Innovating in a Networked World

Roch Guérin
(with input and help from many others)

Dept. Elec. & Sys. Eng.
University of Pennsylvania

What This Talk Is and Is Not

- This talk argues that we have entered an era where (network) innovation will succeed not only based on technical merits, but also as a function of complex economic interactions
  - We need a better understanding of the incentives that govern the adoption of Internet technologies

P.S.: This holds whether we contemplate clean-slate solutions or keep evolving the current Internet

- This is **not** a talk meant to argue that technology innovation is not needed anymore in networking

Outline

- Quantifying the Internet "stage"
- An example of the challenges confronted by technology adoption in large-scale networks
  - The IPv6 migration
- Understanding adoption decisions in a networked world
  - Can the user be the network?
  - Shared or separate networks?
  - Complex or simple networks?

The Internet Incumbent

- Internet users
  - ~117 millions in 1997, ~360 millions in 2000, and ~2.2 billions in 2011 (from ~2% to ~33% of the world's population)
- Registered Internet domains
  - ~15,000 in 1992, ~27 millions in 2000, and ~138 millions in March 2012
- Internet Autonomous Systems
  - ~5,000 ASes in 1996, ~10,000 ASes in 2000, and ~60,000 ASes in 2012
- Core Internet routing tables
  - ~5,000 entries in 1992, ~70,000 entries in 2000, and >400,000 entries in 2012
- Global IP traffic growth
  - ~5 Tera(10^{12})Bytes/month in 1992, ~84 Peta(10^{15})Bytes/month in 2000, and ~28 Exa(10^{18})Bytes/month in 2011

So it is big, still growing, and arguably involves lots of complex interactions

Sources:
- http://www.internetworldstats.com/stats.htm
- http://www.dailychanges.com/
- http://en.wikipedia.org/wiki/Internet_traffic
- http://en.wikipedia.org/wiki/Global_Internet_usage
- http://en.wikipedia.org/wiki/Landline
- http://www.zooknic.com/Domains/counts.html
- http://bgp.potaroo.net

Outline

- Quantifying the Internet "stage"
- An example of the challenges confronted by technology adoption in large-scale networks
  - The IPv6 migration
- Understanding adoption decisions in a networked world
  - Can the user be the network?
  - Shared or separate networks?
  - Complex or simple networks?

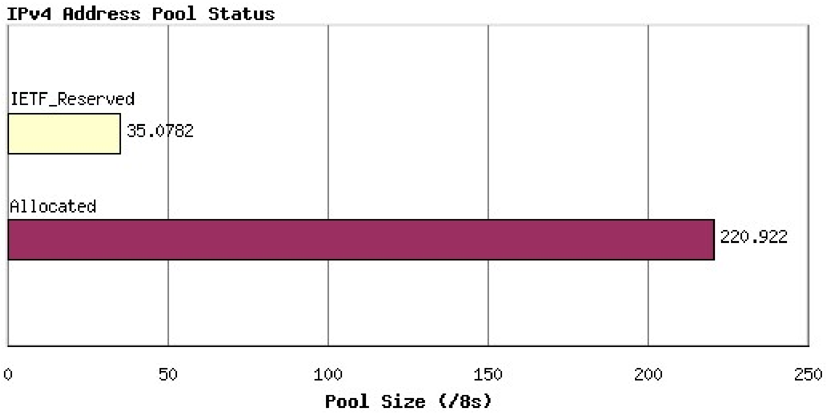
The IPv6 Migration

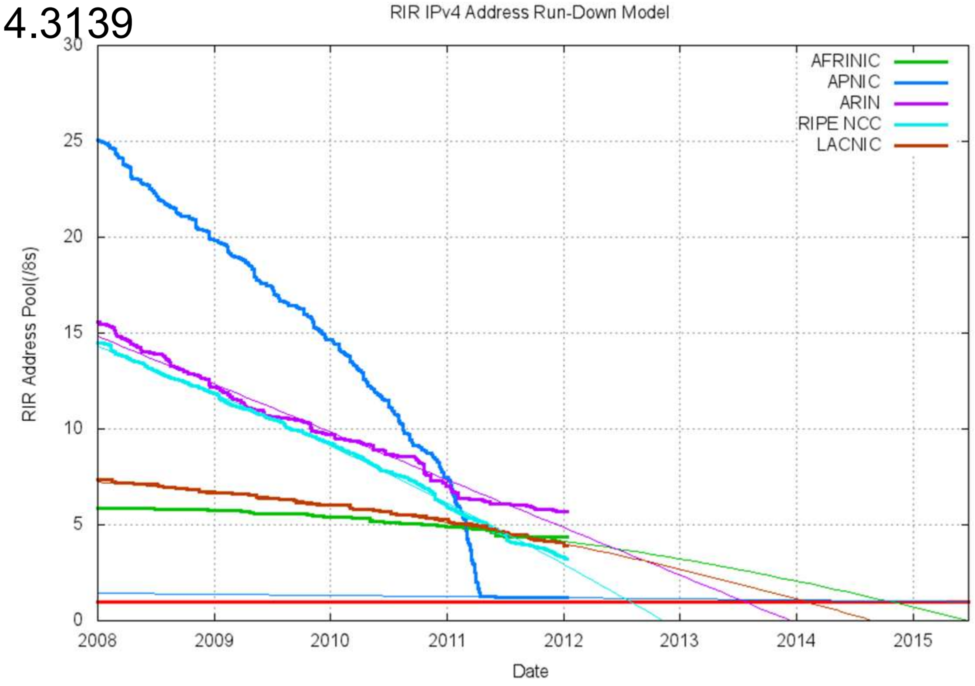
Source: http://www.potaroo.net/tools/ipv4/index.html

IANA pool depleted in February 2011 (No more addresses left!)

