# Do We Need Resource Reservation?

## - Reality, future, and challenges

Ping Pan

(pingpan@research.bell-labs.com)

# Resource Reservation:

- Set aside network resources, such as bandwidth, for a particular data flow.
- Need to signal the network routers on the amount and the quality of services.
    - signaling protocols;
    - active probing;
    - provision in advance;

# Reservation on Reservation:

- Over-provisioning is good enough
  - Not a “technical” issue.
  - Driven by the market.
  - Backbones are practicing it: OC-192 links in Uunet.
- Application can adapt
  - Buffering, fast retransmission, adaptive compression ratio…
  - Available in *vic, vat*….

# Reservation on Reservation: (cont)

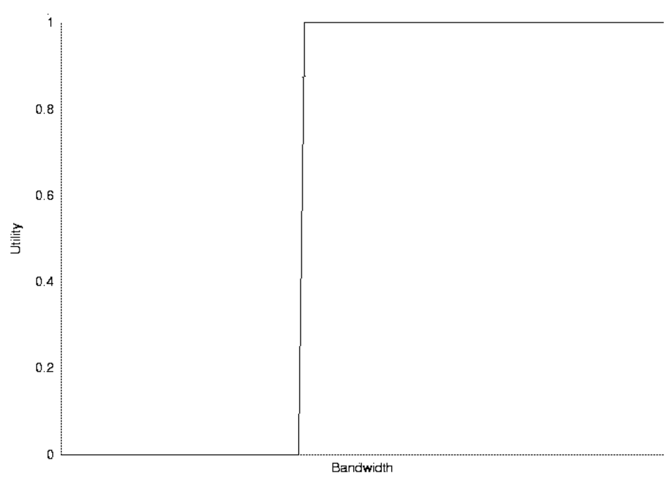
- Real-time traffic volume is small
  - Majority traffic is *elastic* (loose time constraints).
  - “Real-time” streams are cached.
  - Requirements are very strict, thus an over-kill.
- Scaling problem #1: adding processing and memory cost to routers
  - … just take a look at RSVP, and you will understand.

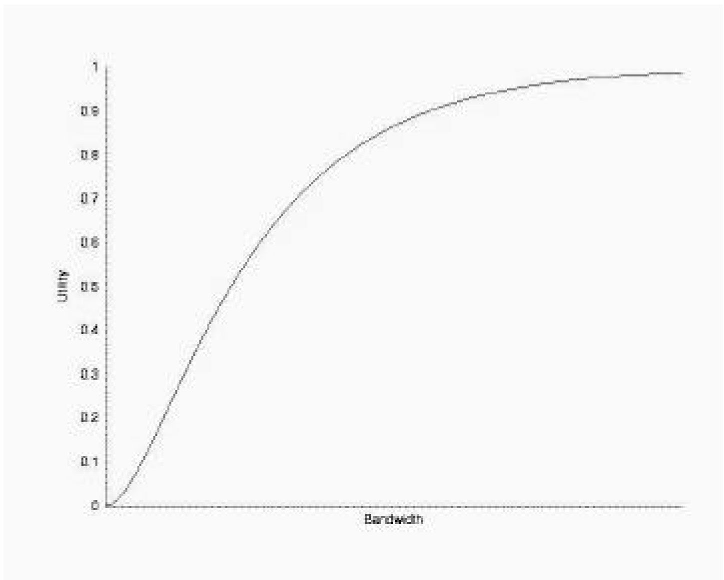
# Reservation on Reservation: (cont)

- Scaling problem #2: adding complexity to service providers
  - Managing "flows" individually requires sufficient accounting and billing procedure.
  - Large number of "flows" is too hard to manage.

# Performance Study:

- Best-Effort vs. Reservation
  - By Shenker and Breslau in Sigcomm'98
  - Define utility function for rigid and adaptive applications:



# Performance Study: (cont.)

- Bandwidth Benefits:
  - Compute the total utility in best-effort and reservation-capable to be B and R.
  - Define *bandwidth gap,* ? , s.t. R = B + ?
- Results:
  - Rigid applications: ? > 0
  - Adaptive applications
    - Possion and exponential load distribution: ? = 0
    - Algebraic load distribution: ? = constant

