



Internet Exchange Points, Technologies and IX.PR
Gauss Research Laboratory
Mehmet Akcin

Outline of Presentation

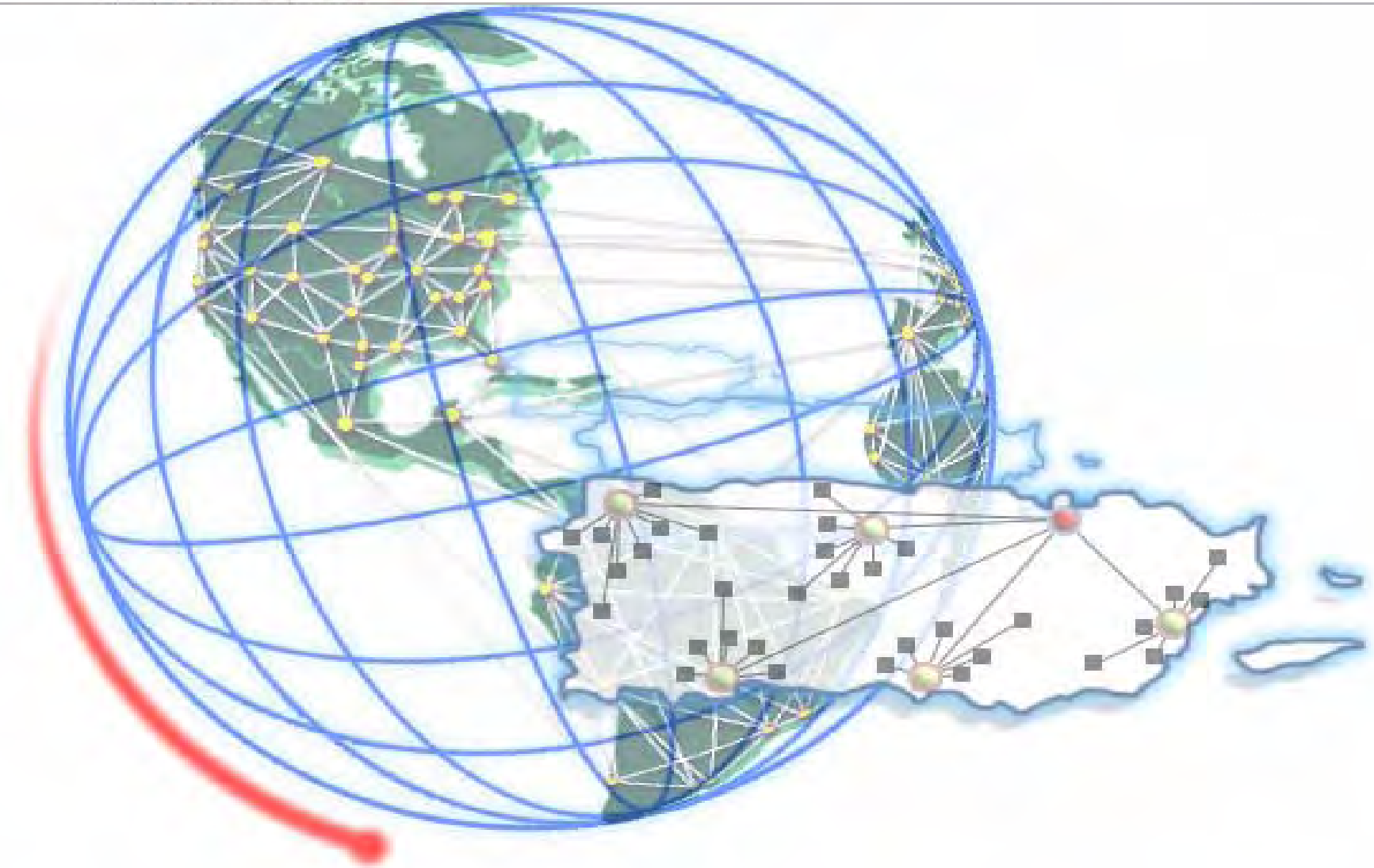
- This workshop will give you detailed information about Internet Exchange points.
- You will understand why an Internet Exchange Point should be established.
- The technologies that are being used in Internet Exchange Points.
- And our project, Internet Exchange of Puerto Rico.

What is an Internet Exchange?