RIR	Projected Exhaustion* Date	Remaining /8s in RIR Pool
APNIC:	**19-Apr-2011 (!)**	1.1896
RIPENCC:	**08-Aug-2012**	2.1605
ARIN:	**24-Jun-2013**	4.6591
LACNIC:	**01-Feb-2014**	3.6378
AFRINIC:	**09-Nov-2014**	4.3139

* Reaches last /8





A Well-Known Problem With A Well-Understood Technical Solution

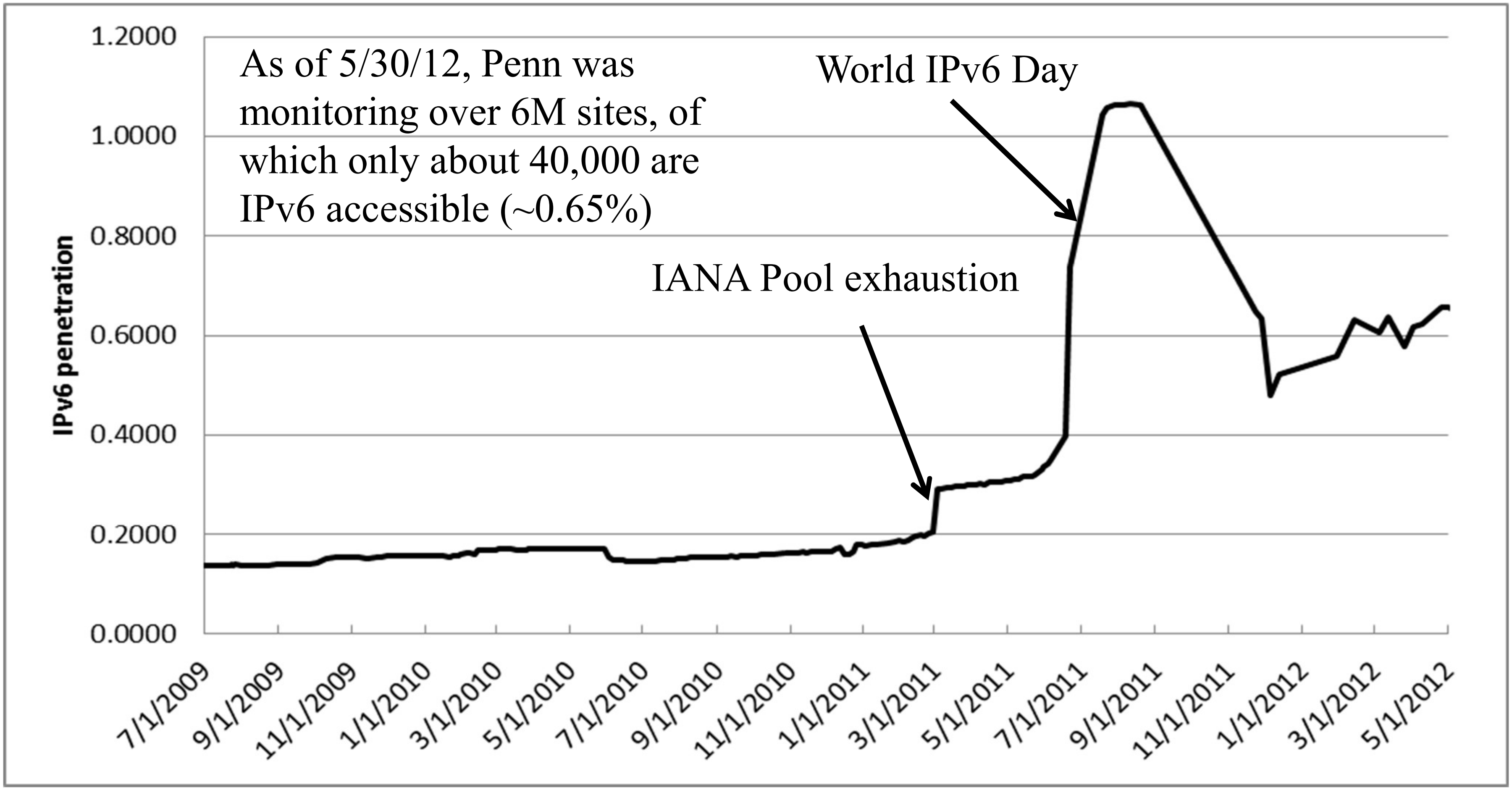
- IPv4 address exhaustion has been repeatedly forecast
  - Maybe too many times
- A solution (IPv6) was standardized in 1995 (RFC 1883)
- Equipment vendors (routers & hosts) were slow in implementing it, but IPv6 has now been systematically available for more or less over 5 years
- So the transition to IPv6 should be relatively straightforward

- Well, not exactly…

Representative IPv6 Status

IPv6 web accessibility from Penn

(http://mnlab-ipv6.seas.upenn.edu)



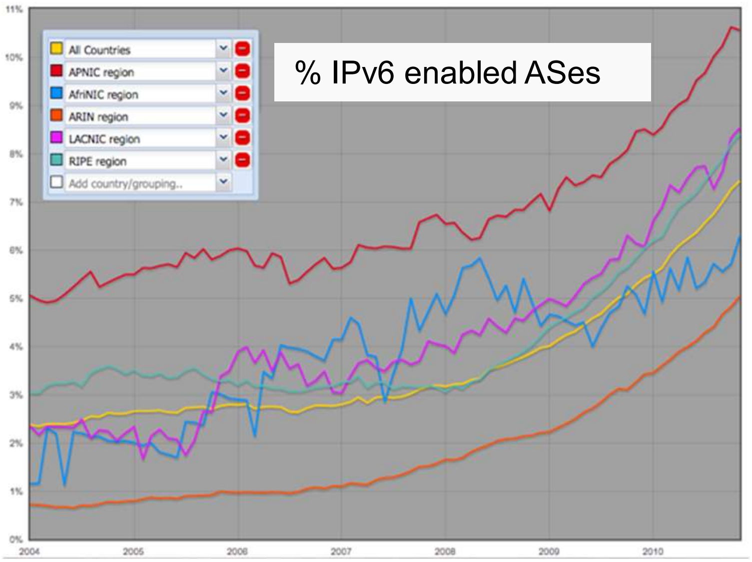
As of 5/30/12, Penn was monitoring over 6M sites, of which only about 40,000 are IPv6 accessible (~0.65%)

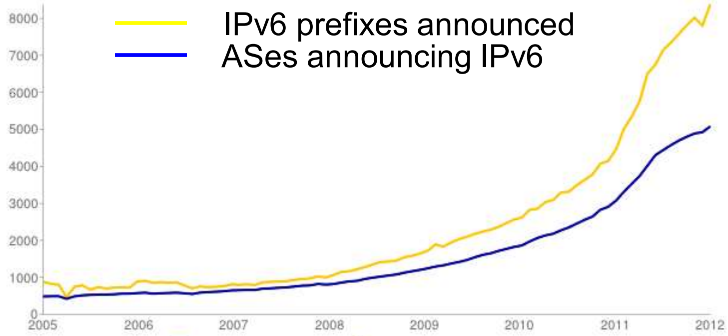
World IPv6 Day

IANA Pool exhaustion

IPv6 Adoption By the Major Internet Stakeholders

- Internet Service Providers are among those who stand to gain the most from IPv6 adoption
    - They need more (IPv6) addresses to sign-up new customers
    - But they have not really been jumping on the IPv6 bandwagon
- Maybe the technology is not ready after all, *i.e.,* this is still a technology problem





† From http://www.ipv6actnow.org/statistics/

Is IPv6 Technologically Ready?