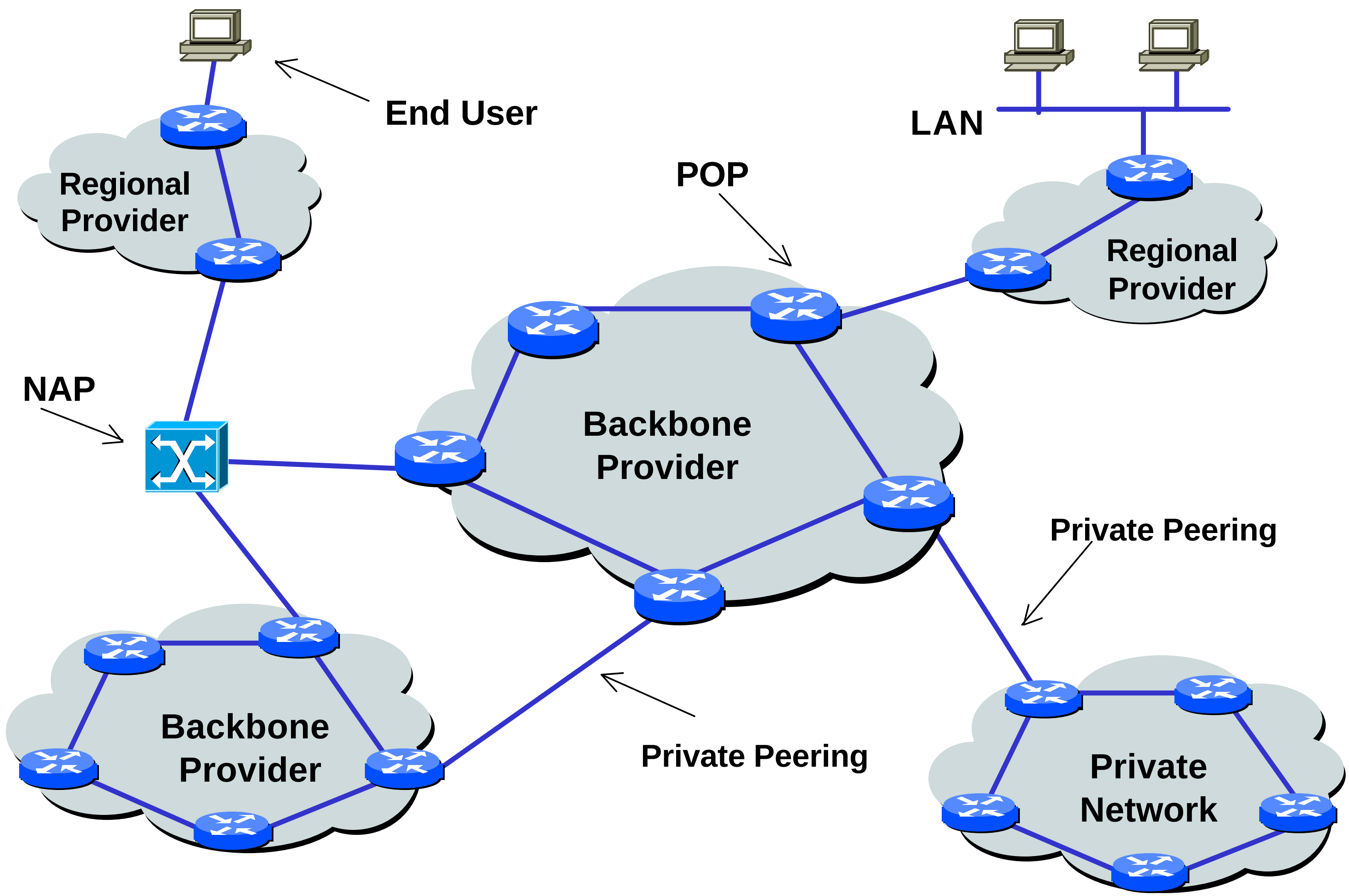
# Current Network Condition

- The Paxson Study:
  - Between1994 and 1997, from 35 wide-spread sites, collected over 20,000 TCP connection traces.
  - Define *available bandwidth,* as the proportion of the total network resources that were consumed by a connection itself.
    - 1 means resource available; 0, none.
- Results:
  - The range of available bandwidth is very broad, from very little to almost 1.
  - End-to-end delay variations: primarily between100-1000 msec, extended frequently to much larger times.