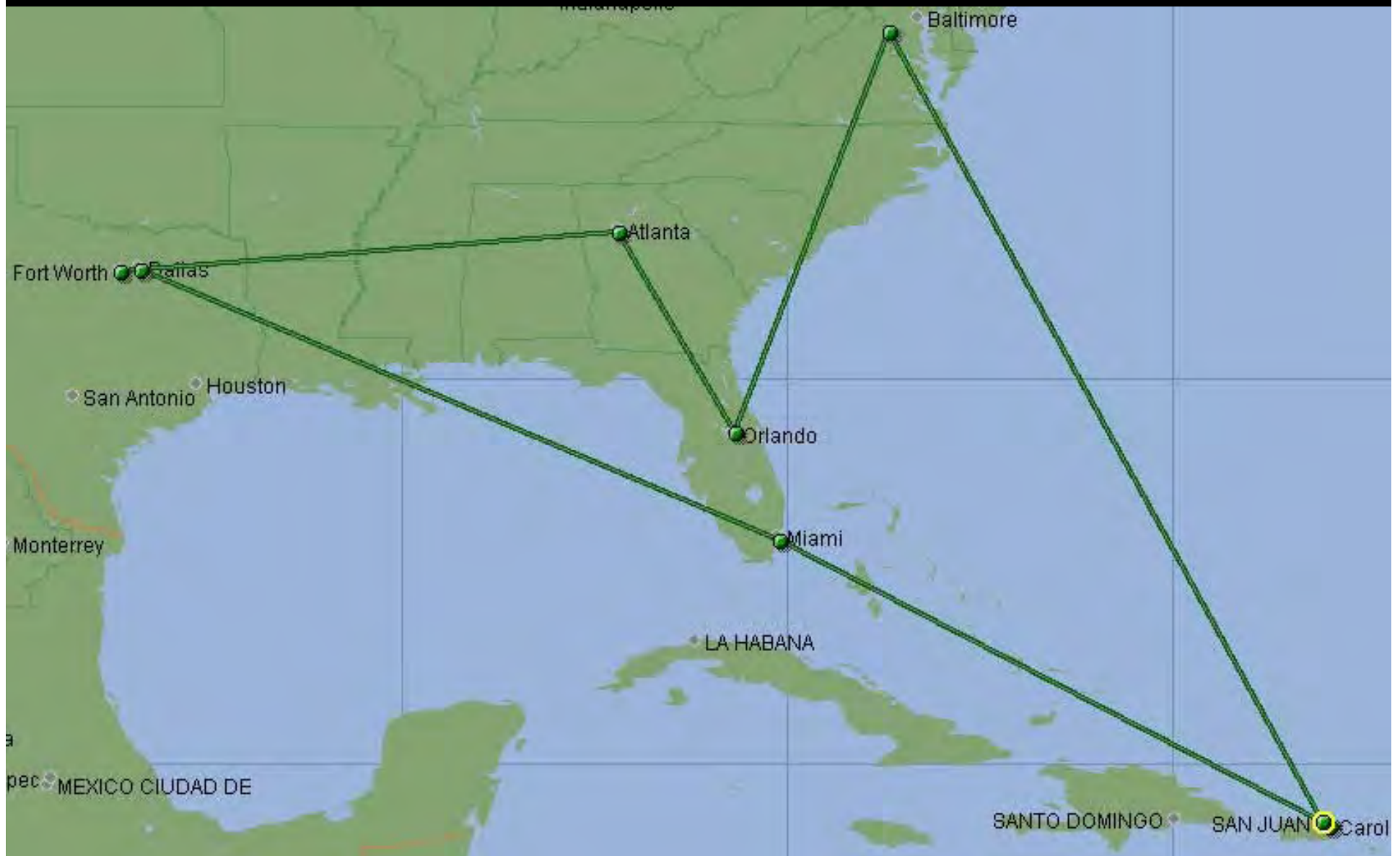
- Multiple ISPs locate backbone IP router nodes in single building operated by co-location provider
- In-building connections
 - to shared interconnect fabric (using Ethernet LAN switching technology)
 - over point-to-point private interconnections
- Routing information, and hence traffic, is exchanged bilaterally between ISPs
- Exchange operator may or may not be same organization as co-location provider
- Co-location provider will generally have other customers:
 - carriers, hosting, content distributors, NS registries/registrars

ix.PR
Internet Exchange
of Puerto Rico

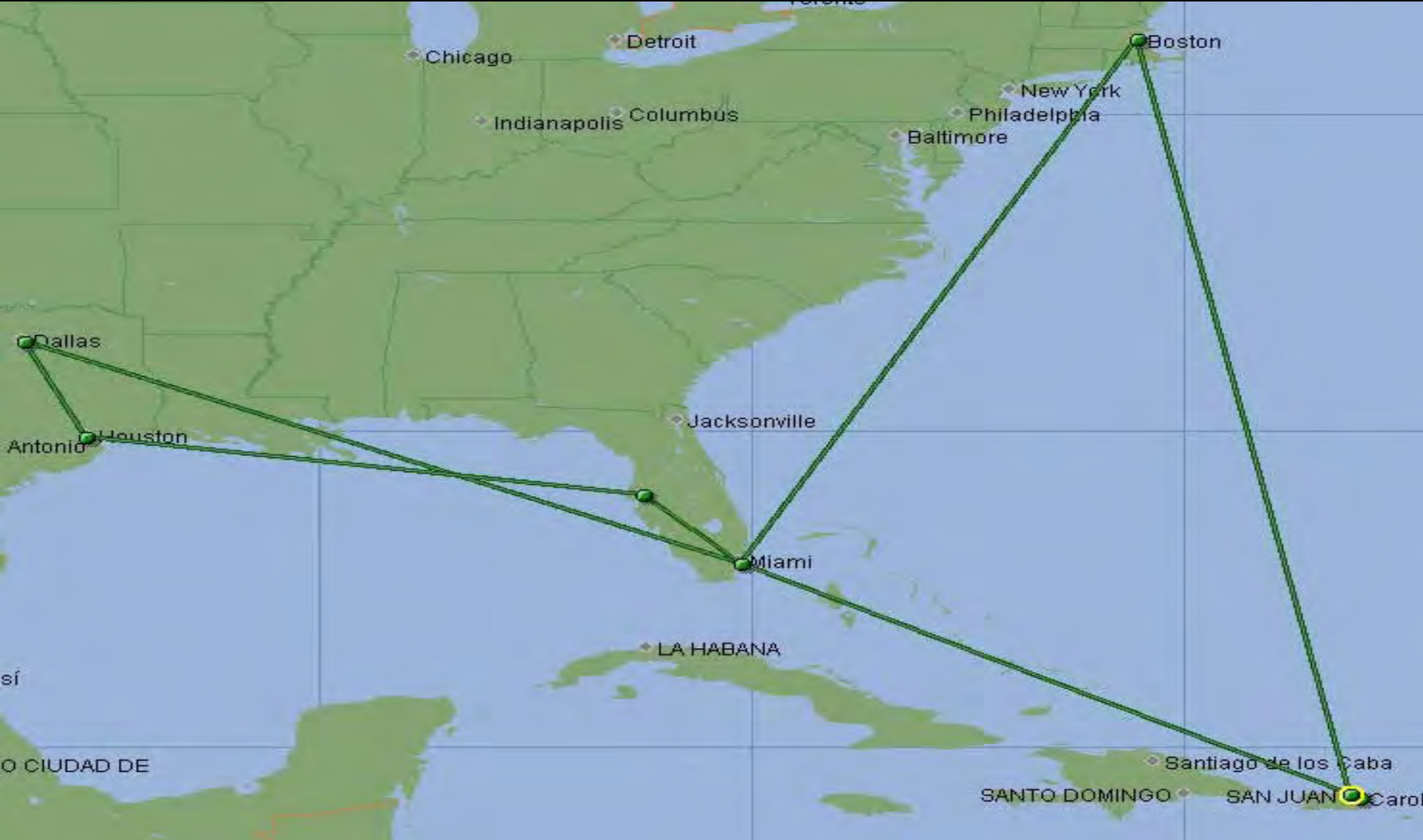
010001010100101001011011010001001
000010101011010001001001001001000



