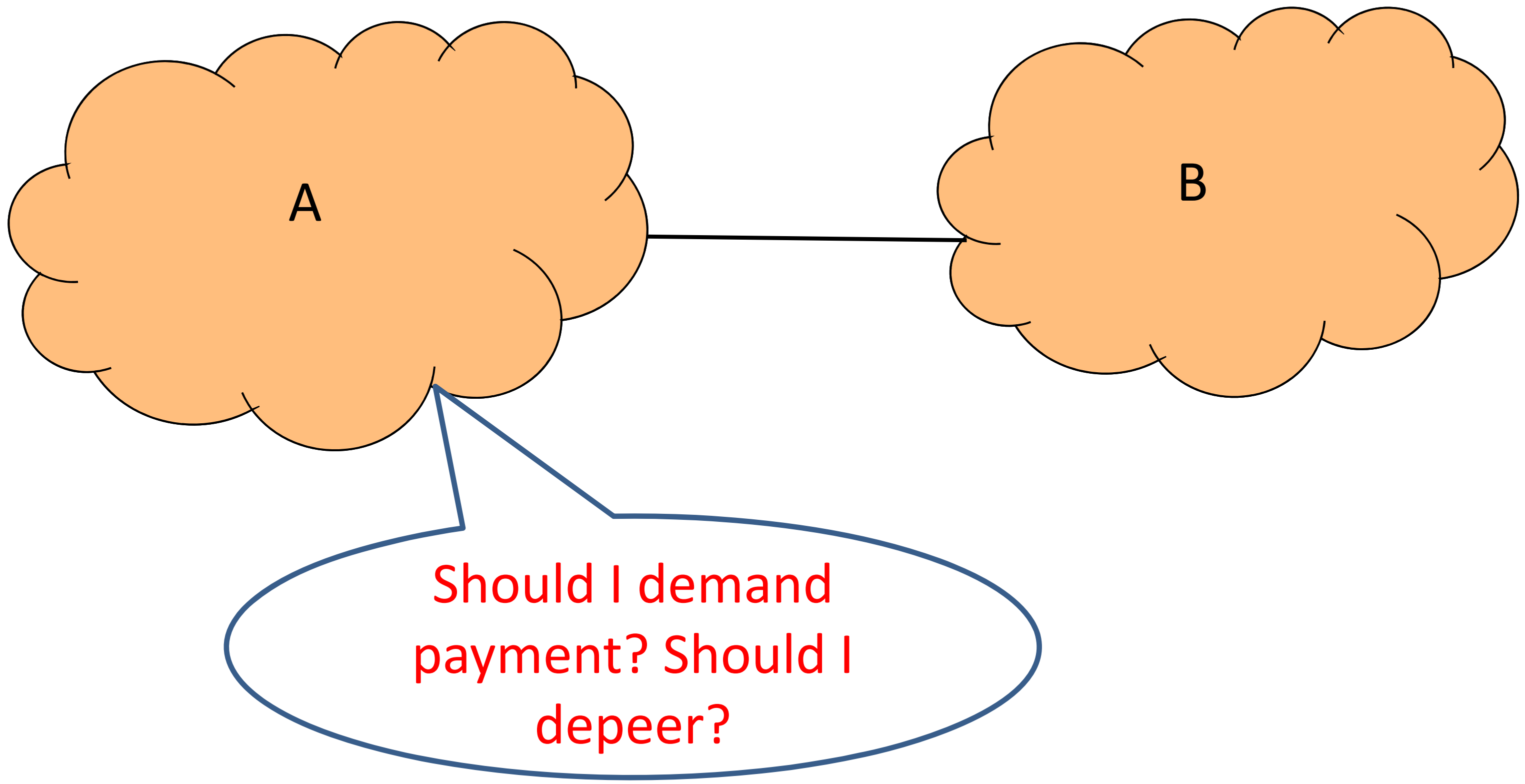
Peering Uncertainty – Current Peers



Peering Uncertainty – Current Peers

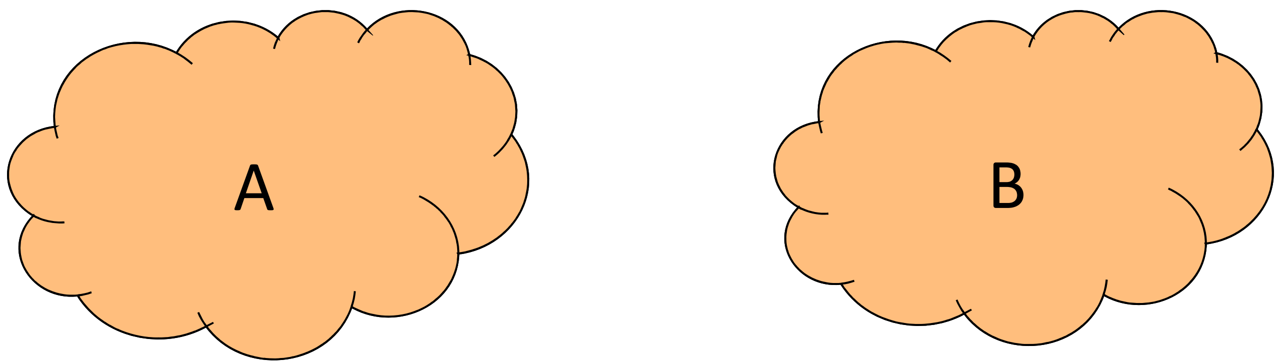


Peering Uncertainty – Current Peers

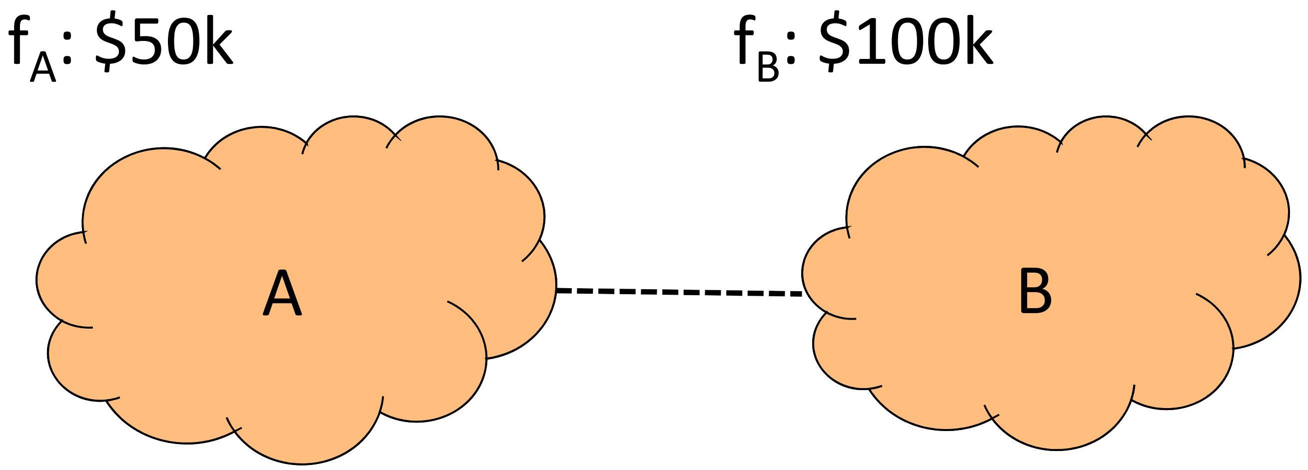


Negative Peering Value

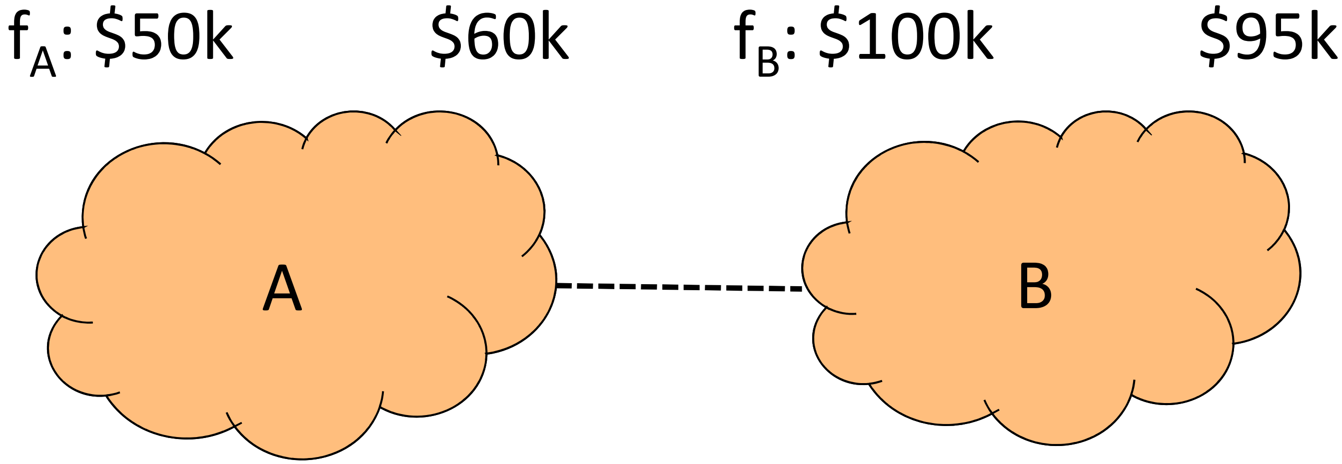
$f_A: \$50k$ $f_B: \$100k$



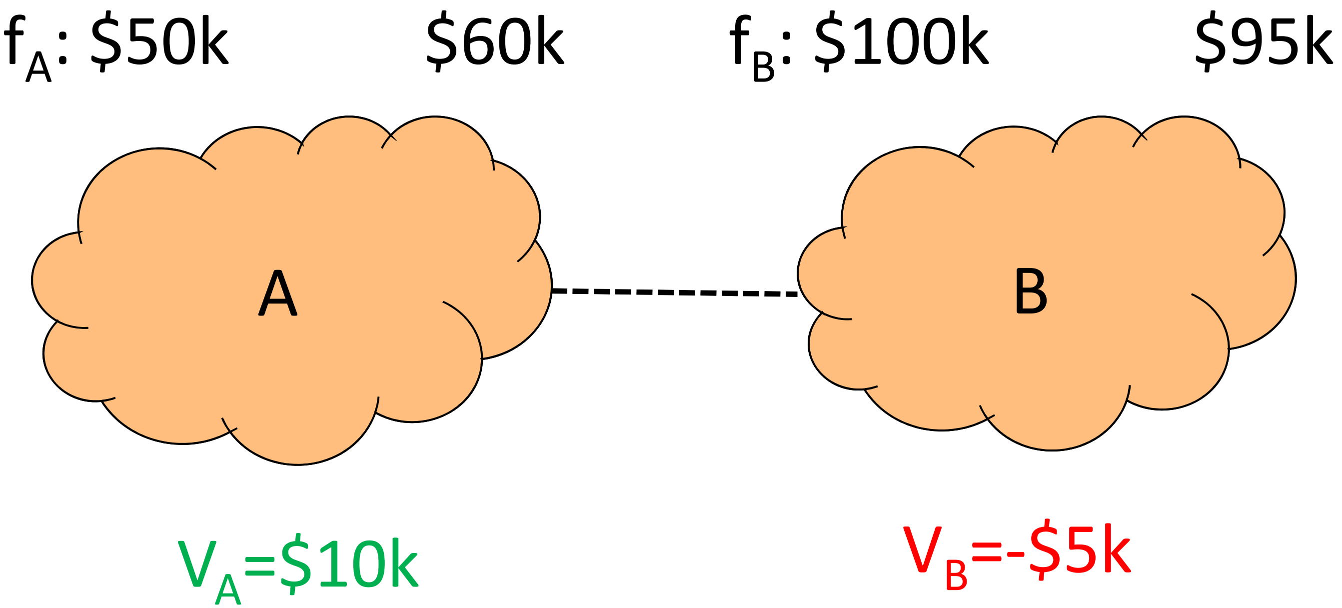
Negative Peering Value



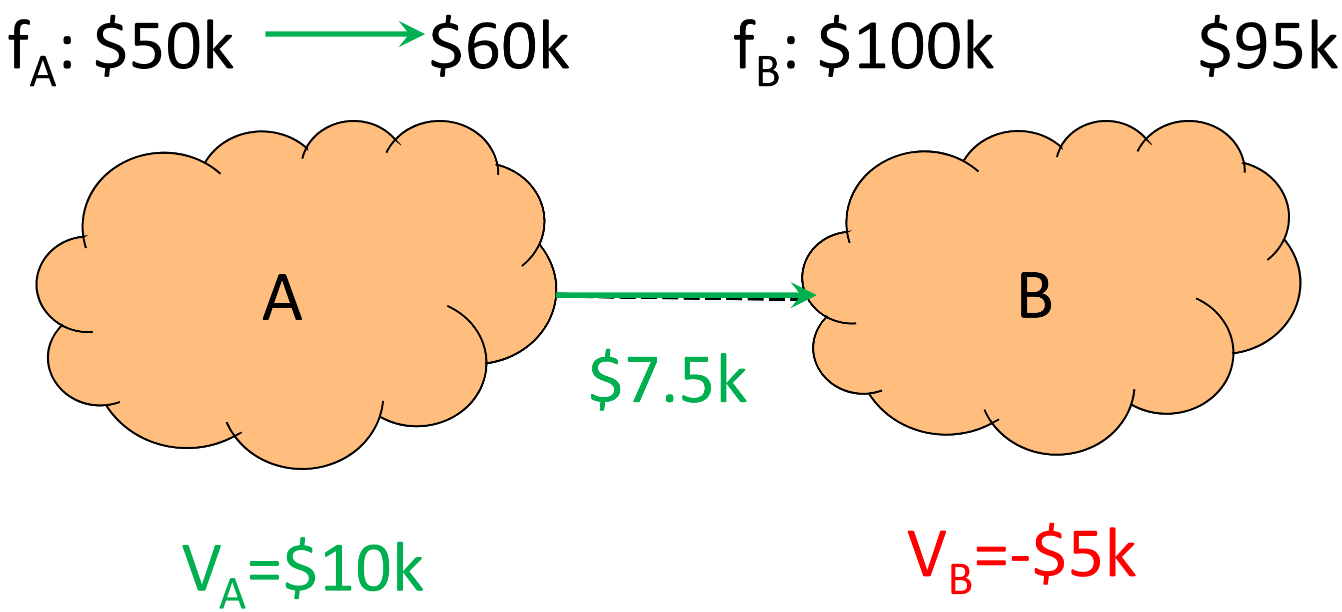
Negative Peering Value



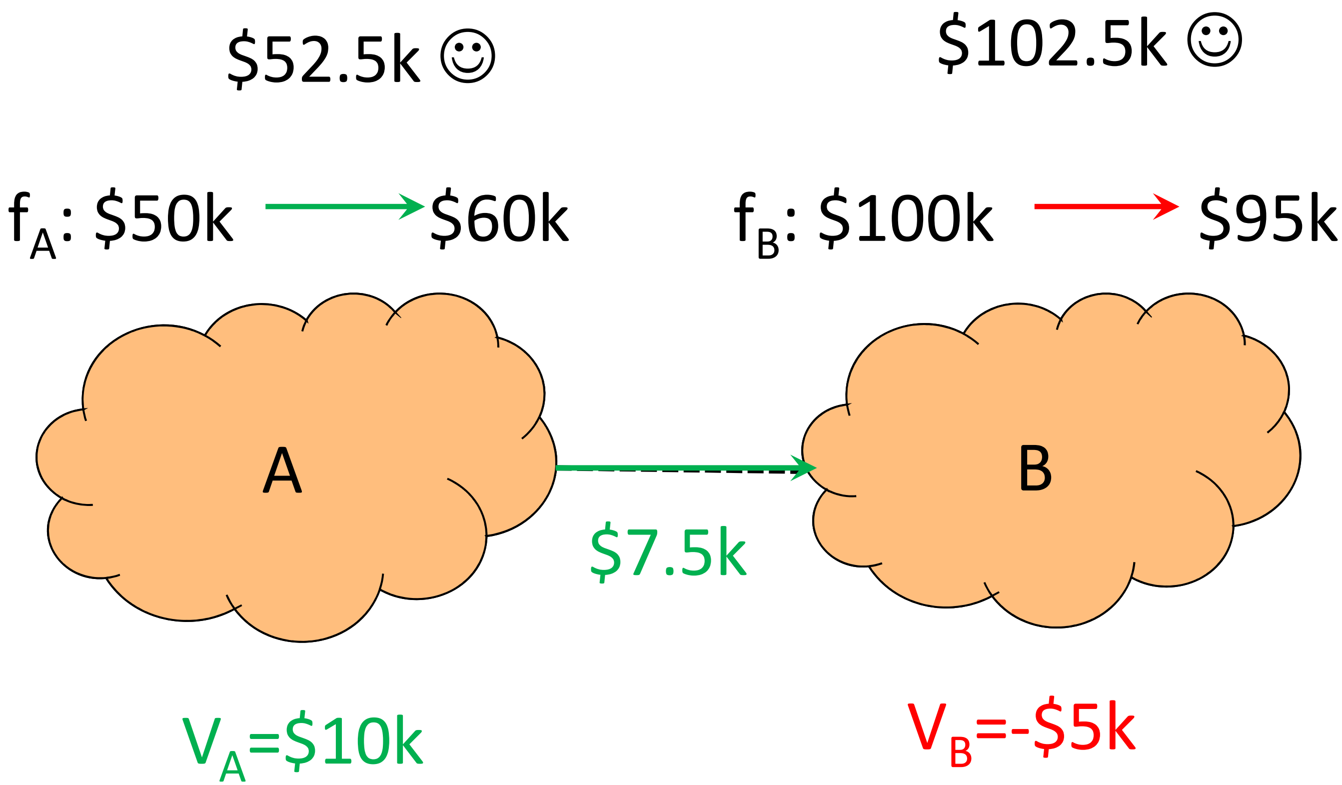
Negative Peering Value



Negative Peering Value



Negative Peering Value



Measuring Peering Value

- How do A and B measure V_A and V_B?
- With Peering trials:
  - Collect: netflow, routing data