# Current Network Condition (cont.)

- Implications:
  - Network has both low and high bandwidth links;
  - When going over low-bandwidth links, congestion.
- Questions:
  1. Where are the bandwidth bottleneck links?
  2. Can we simply add bandwidth to the bottleneck links?
  3. How real is end-user traffic actually going through bottleneck links?

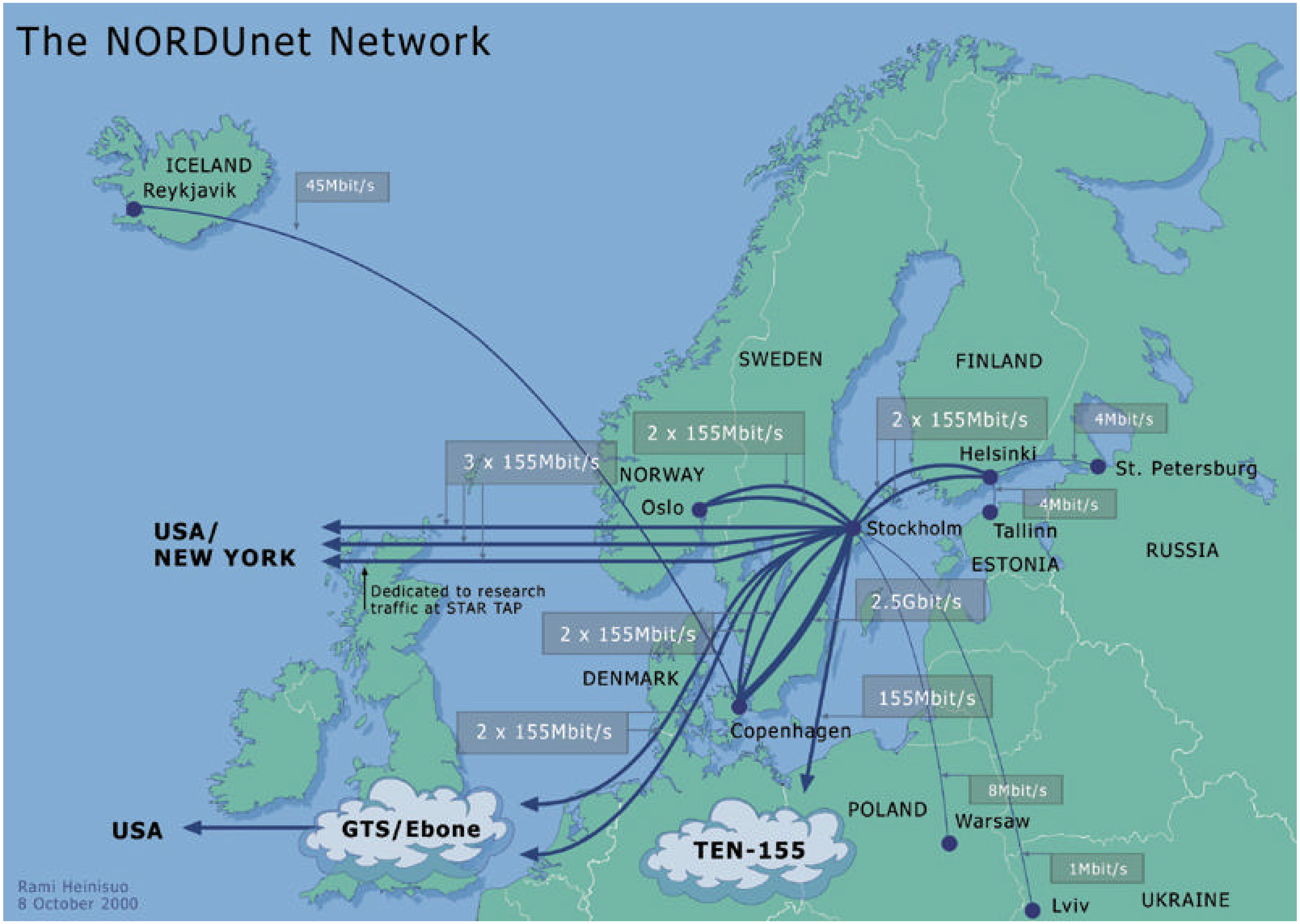