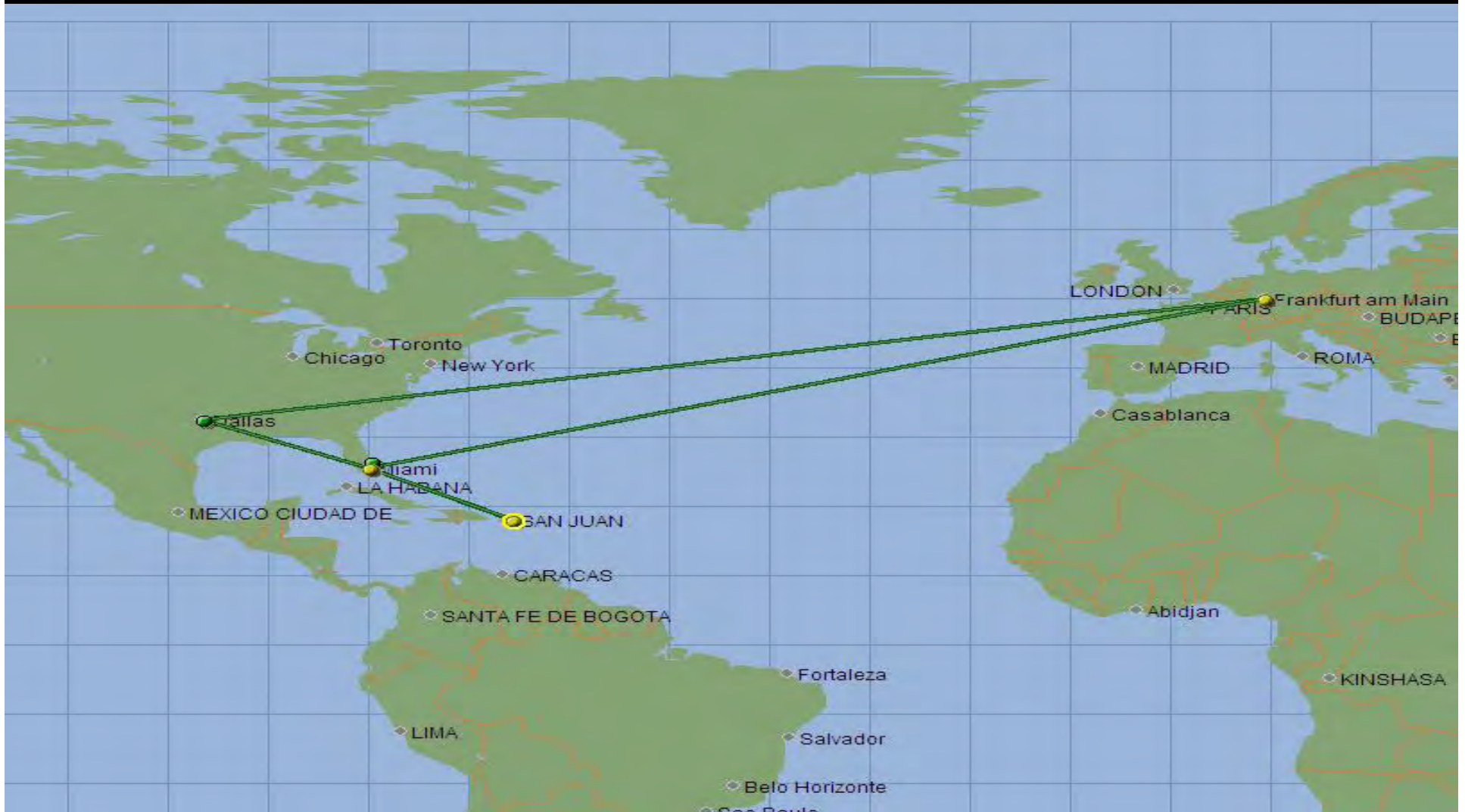
Current Connectivity Examples



Hmm? Not bad.. But we got worst!



Whew! Transatlantic link to exchange data with your neighbor!



Another one..



Internet Exchange Advantages

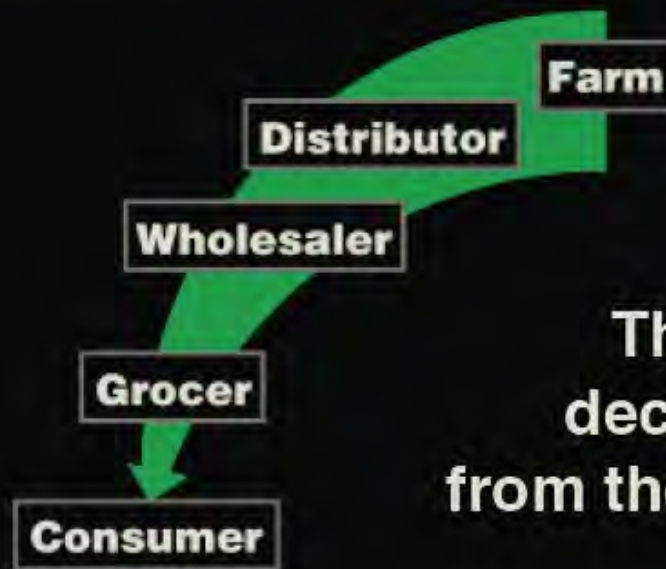
- Keeps domestic traffic within a country/region without having to take indirect international route
- Typically 20-35% (45 to 55% in Puerto Rico) of traffic can be domestic
- Reduced bandwidth costs
- Improved throughput and latency performance
- Economies of scale
- Commercial basis of traffic exchange between ISPs across IXP usually via cost-saving peering
- Critical mass of ISPs in a single location creates competitive market in provision of capacity, transit and services

Why IX is necessary?