- Q: Is poorly performing IPv6 technology justifying ISPs limited enthusiasm for IPv6?
- Approach
  - Assessing the extent to which IPv6 performs “as expected” when the rest of the Internet does
  - Focus on web access as a representative Internet application
    - Compare IPv6 and IPv4 web access performance from different locations and to many different web sites
    - Quantify differences and explore possible causes

Measurement Vantage Points

Vantage Points		Date on-line	AS_PATH	Type
Comcast	(B)	2/4/11	Y	Commercial
Loughborough U.	(D)	4/29/11	Y	Academic
Penn	(A)	7/22/09	Y	Academic
UPC Broadband	(C)	2/28/11	Y	Commercial
Go6-Slovenia	(E)	5/19/11	N	Commercial
Tsinghua U.	(F)	3/22/11	N	Academic

Measurement Data Overview

- From each vantage-point
  - Target top 1M web sites (from Alexa) and a few others
  - Record download speeds for all web sites accessible over both IPv6 and IPv4
  - Gather monitoring data over several months
  - Compare IPv6 and IPv4 AS_PATHs
- Minor monitoring differences across vantage points
  - Different start dates
  - Asynchronous sampling of Alexa (Alexa churn)
  - Local site additions
- Monitoring statistics
  - Confidence targets for individual monitoring rounds
  - Confidence targets for site performance across monitoring rounds (average out temporal variations)
  - Sites that fail to meet confidence targets are eliminated

Many sites map to the same IP address, *e.g.,* hosting service

Vantage Points	# unique IPs
Comcast	844,355
Loughborough U.	883,413
Penn	1,633,606
UPC Broadband	946,977
Go6-Slovenia	850,954
Tsinghua U.	917,582

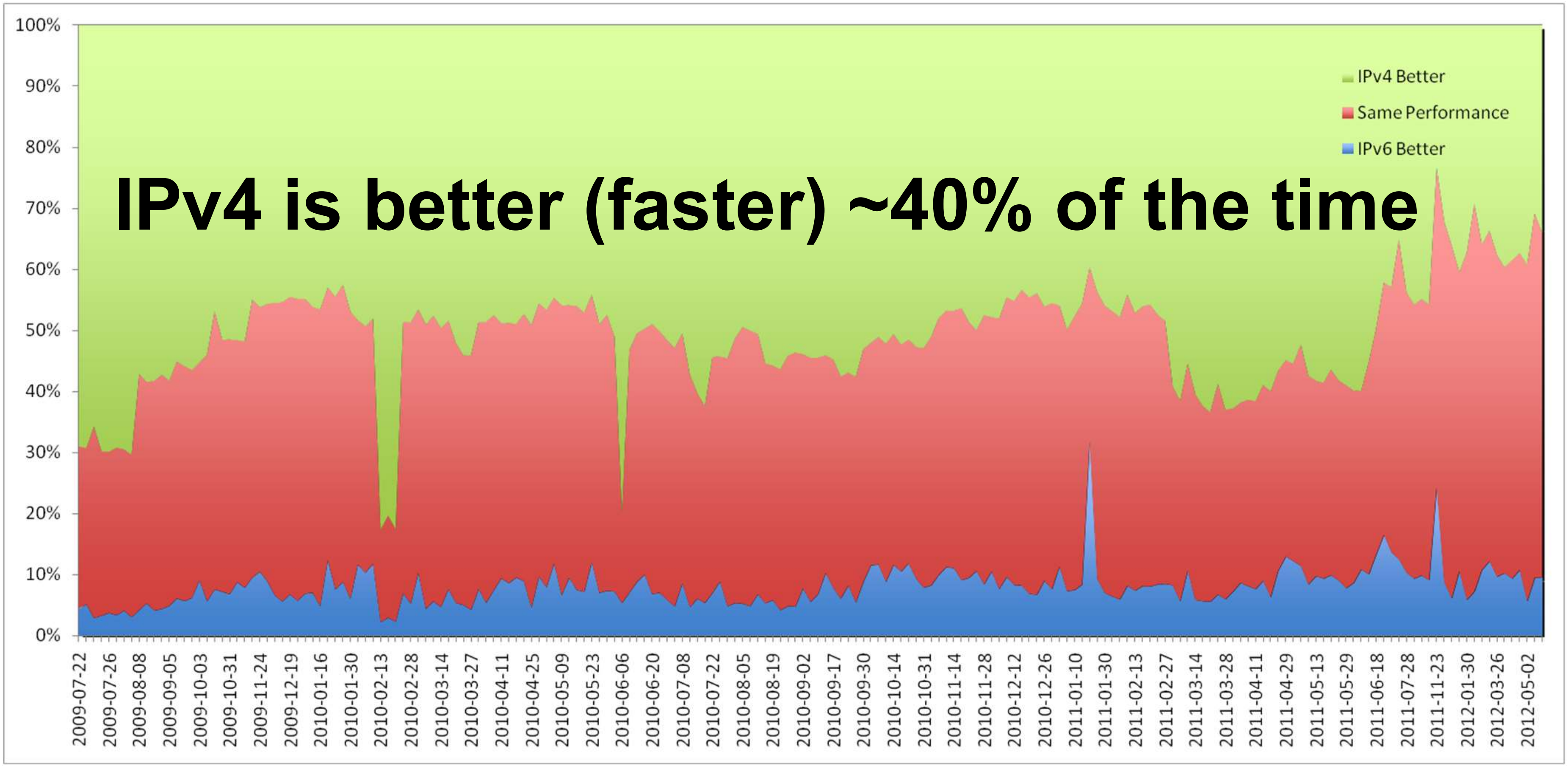
Measurement Data Scope

Meet confidence target

# IPv6+IPv4	Comcast	LU	Penn	UPCB	All
Sites (total)	4,568	5,069	12,385	7,843	-
Sites (kept)	3,525	3,906	7,994	4,418	-
Dest. ASes (IPv4)	724	801	1,047	766	1,364
Dest. ASes (IPv6)	592	642	727	609	1,010
ASes crossed (IPv4)	922	1,019	1,332	988	1,785
ASes crossed (IPv6)	742	764	849	746	1,208