  - Know: topology, costs, transit providers
- With peering trials, A and B can measure their own value for the peering link (V_A and V_B) reasonably well
- Hard for A to accurately measure V_B (and vice versa)

Hiding peering value

- Assume true $V_A + V_B > 0$ and $V_B > V_A$
  - A should get paid $(V_B - V_A)/2$

- If A estimates V_B correctly, and claims its peering value is V_L, where $V_L << V_A$
  - B is willing to pay more: $(V_B - V_L)/2$ ☺

- If A doesn't estimate V_B correctly, and $V_L + V_B < 0$, the peering link is not feasible!
  - A loses out on any payment ☹

- Does the risk of losing out on payment create an incentive to disclose the true peering value?

Peering Policies

- What peering policies do networks use? How does this depend on network type?
- Do they peer at IXPs? How many IXPs are they present at?
- PeeringDB: Public database where ISPs volunteer information about business type, traffic volumes, peering policies
- Collecting peeringDB snapshots periodically
- Goal is to study how peering policies evolve

peeringDB

Company Information	
Company Name	AT&T US - AS7132
Also Known As	ATTIS, SBC Internet Services, SBCNT
Company Website	http://www.att.com
Primary ASN	7132
IRR Record	AS-SBCIS-7132
Network Type	Cable/DSL/ISP
Approx Prefixes	2900
Traffic Levels	100+ Gbps
Traffic Ratios	Mostly Inbound
Geographic Scope	Global
Looking Glass URL	route-server.sbcglobal.net
Route Server URL	