- Because Internet is Perishable
 - The quality (as opposed to quantity-per-time) characteristics of an Internet connection are *loss*, *latency*, *jitter*, and *out-of-order delivery*.
 - Loss increases as a function of the number and reliability of components in the path, and the amount of contention for capacity.
 - Latency increases as a function of distance, and degree of utilization of transmission buffers by competing traffic sources.
 - Jitter is the degree of variability in loss and latency, which negatively affects the efficacy and efficiency of the encoding schemes which mitigate their effects. Jitter increases relative to the ratio of traffic burstiness to number of sources.
 - Out-of-order delivery is the portion of packets which arrive later than other, subsequently-transmitted packets. It increases as a function of the difference in queuing delay on parallel paths.
- **All of these properties become worse with time and distance**, which is a reasonable definition of a perishable commodity.

Fruit Example..

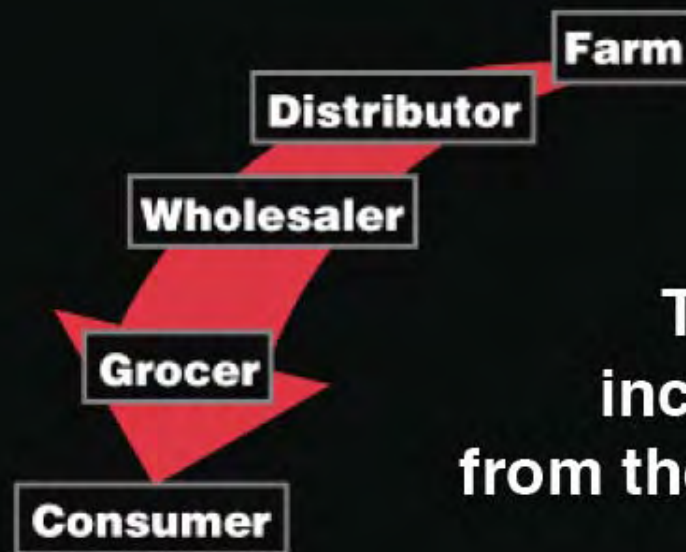
Value decreases with time & distance



The **value** of a piece of fruit decreases, the further it gets from the farm which produced it.

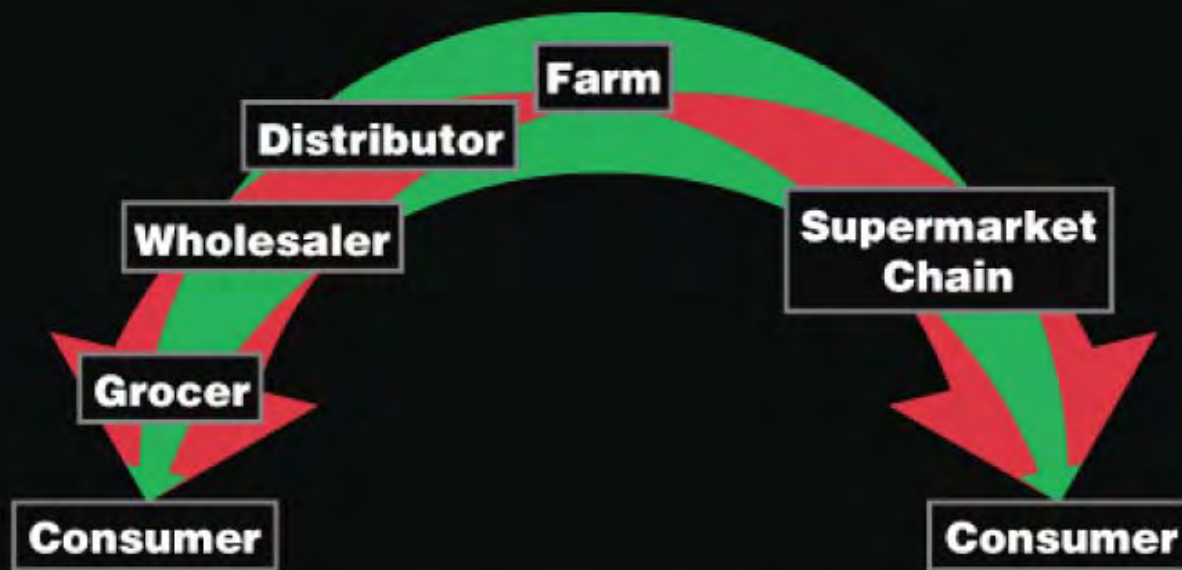
The shelf-life which the consumer can expect decreases, and eventually it becomes overripe, then rotten.

Cost increases with time & distance



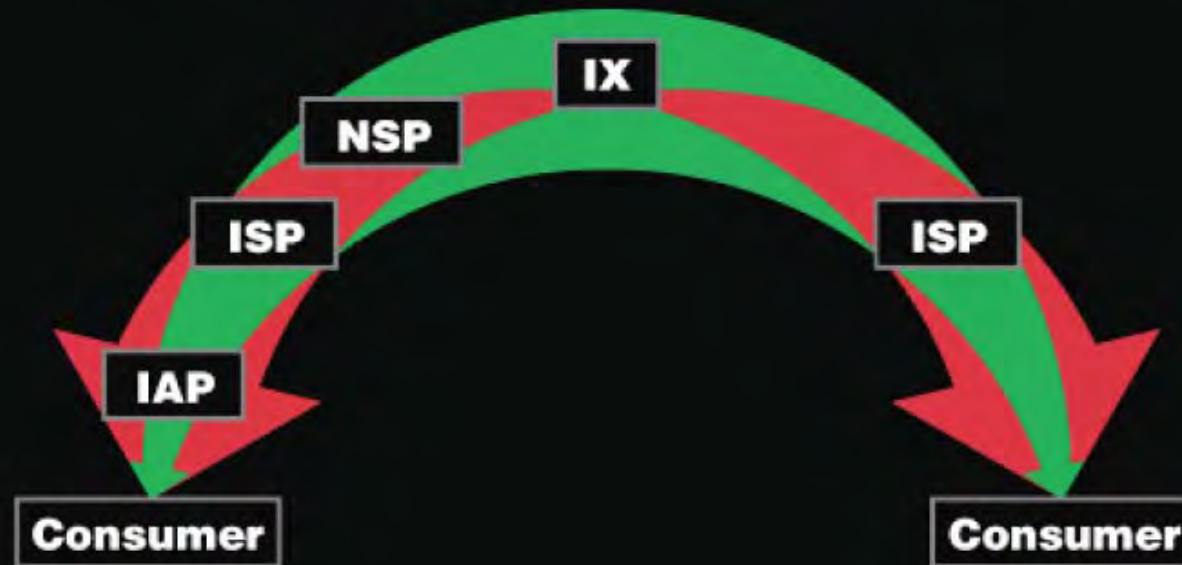
The **cost** of a piece of fruit increases, the further it gets from the farm which produced it.