# Internet Structure



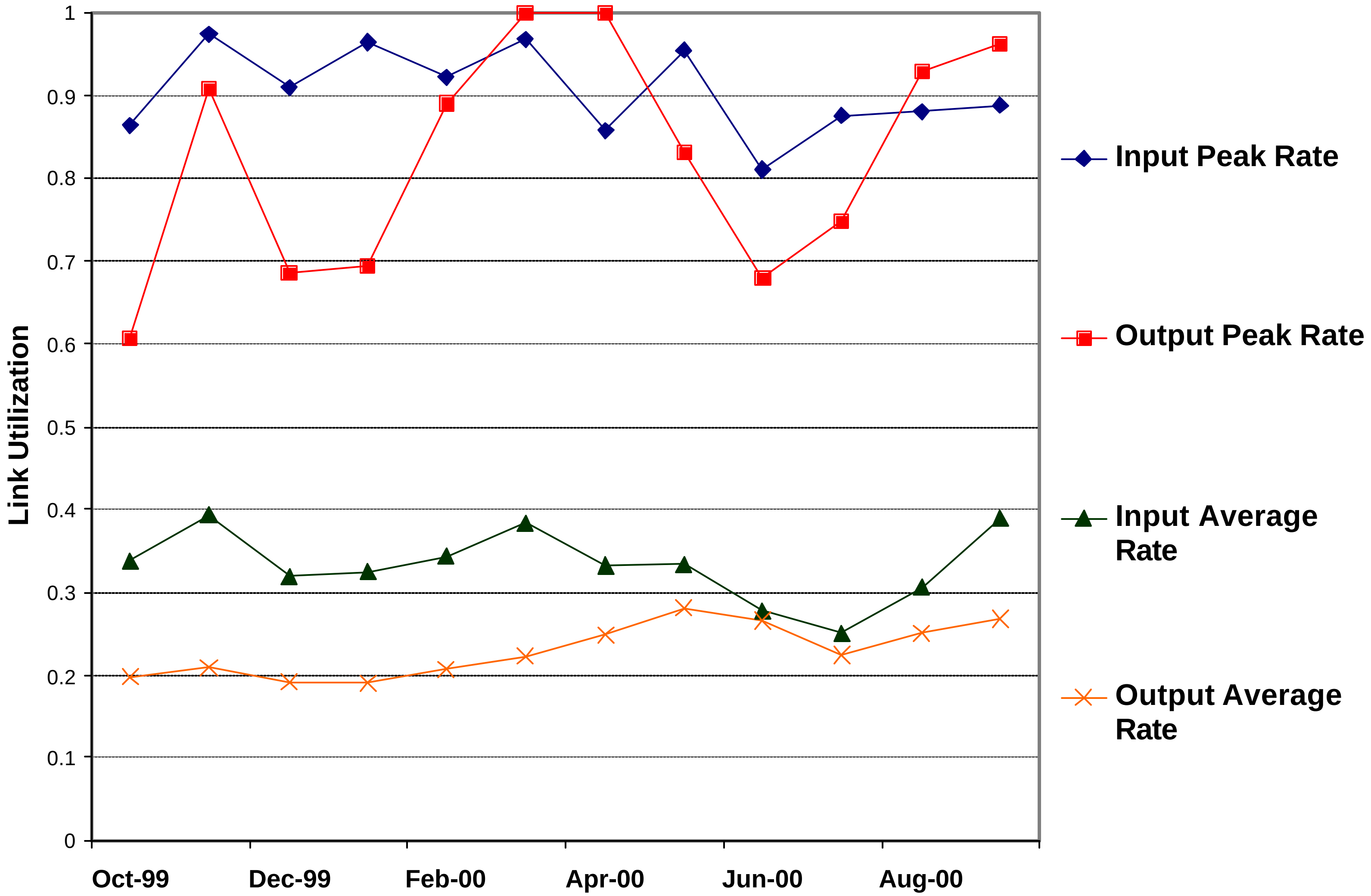
# Bandwidth Bottlenecks

- Internet links:
  - Backbones (OC-3 to OC-192, and more)
  - Private networks (full-meshed T1, T3)
  - LAN (10M and 100M Ethernets)
- Network utilization according to Odlyzko
  - Backbones (10-15 %)
  - Private networks (3-5 %) (for low transaction latency)
  - LAN's (1%)
- *Access Links:* interconnect SP's and connect private networks to SP's.