Notes	Email peering@attglobal.net for non-US peering requests - AS2686, AS2687 & AS2688. Please include your ASN in the request.
Protocols Supported	Unicast IPv4 ☑ Multicast ☐ IPv6 ☐
Date Last Updated	2008-01-11 18:47:32 UTC
Peering Policy Information	
Peering Policy URL	http://www.att.com/peering/
General Policy	Selective
Multiple Locations	Required - US
Ratio Requirement	Yes
Contract Requirement	Not Required
Contact Information	

Public Peering Exchange Points

Exchange Point Name	ASN	IP Ad
Equinix Ashburn	7132	206.22
Equinix Ashburn	7132	206.22
Equinix Chicago	7132	206.22
Equinix Dallas	7132	206.22
Equinix Los Angeles	7132	206.22
Equinix Newark	7132	206.22
Equinix Palo Alto (PAIX)	7132	198.32
Equinix San Jose	7132	206.22

Private Peering Facilities

Facility Name	ASN	City
Equinix Ashburn (DC1-DC5)	7132	Ashburn
Equinix Chicago (CH1/CH2)	7132	Chicago
Equinix Dallas (DA1)	7132	Dallas
Equinix Los Angeles (LA1)	7132	Los Angeles
Equinix Newark (NY1)	7132	Newark
Equinix Palo Alto (The Ocho)	7132	Palo Alto
Equinix San Jose (SV1)	7132	San Jose

References

- "The Internet is Flat: Modeling the Transition from a Transit Hierarchy to a Peering Mesh"
  - A. Dhamdhere, C. Dovrolis [CoNEXT 2010]
- "A Value-based Framework for Internet Peering Agreements"
  - A. Dhamdhere, C. Dovrolis, P. Francois [ITC 2010]
- "The Economics of Transit and Peering Interconnections in the Internet"
  - C. Dovrolis, K. Claffy, A. Dhamdhere [NSF NETSE 2010-2013]