Salaries and hourly labor, warehouse leasing, petrol, lorry amortization, loss and spoilage, insurance, and other factors contribute additively.



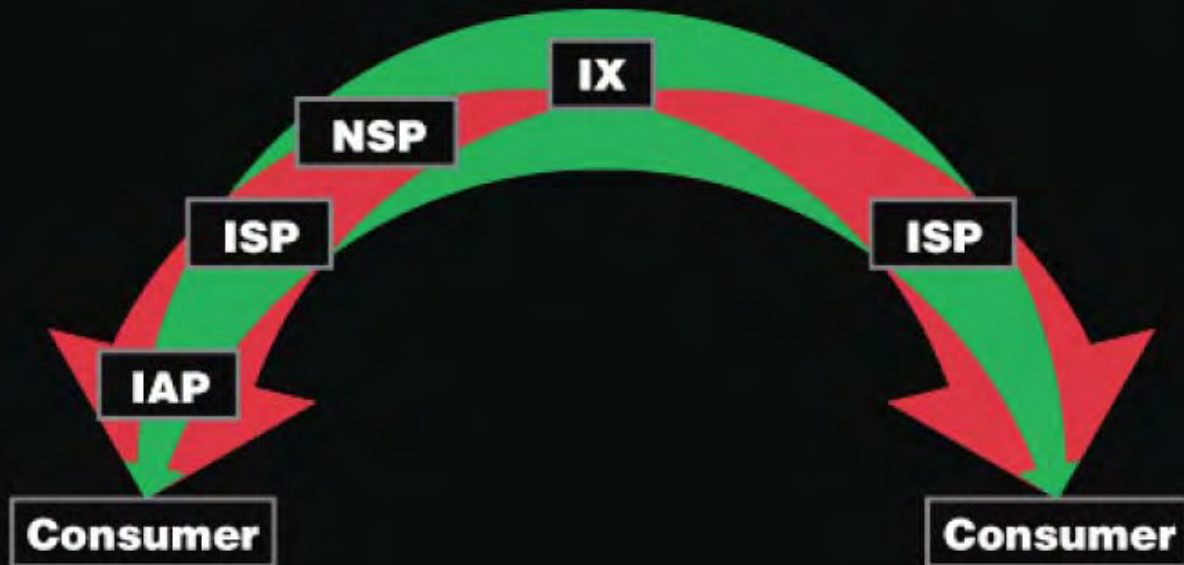
In a competitive environment, retail price is limited by competition, so time and distance influence the price more than the number of middlemen.

The problem is the same:

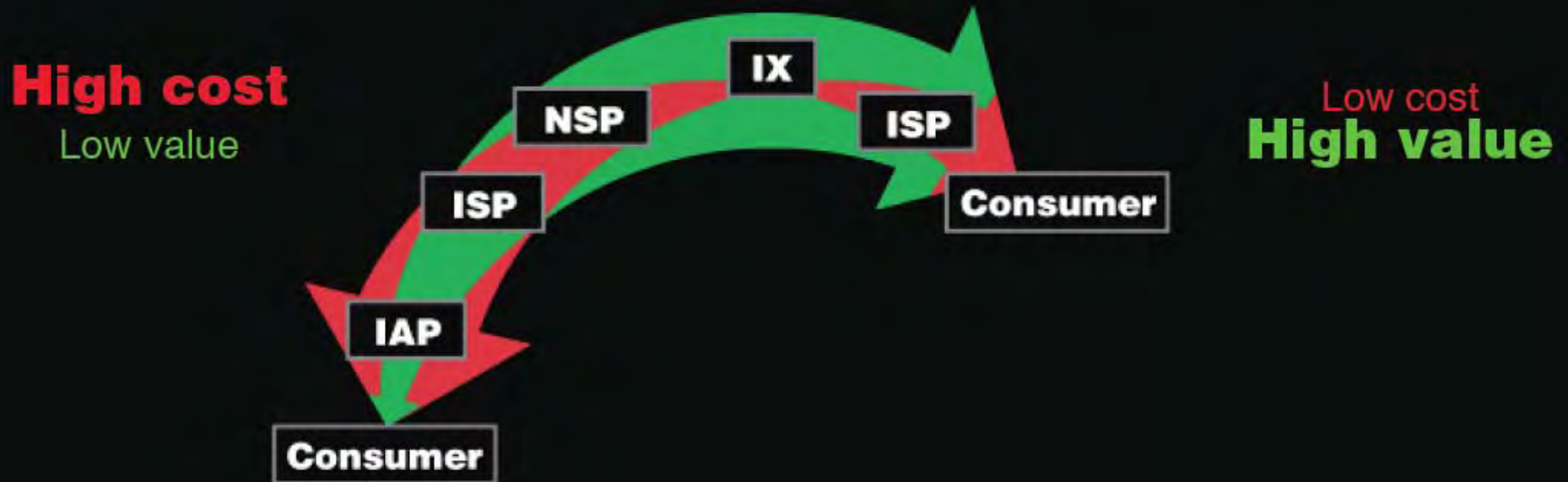


ISPs form a delivery chain, bringing perishable bits to the consumers who purchase them.

So how do we improve things?



Bring the customer nearer an IX...



...or bring an IX nearer the customer.