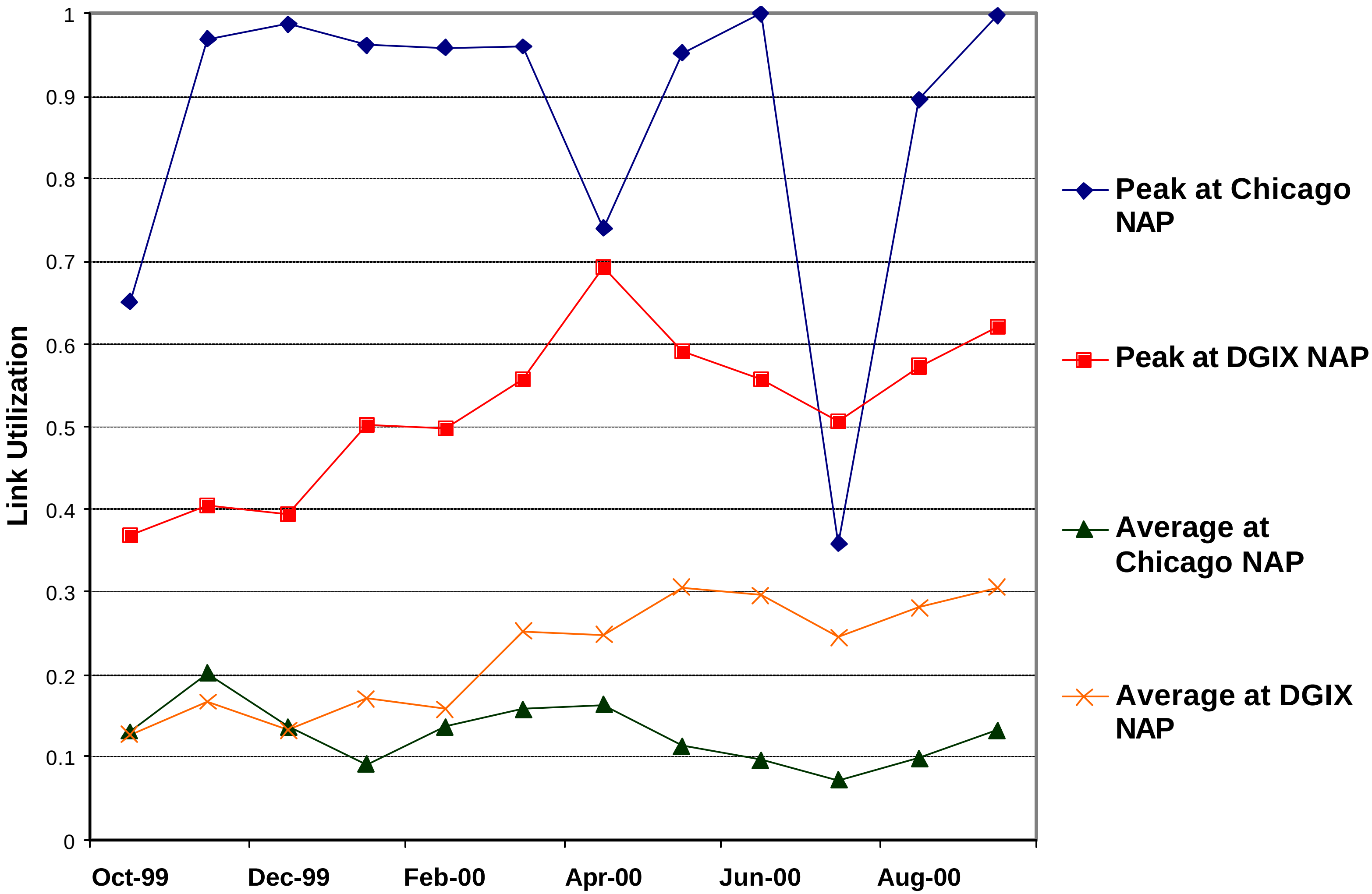
# NORDUnet Traffic Analysis