P.S.: Removing sites that did not meet confidence targets did not introduce noticeable bias

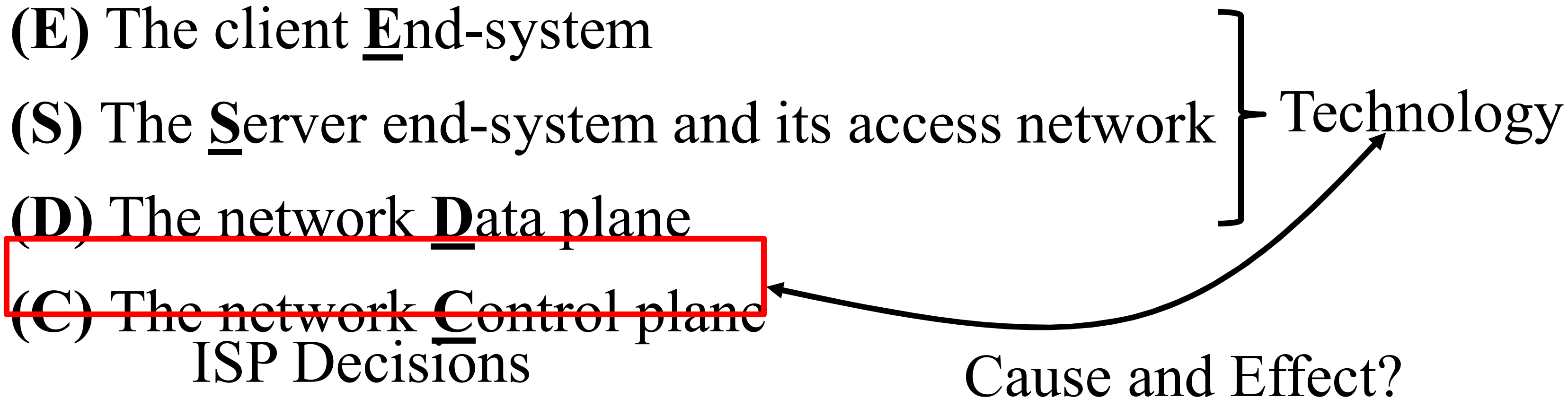
A Bird's Eye View of the Findings



IPv4 is better (faster) ~40% of the time

Is IPv6 Performance Lag Due To Technology?

- Four major factors can affect how IPv6 performs compared to IPv4



(E) The client **E**nd-system

(S) The **S**erver end-system and its access network

(D) The network **D**ata plane

(C) The network **C**ontrol plane

Methodology Overview

- Classify web sites based on whether or not their IPv6 and IPv4 ***"locations"*** and ***"paths"*** differ
  - Same (different) location, *i.e.,* SL (DL) ≡ Same (different) destination AS
    - For SL sites (DL sites have obviously different AS_PATHs)
      Same (different) path, *i.e.,* SP (DP) ≡ Same (different) AS_PATH

- How do IPv6 and IPv4 compare within SP?
  - For SP sites, **(C)** is absent
- Are the results different when we consider DP sites?
  - Differences are likely caused by **(C)**

IPv6 vs. IPv4 Performance Within SP

When paths are identical, IPv6 and IPv4 perform similarly

	Comcast	LU	Penn	UPCB
IPv6 ≈ IPv4*	80.7%	70.2%	81.3%	79.8%
Zero mode	6%	10.8%	9.4%	7.3%
Small # sites	13.3%	19%	9.3%	12.9%
# ASes	233	248	75	124
†Cross-check ☑	129	164	47	82
†Cross-check ☒	0	0	0	0

* IPv6 ≈ IPv4: IPv6 performance is within 10% confidence interval of IPv4 performance, or IPv6 outperforms IPv4

† Cross-checking looks for (in)consistent results for ASes found in the same "category," *i.e.,* SP or DP, from different vantage points

World IPv6 Day (6/08/11) Validation (Sites in SP)

World IPv6 Day IPv6 traffic was significantly higher, *i.e.*, data plane performance was tested more extensively

	LU	Penn	UPCB
IPv6 ≈ IPv4	85.7%	92.3%	72.2%
Other	14.3%	7.7%	27.8%
#ASes	42	13	36
Cross-check ☑	17	8	13

IPv6 vs. IPv4 Performance Within DP

A very different result, when IPv6 and IPv4 follow different paths!

	Comcast	LU	Penn	UPCB
IPv6 ≈ IPv4	11%	10%	3%	8%
Zero mode	5%	3%	12%	6%
# ASes	233	248	75	124

- World IPv6 Day Results

	LU	Penn	UPCB
IPv6 ≈ IPv4 (DP)	48.9%	53.5%	51.0%
#ASes	92	114	102
IPv6 ≈ IPv4 (SP)	85.7%	92.3%	72.2%

Recall SP figures →

What Can We Conclude?

- By-and-large, when ISPs do their part, *i.e.*, routing is the same, IPv6 (web access) performance is on par with that of IPv4
- Concerns for the maturity of IPv6 (network) technology, therefore, cannot really explain ISPs limited enthusiasm for turning IPv6 on
- In short, it is not a technology problem

Why is Migration to IPv6 Hard?

- Many possible reasons, but complex interactions between incentives from different stakeholders play an important role
- Internet stakeholders
  - Internet Content Providers (ICPs)
    - They derive revenue from users, which depends partly on connectivity quality
    - Converting to IPv6 has a cost (direct or indirect)
  - Internet Service Providers (ISPs)
    - Revenue comes from connecting users (and content providers)
    - Costs for operating the network and deploying translation gateways (IPv4-IPv4 or IPv6-IPv4)
  - Users
    - Connect primarily to access content and services
    - They are sensitive to connectivity cost and quality
- Sample interactions
  - ISP gives IPv6 address to new users – they cannot access the bulk of the ICPs
    - ISP provides translation gateways (IPv6-IPv4)
      - Their cost grows with the volume of translation traffic
    - Keeping translation traffic low requires that ICPs adopt IPv6 as more new users join
      - If gateway quality is low, ICPs have incentives to adopt IPv6, but users are (initially) unhappy, *i.e.,* fewer users
      - If gateway quality is high, users are happy, but ICPs have no incentives to adopt IPv6
  - ISP gives private IPv4 address to new users
    - ISP provides translation gateways (IPv4-IPv4)
    - ICPs have obviously no incentives to adopt IPv6
    - Volume of translation traffic keeps growing