History of IXs

- **Metropolitan Access Experiment**
- **Metropolitan Area Ethernet**
- **Metropolitan Area Exchange**
- **WorldCom MAE-East™**
 - Washington, D.C.**
 - 10mb shared FOIRL(Magnums) into assorted switches**
 - No fixed topology, four initial locations**
 - MFS fiber plant**
 - Shared administration**
 - Late 1992**
 - Sprint/ICM, Altnet, PSI, SURAnet, NSFnet**

First Exchanges

- **MAE-West / Federal Internet Exchange**

 - San Jose / Mountain View (1994)

 - FDDI “dumbbell” ring

 - Bridged to 10mb Ethernet in many locations

 - Two locations, two administrations

- **Commercial Internet Exchange**

 - Palo Alto (1994)

 - Layer-3 MMLPA

 - Commodity DS1 (T1) lines into a Cisco 7010

 - 1991 Not-for-profit industry association

 - Alternet, PSI, Cerfnet

First Exchanges

- **Hong Kong Internet Exchange**
 - Chinese University of Hong Kong**
 - Single location Ethernet switch**
 - Administered by the university**
 - First major free exchange (1995)**

IXP Governance and Commercial Models

- In most markets, IXs are a natural monopoly
 - problem of trust between competitors
 - risks of abuse and conflicts of interest
- Successful IXs are not usually:
 - owned, operated or housed by a single ISP or carrier
 - ISPs or wholesale IP transit providers
 - national or international backbones
- Co-location facility neutrality:
 - normally (mainly in Europe) these are buildings operated by independent commercial companies
 - though sometimes (mainly in US) co-los operate IXs
 - IXs tend not to be in carrier co-lo facilities

Successful IXP Neutrality Principles

- Does not compete with its ISP members/customers
- Does not discriminate between its ISP members/customers
- Does not move traffic between countries or continents.
- Does not make exclusive arrangements with:
 - ISPs
 - Carriers
 - Co-lo Providers
- Does not provide IP transit routing
- Does not take share of ISPs' transit revenues
- Only interconnects between metro area co-lo sites
- May be present at multiple co-lo sites and providers

Governance/Commercial Models

- Operated by public sector national academic network
 - BNIX, GIGAPIX
- Not-for-profit membership associations of participating ISPs
 - LINX, AMS-IX
 - Over 90% of the 400+ IXPs globally work this way !
- Service within commercial co-location operator
 - Equinix, PAIX, IX Europe
- Companies whose shareholders are participating ISPs
 - MIX, JPIX
- Independent neutral commercial companies
 - XchangePoint, JPNAP

Internet Exchanges in World

- IXP operators are typically:
 - neutral
 - not-for-profit membership organizations
 - do not run hosting/co-location facilities
 - not same organization as co-location provider

Technologies in IX

- Initially (1992-4):
 - 10Mb/s Ethernet from ISP router to IXP switch
 - FDDI between IXP switches
 - Single switch in single location
- 100Mb/s Ethernet mostly replaced these 5+ years ago
- Some use of ATM meantime
- 1Gb/s Ethernet now common access technology
- 1Gb/s Ethernet also used in core of networks
- 10Gb/s Ethernet increasingly common in IXP cores
- Some limited use of DWDM and MPLS

Gigabit Ethernet

- Cost-effective and simple high bandwidth
- Most common technology for many ISPs accessing major IPPs
- Works well for local and metropolitan distances
- Proven and deployed at most major IPPs
- Almost universally used for IPP inter-switch links
- Technology is mature and price dropping
 - e.g. 1Gb/s over copper
- Cost-effective high-performance switches available from various vendors:
 - Cisco, Extreme, Foundry, etc..

Requirements to participate in IX

- Your own Autonomous System (AS) number
 - you need this if you take service from >1 ISP anyway
- Your own IP address space
 - need to become “registrar” of NRO member registry
e.g. AFRINIC, RIPE NCC, ARIN
- Router(s) which can do BGP
 - most medium/large Cisco/Juniper routers
 - also open-source based PC platforms
- Space in one of the co-lo facilities at which it is present

IXs and TLD's Name Servers

- Many root, gTLD and ccTLD name servers are operated at major Internet Exchanges
- Significant advantages:
 - Low latency
 - Cheap high bandwidth
 - Rich connectivity
 - Many players can enjoy direct connection
 - Increases resilience and robustness
- RIPE NCC hosted first instance of “k” root server at LINX
 - now also in Amsterdam, Frankfurt, Reykjavik...
- Nominet's “.uk” primary and secondary ccTLD servers at both major London IXPs
- Various instances of f, i, j & k root servers at many Euro IXPs

IX Resources

- Packet Clearing House
 - <http://www.pch.net>
- RIPE EIX (European Internet eXchange) Working Group
 - <http://www.ripe.net/ripe/wg/eix/>
- Euro-IX Association of IXP Operators
 - <http://www.euro-ix.net>
- Global IXP Directory
 - <http://www.ep.net>
- But the single most useful resource is self-help from the IXP and ISP community

Services offered thru IXs

- **Route-server**
- **Looking-glass**
- **Measurement and instrumentation**
- **Network Time Protocol**
- **Web cache parent**
- **News server**
- **Root server mirror**
- **Software Mirrors**

IX Economics

- Exchanges must be inexpensive, not reliable.
- There should be multiple exchange point *operators*, but only one *switch fabric* per geographic region.
- Open peering promotes growth; closed peering causes stagnation.
- All but a half-dozen ISPs purchase transit.
- Thus all peering is simply an economic optimization versus transit. It exists simply to reduce the average per-bit cost. Shorter paths are a collateral benefit.
- The effectiveness of peering can be directly measured as a function of its reduction of per-bit cost versus transit.
- Thus making peering inexpensive is more beneficial to its effectiveness than making it reliable.

IX Economics

Example:

Transit may cost \$0.50/gigabit.

Reliable (99.999%, 26 sec/month downtime) peering may cost \$0.40/gigabit.

Unreliable (99.9%, 45 min/month downtime) peering may cost \$0.005/gigabit.

- Transit cost: \$500/megabit/second/month at 40% utilization = \$0.50/gigabit.
- Reliable exchange cost: \$10,000/month for 100 megabits/second at 15% utilization = \$0.40/gigabit.
- Unreliable exchange cost: \$500/month for 100 megabits/second at 50% utilization = \$0.005/gigabit.