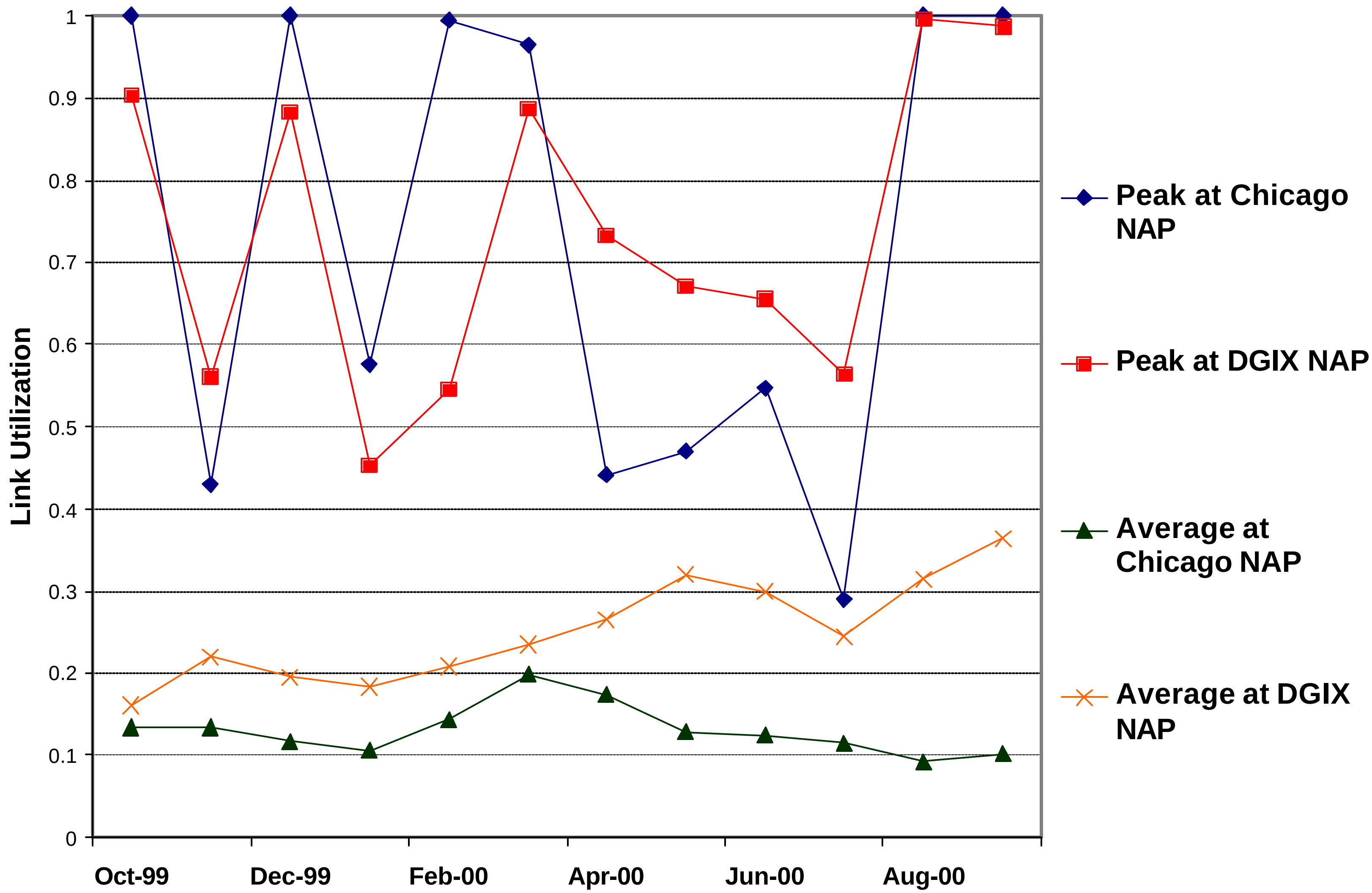
# NORDUnet: Trans-Atlantic Links (3 OC-3)