The Net of It

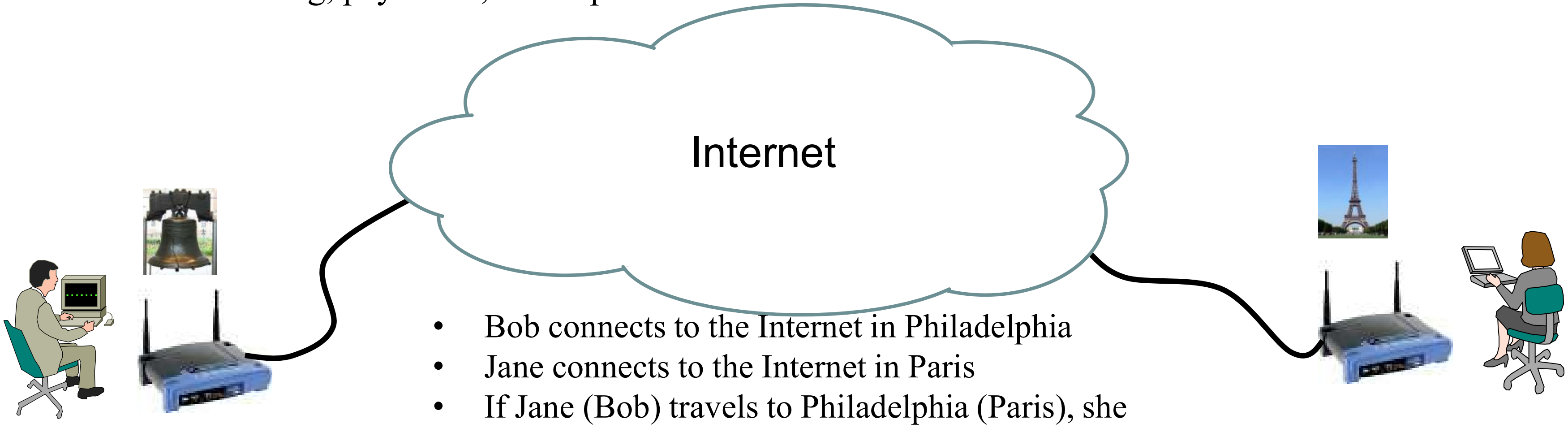
- Large-scale, complex network systems give rise to a wide-range of interactions that can affect technology adoption
- Understanding those effects and how they impact the deployment of new technologies is as important as the technology itself
  - I’ll try to give a few “constructive” examples next

Outline

- Quantifying the Internet "stage"
- An example of the challenges confronted by technology adoption in large-scale networks
  - The IPv6 migration
- Understanding adoption decisions in a networked world
  - Can the user be the network?
  - Shared or separate networks?
  - Complex or simple networks?

User-Provided Connectivity (UPC)

- Users allow others to access their own connectivity in exchange for some compensation
  - Community-based networks, FON, Keywifi
- Different compensation schemes
  - Cost sharing, payments, or reciprocation



- Bob connects to the Internet in Philadelphia
- Jane connects to the Internet in Paris
- If Jane (Bob) travels to Philadelphia (Paris), she (he) can use Bob's (Jane's) Internet connection

- Service exhibits strong externalities that can affect its eventual success
  - Positive externalities: More users means more options to connect while on the road
  - Negative externalities: More users means higher likelihood to have to share connectivity

Challenges in Modeling UPC Adoption

- The standard approach is to model user adoption within a large heterogeneous population, based on the notion of *utility, i.e.,* users adopt if their utility is non-negative
- In the context of UPC, a user's utility should capture
  - Value of basic (home) connectivity
  - Service price
  - Value of connectivity while "roaming"
  - Impact of roaming traffic on connectivity
  - Incentives for accommodating roaming traffic

Externalities

the challenge is in incorporating the impact of externalities

A Standard Utility Based Model

- Utility function of user with roaming parameter θ

$$U(\theta) = F(\theta,x)+G(m) - p_\theta$$

  - A user adopts if $U(\theta) \geq 0$
  - θ is a random variable that identifies a user's roaming profile
    - Known distribution
    - $\theta \in [0,1]$, $\theta = 0$ (never roams), $\theta = 1$ (always roaming)
  - $F(.,.)$ is utility of connectivity (at home and while roaming)
    - x is current level of adoption (coverage)
  - $G(.)$ accounts for negative impact of roaming traffic, and positive impact of possible compensation
    - m is current volume of roaming traffic (depends on number **and** identity –their θ values– of adopters)
  - p_θ is price charged to user with roaming value θ

A Simple Instantiation

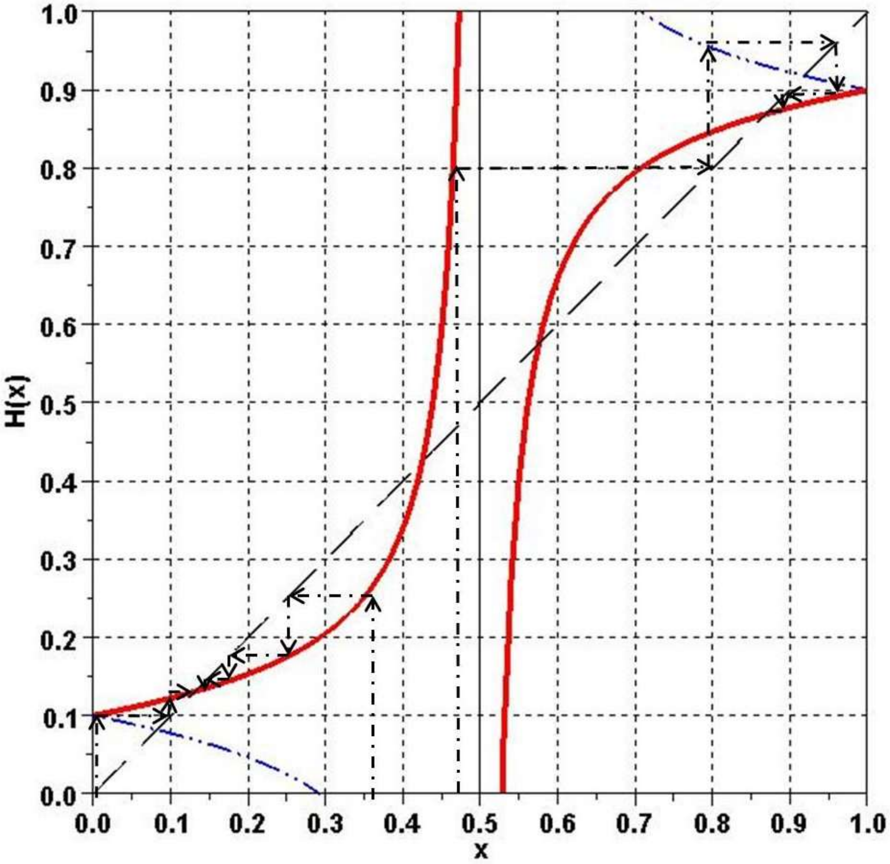
- Linear (positive and negative) externalities