IX Economics

- Example ISP ships 10 terabits/month (approximately 40 megabits/second average)
- If exclusively by transit, \$5,000/month.
- If 50.001% by transit and 49.999% by reliable exchange, \$12,500.05/month.
- If 50.1% by transit and 49.9% by unreliable exchange, \$3,005/month.
- **Reliable exchange saves \$4.95 of transit, but costs \$9,500 extra each month.**

Open Peering is the Only Effective Way to Create Value

- Any one ISP's customers make up an insignificantly small portion of the Internet.
- The amount your customers are paying to reach your other customers is insignificant, relative to the amount they're paying to reach everyone else's customers.
- The value which you as an ISP have to sell to your customers is the sum of the bandwidth at each of the exits of your network, weighted by the number of routes available through each.
- There are three ways to increase the value which you have to sell to customers:

Buy it (purchase transit)

Sell it (sell transit) or

Peer

Open Peering is the Only Effective Way to Create Value

- Purchasing transit is expensive. Although it is generally necessary and desirable to purchase some transit, economic optimization requires that it be used as little as possible.
- A network cannot survive by reselling transit alone, as that would be an unnecessary middleman position
- Selling transit is necessary and desirable, as that increases the number and size of your customer base, the group of people who pay you money.
- However, you cannot sell as fast as the Internet grows overall, so the portion of the Internet which consists of your customers will decline over time.
- Thus selling transit is too slow a means of increasing value. It also constitutes a chicken-and-egg problem: if you depend upon sales for growth of value, and depend upon growth of value to fulfill new sales, you cannot gain momentum.

Open Peering is the Only Effective Way to Create Value

- Adding peering bandwidth both costs less and can be achieved more quickly than adding either purchased or sold transit bandwidth, since it's both geographically aggregated and temporally flexible.
- Switching from an open peering policy to a closed peering policy will necessarily retard the growth-rate of your network, both in absolute terms, and relative to your competitors who are growing through the addition of new peering bandwidth.
- **No network has ever been profitable while pursuing a closed-peering strategy.**

Collateral Lesson:

Peering and Sale of Transit are Complementary, not Mutually Exclusive

- There exist a set of related fallacious beliefs which cause innumerate people in this industry to lose money:

That it is not advantageous to peer with one's customers,

that refusing to peer with another ISP can do them disproportionately more harm than it does one's self,

and most ridiculously, that if you refuse someone peering, they might become your customer instead.

Collateral Lesson 1: Peering with Customers is Good

- Any peering increases the amount of bandwidth you have available to sell to your customers.
- If you peer with a customer, it increases the amount of bandwidth which you can sell to other customers.

Collateral Lesson 1: Peering with Customers is Good

- Peering with a customer means offering them free routes *to* your other customers within the same region.
- If 0.1% of your traffic is between your own customers, and you peer with 10% of your customers, 0.001% of your traffic is between a pair of customers which are both also peers.

Collateral Lesson 1: Peering with Customers is Good

- Thus by peering with 10% of your customers, you increase the bandwidth you have to sell to customers 0.1% of the time. In exchange, you either sacrifice payment for 0.001% of your traffic or need to create a new billing method for it. These numbers are both insignificant.
- What's important is that it allows you to have a uniform peering policy without having to special-case an exception class.

Collateral Lesson 2: Refusing Peering Hurts You Both

- When an ISP refuses to peer with another ISP, both are hurt.
- ISPs which refuse to peer are generally failing to peer with a set of other ISPs which *collectively* advertise more routes than the ISP which is failing to peer.

Collateral Lesson 3:

If You Refuse Someone Peering, You Create a Customer for your Competitor

- If ISP “A” refuses to peer with ISP “B,” two possibilities exist:
 - B will buy transit from one of A’s competitors to reach A, or
 - B will peer with one of A’s transit providers to reach A.
- At best, A loses the possibility of selling B transit, and creates a customer for one of their own competitors.
- At worst, A loses the possibility of selling B transit, and has to pay to receive traffic which B can send for free in any volume.

IX Economics Summary

- Only an inexpensive exchange can succeed.
- Only one switch fabric should exist in a region.
- Connections should be offered by multiple exchange point operators with different facilities at different price points.
- ISPs which wish to grow and be profitable must peer with everyone they can.

Internet Exchanges World Wide

- **More Than 400 IXs Today**
- ~ 150 in North America
- ~ 150 in Europe
- ~ 80 in Asia-Pacific
- 11 in Africa
- 10 in Latin America
- ~ 2 in the Middle-East
- 1 in Caribbean