$$U(\theta) = (1 - \theta)\gamma + \rho\theta x + (b - c)m - p$$

- θ : Uniformly distributed in [0,1]
- γ : Value of home connectivity – weighed by a user's roaming frequency
- ρ : Value of connectivity while roaming – weighed by coverage x (odds of finding service), and user's roaming characteristic θ (frequency of use)
- b : Compensation for providing access to roaming traffic (proportional to volume of roaming traffic m)
- c : Impact of volume of roaming traffic, m, on user connectivity
- m : Roaming traffic (uniformly distributed across users' home connections)
- p : Service price (identical for all users)

Sample Adoption Evolution

- Adoption evolves iteratively
  - Adoption at n+1 is a function of *state* at *n, i.e.,* $X_n=(x_n,y)$: number and type adopters, $X_{n+1}=H(X_n)$
- Function(s) $H(X)$ determine adoption evolution
- Equilibria at fixed points, $H(X)=X$, or boundary points

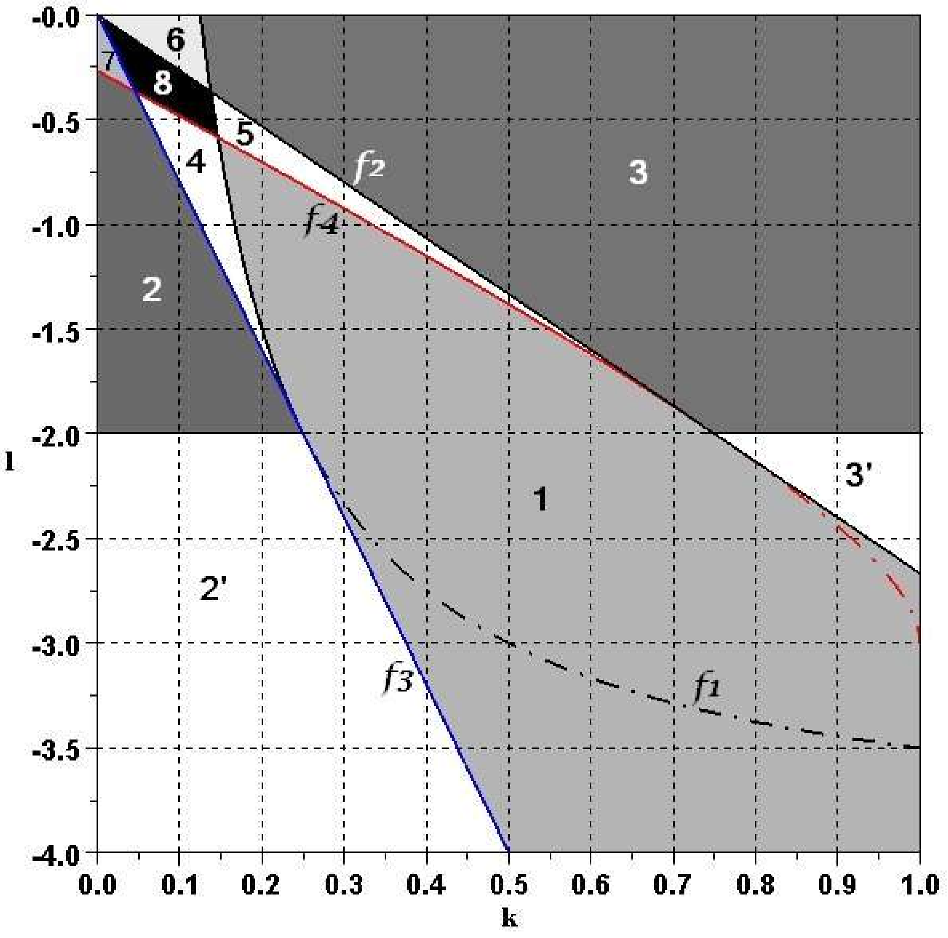


Insight Into Adoption Outcomes

- Characterize regions of the (k,l) plane (exogenous parameters)
  - $k = (\gamma - p)/\gamma$ and $l = (b - c)/\gamma$

  giving rise to different outcomes

Cases	$[0, 1/2)$	$[1/2, 1]$
1	—	—
2	●	—
2’	↻	—
3	—	●
3’	—	↻
4	●, ○	—
5	—	●, ○
6	●, ○	●
7	●	●, ○
8	●, ○	●, ○



A Representative Insight

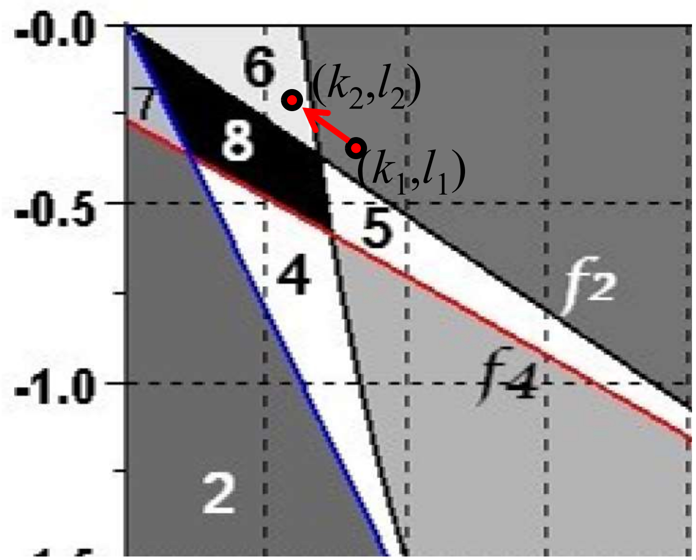
- Recall a user's utility function
  - $U(\theta) = k + lm + \theta(2x - 1)$

    where $k = \gamma - p$ and $l = b - c$, with b corresponding to incentives to offset the impact of roaming traffic
  - At equilibrium bm is "equivalent" to a price decrease, *i.e.*, $p' = p - bm$

- Trading price for incentives can impact adoption ***dynamics***
  - In can result in the emergence of a ***second*** low adoption equilibrium

$(k_1,l_1) = (\gamma - p_1, -c)$, *i.e.*, $b=0$

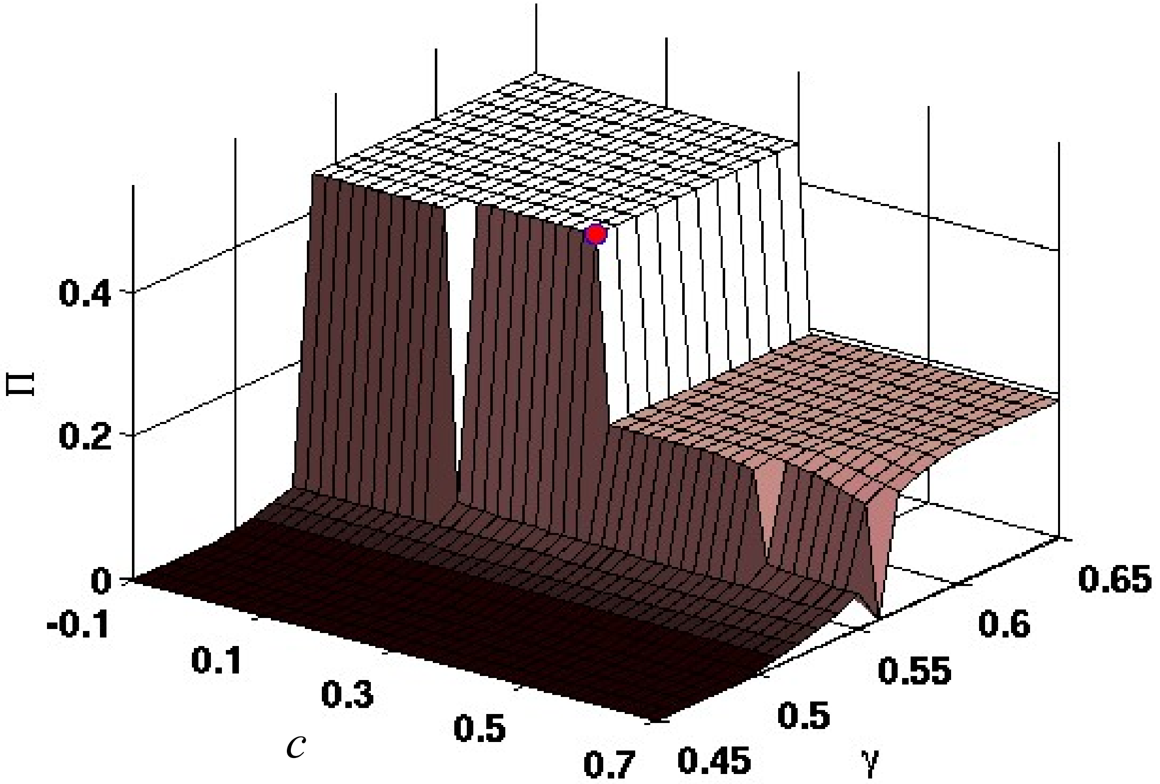
$(k_2,l_2) = (\gamma - (p_1 + \boldsymbol{bm}), \boldsymbol{b} - c)$



- In general, positive and negative externalities make predicting outcomes difficult

Another Representative Insight

- Price optimization for the service is ***extremely*** fragile to estimation errors in service valuation parameters



Outline

- Quantifying the Internet "stage"
- An example of the challenges confronted by technology adoption in large-scale networks
  - The IPv6 migration
- Understanding adoption decisions in a networked world
  - Can the user be the network?
  - Shared or separate networks?
  - Complex or simple networks?

Shared or Separate Networks

- The Internet is arguably a successful example of a *shared* network, *i.e.,* multiple services on the same infrastructure
- When is a shared network (infrastructure, *e.g.,* cloud) as opposed to separate networks preferable?
  - There are both economies and diseconomies of scope (and scale) involved in sharing a common infrastructure
  - We are currently seeing different answers to that question in various contexts, *e.g.,* Information Technology & Operation Technology convergence, triple and quadruple play offerings, etc.
- Can we develop insight into factors that affect the answer?

A Base Formulation

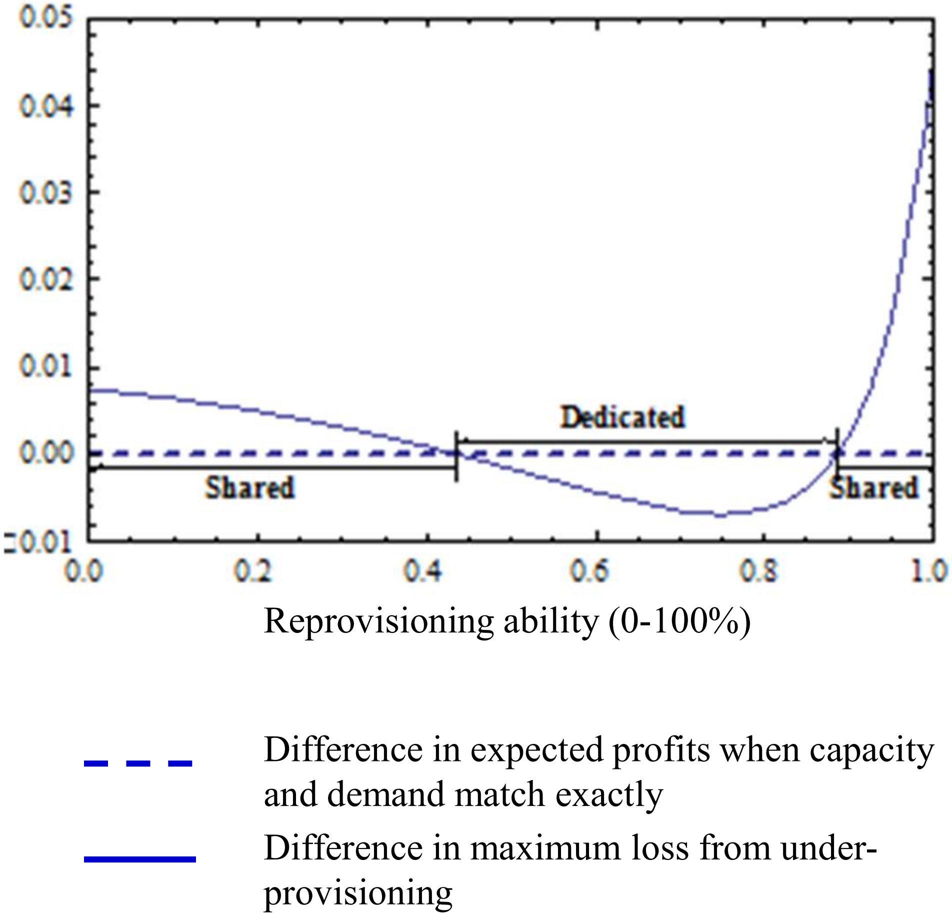
- An existing service (predictable demand) and network
- A new service with uncertain demand
  - Positive demand externalities when services are integrated on the same network
  - Some ability to "reprovision" network capacity in the presence of excess demand (penalty for under-provisioning)
- Economies and diseconomies of scope for both integrated and separate network choices
- Network provider's goal
  - Consider two network options
    1. Deploy a separate network dedicated to the new service
    2. Upgrade existing network to handle both services
  - Select and dimension network option that maximizes profit