North American Internet Exchanges



European Internet Exchanges



Asia-Pacific Internet Exchanges



African Internet Exchanges



Latin American Internet Exchanges

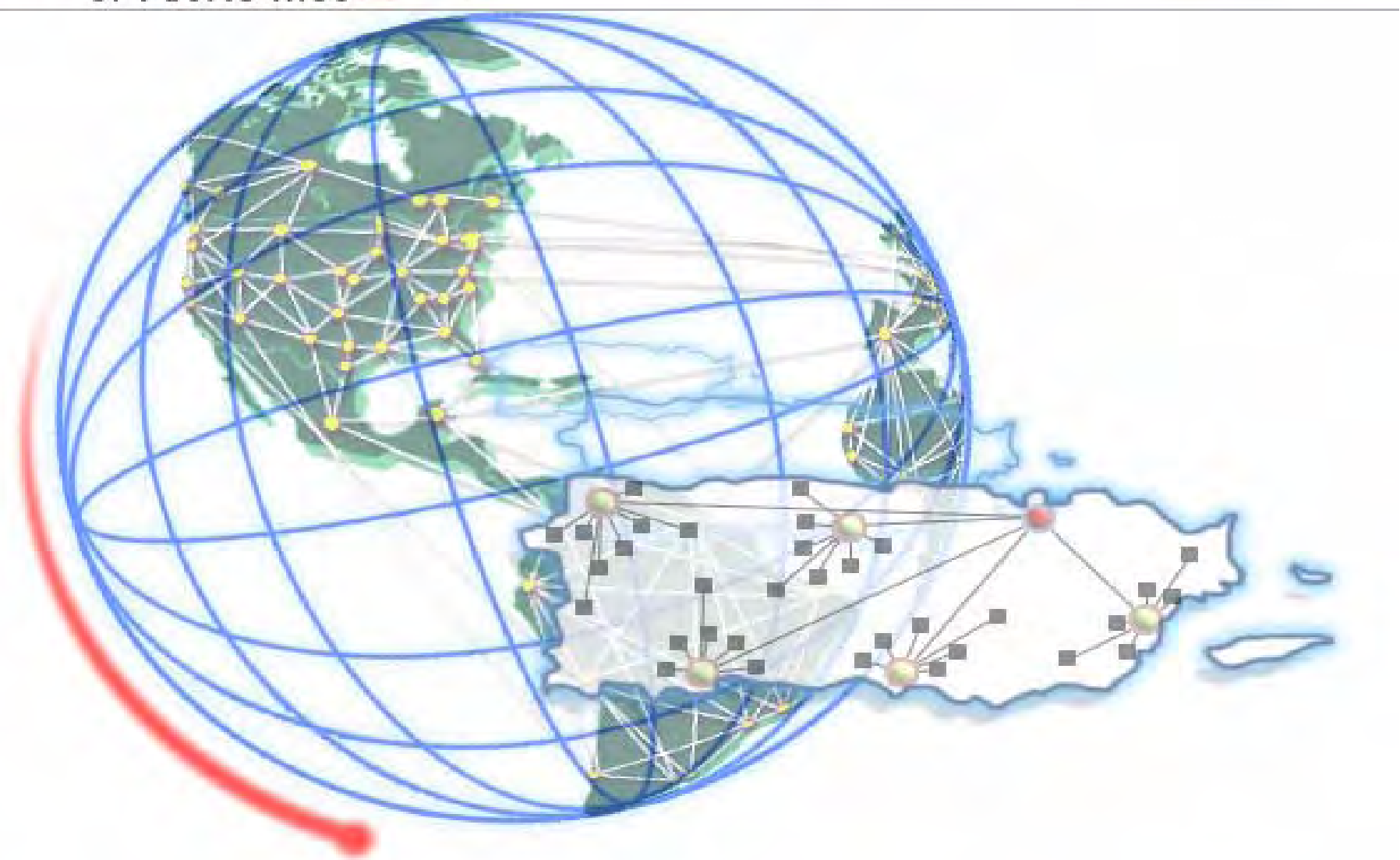


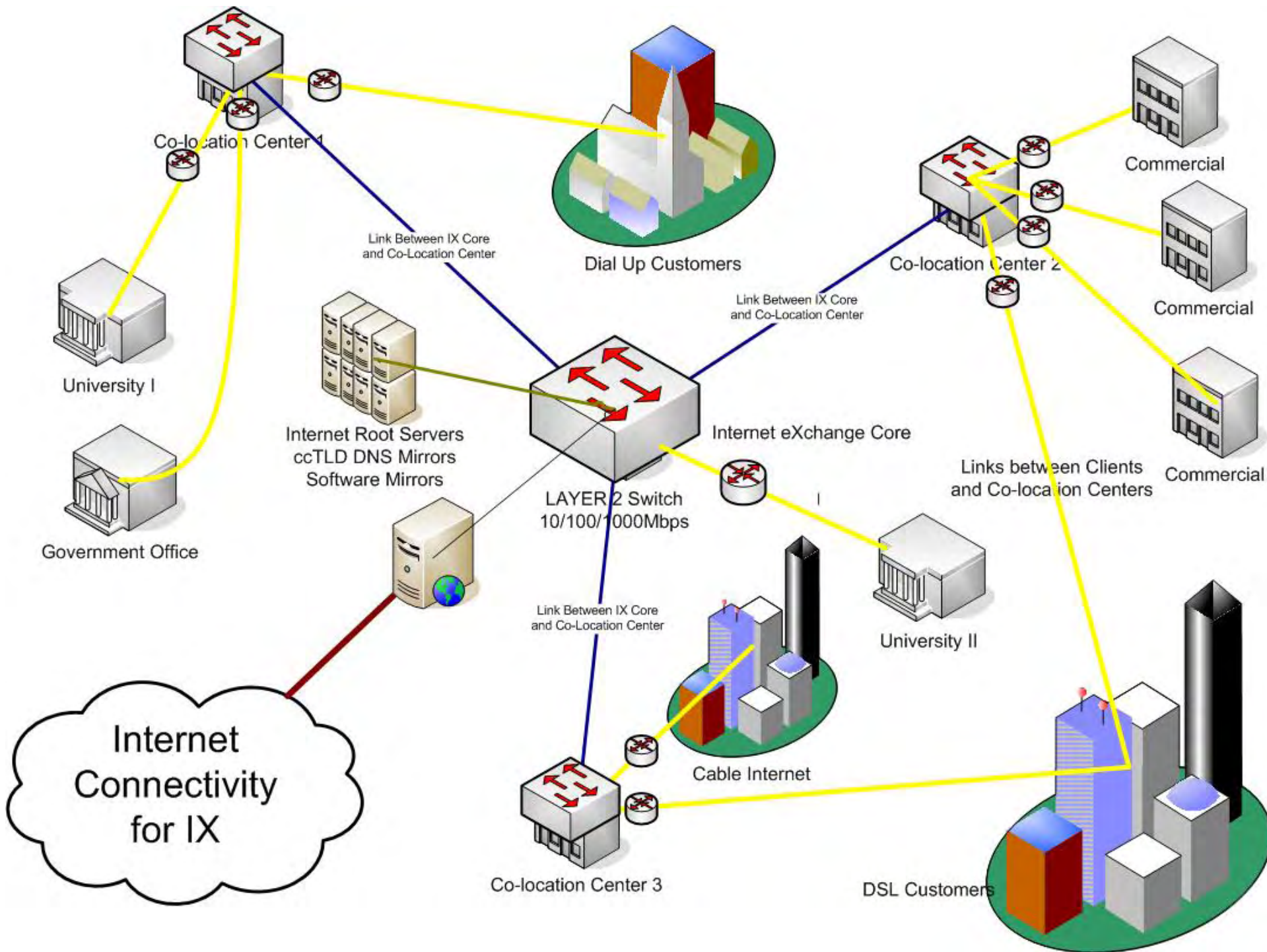
IX.PR is the first Caribbean Exchange



ix.PR
Internet Exchange
of Puerto Rico

010001010100101001011011010001001
0000010101010001001001001001000





References

- Bill Woodcock's (Director of PCH.net Research) presentation on Internet Exchanges, NANOG 35
- EP.net
- LINX (London Internet Exchange)
- European Internet Exchange Association
- Wikipedia,
http://en.wikipedia.org/wiki/Internet_exchange_point