A Simple Model

- Capture provider's problem as a three stage sequential process (standard in capacity planning and flexible manufacturing literature)
  - Stage 1: Select network option
  - Stage 2: Provision network capacity for unknown demand
  - Stage 3: Allocate realized demand to provisioned capacity and reprovision if needed
- Solution proceeds in reverse order
  1. Perform reprovisioning given realized demand, provisioned capacity, and network option – this yields a profit figure for each network option
  2. Compute provisioning capacity that maximizes expected profit for each network option
  3. Select network option that yields largest profit

Representative Insight

- Which network option is best depends primarily on two key operational metrics
  - Contribution margin (price less variable costs)
  - Return on capacity (ratio of contribution margin and unit capacity cost)
- More interestingly and less expected, the extent to which the network can be easily reprovisioned to handle excess demand can also affect the outcome



Outline

- Quantifying the Internet "stage"
- An example of the challenges confronted by technology adoption in large-scale networks
  - The IPv6 migration
- Understanding adoption decisions in a networked world
  - Can the user be the network?
  - Shared or separate networks?
  - Complex or simple networks?

The Smart or Dumb Network Question

- The Internet’s success has often been attributed to the fact that it is a “dumb” network, *i.e.,* the narrow waist paradigm
- There is, however, a trade-off between the cost and usefulness of adding functionality to the network
  - Simple networks are cheaper, but require users to individually develop any additional functionality they need
  - Smart network are more expensive, but their features can lower the cost of developing new services

Which option is better, when?

Capturing the Cost-Benefit Trade-Off of Smart vs. Dumb Networks

- A two-sided market model
  - The network as the “platform”
  - Users and content/application developers as the two sides of the market

- The network incurs cost for adding features, and generates revenues from users and content developers
- Content developers pay for network access and incur development costs that depend on available network features. They have revenues that grow in proportion to the number of network users
- Users pay for network access from which they derive a utility that grows with the available content

- The network provider sets access prices and selects what features to offer so as to optimize profit

Representative Insight

- For given network and content developer cost functions, there is an “optimal” number of features
  - The marginal cost increase to the network of adding a new feature is equal to the marginal decrease in development cost across all content developers connected to the network
- The more interesting outcome is that the answer of whether a dumb or a smart network is better, is highly sensitive to the relative rate of change (as the number of network features increases) in the development costs of both network and developers
  - In other words, deciding on the “right” network design is likely to remain a challenging exercise

In Conclusion

- The intent was not to argue that we should focus solely on understanding the complex interactions that arise in networks
  - Without a steady influx of technology innovation the current momentum will eventually stall
- The goal was to highlight that successfully deploying new technologies in networks calls for understanding the many “network effects,” a.k.a. externalities, that influence their adoption
  - The value of a network (technology) often depends on how many others are using, and/or is unlocked only when adoption crosses a certain threshold
  - This is not unique to networks (or the Internet) and affects many large-scale systems
  - More importantly, there is a rich literature and a large set of available tools that can help with such an investigation, *i.e.,* we don’t have to reinvent the wheel

References & Collaborators

[1]. S. Sen, R. Guerin, and K. Hosanagar, *"Shared Versus Separate Networks – The Impact of Reprovisioning."* Proc. ACM ReArch'09 Workshop, Rome, Italy, December 2009.

[2]. R. Guerin and K. Hosanagar, *"Fostering IPv6 Migration Through Network Quality Differentials."* ACM SIGCOMM Computer Communication Review, Vol. 40, No. 3, July 2010.

[3]. S. Sen, Y. Jin, R. Guerin, and K. Hosanagar, *"Modeling the Dynamics of Network Technology Adoption and the Role of Converters,"* IEEE/ACM Transactions on Networking, Vol. 18, No. 6, December 2010.

[4]. M. H. Afrasiabi and R. Guerin, *"Exploring User-Provided Connectivity - A Simple Model."* Proc. ICQT'11 Workshop, Paris, France, October 2011.

[5]. S. Sen, R. Guerin, and K. Hosanagar, *"Functionality-rich versus minimalist platforms: a two-sided market analysis."* ACM SIGCOMM Computer Communication Review, Vol. 41, No. 5, October 2011.

[6]. M. Nikkhah, R. Guerin, Y. Lee, and R. Woundy, *"Assessing IPv6 Through Web Access – A Measurement Study and Its Findings."* Proc. ACM CoNEXT 2011 Conference, Tokyo, Japan, December 2011.

[7]. M. H. Afrasiabi and R. Guerin, *"Pricing Strategies for User-Provided Connectivity Services."* Proc. IEEE INFOCOM 2012 mini-conference, Orlando, FL, March 2012.

Collaborators

- Hadi Afrasiabi, Penn (Ph.D. student)
- Kartik Hosanagar, Wharton, Penn
- Youngmi Jin, KAIST (former post-doc)
- Yiu Lee, Comcast
- Xinxin Li, U. Connecticut
- Mehdi Nikkhah, Penn (Ph.D. student)
- Soumya Sen, post-doc Princeton U. (former Ph.D. student)
- Rich Woundy, Comcast