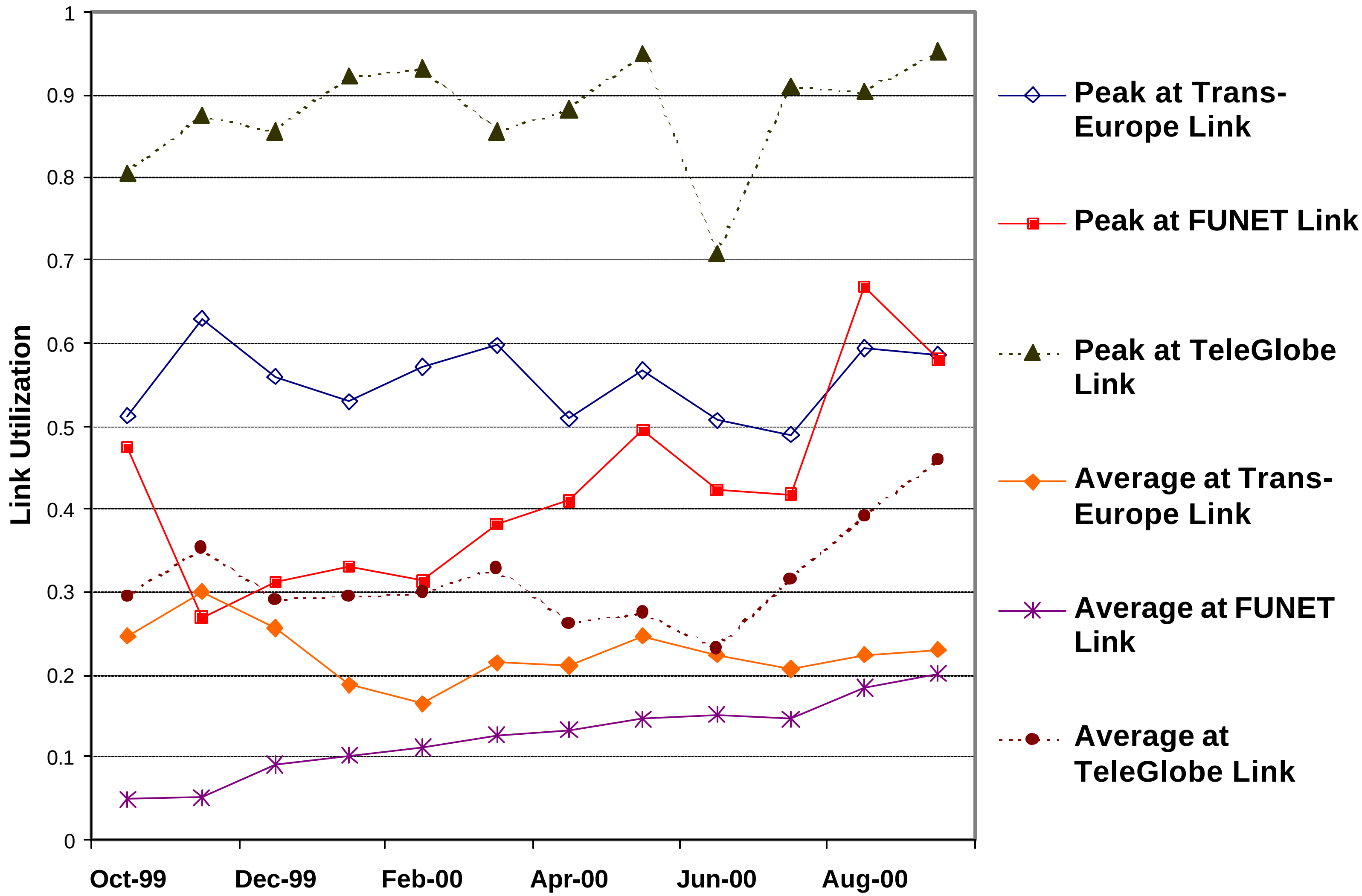
# NORDUnet: Input Traffic At NAP's



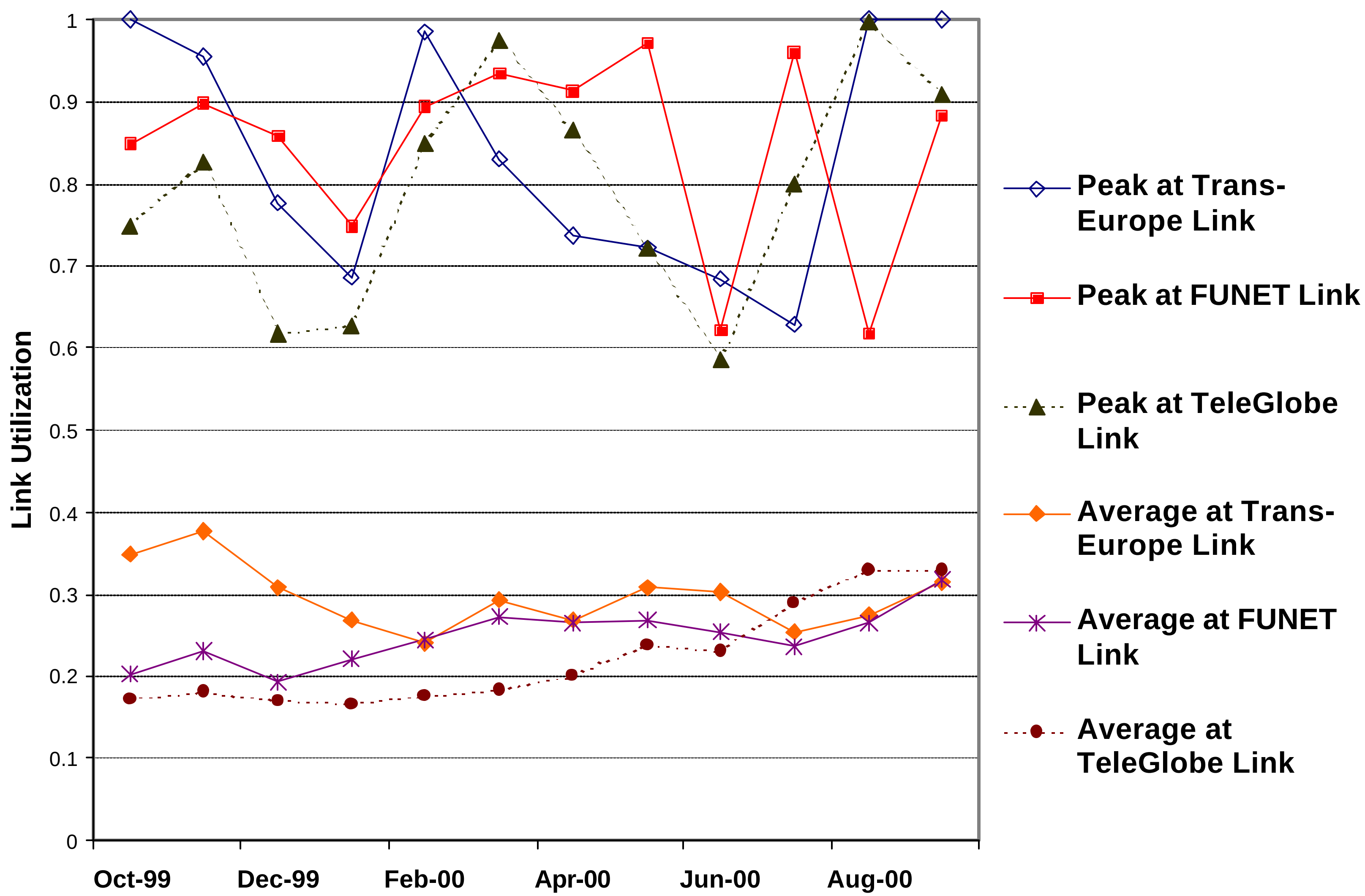
# NORDUnet: Output Traffic At NAP's



# NORDUnet: Input Traffic At Peering



# NORDUnet: Output Traffic At Peering



# NORDUnet Traffic Analysis

- Taken from 12/36 busiest links
- Results:
  - All access links, including trans-Atlantic links, can get congested.
  - Average utilization is low: 20-30%
  - Peak utilization can be high: up to 100%
  - Peak duration can be very long:
    - Chicago NAP congested once in 8/00, lasted 7 hours.
    - TeleGlobe links congested every workday in 8/00 and 9/00
- Reasons:
  - Frequent re-configuration and upgrading.
  - Load balancing to protect own users.

# Bottleneck Link Summary

- From several other ISP’s, we found
  - Average link utilization is always low
  - Occasionally having long-lasting peak utilization.
- Determining Factors:
  - Congestion Ratio = peak/average
    - Can be as high as 5.
  - Congestion Duration

# Bandwidth Pricing

- Reality: leased bandwidth price has *not* been dropping *consistently* and *dramatically*.
- Facts:
  - 300 mile T1 price:
    - 1987: $10,000/month
    - 1992: $4,000/month
    - 1998: $6,000/month (thanks to high Internet demand)
  - 100-mile cabling cost in 1998: $65,000
  - US is 4 times cheaper than Europe national price, and 16 times cheaper than Europe international price.

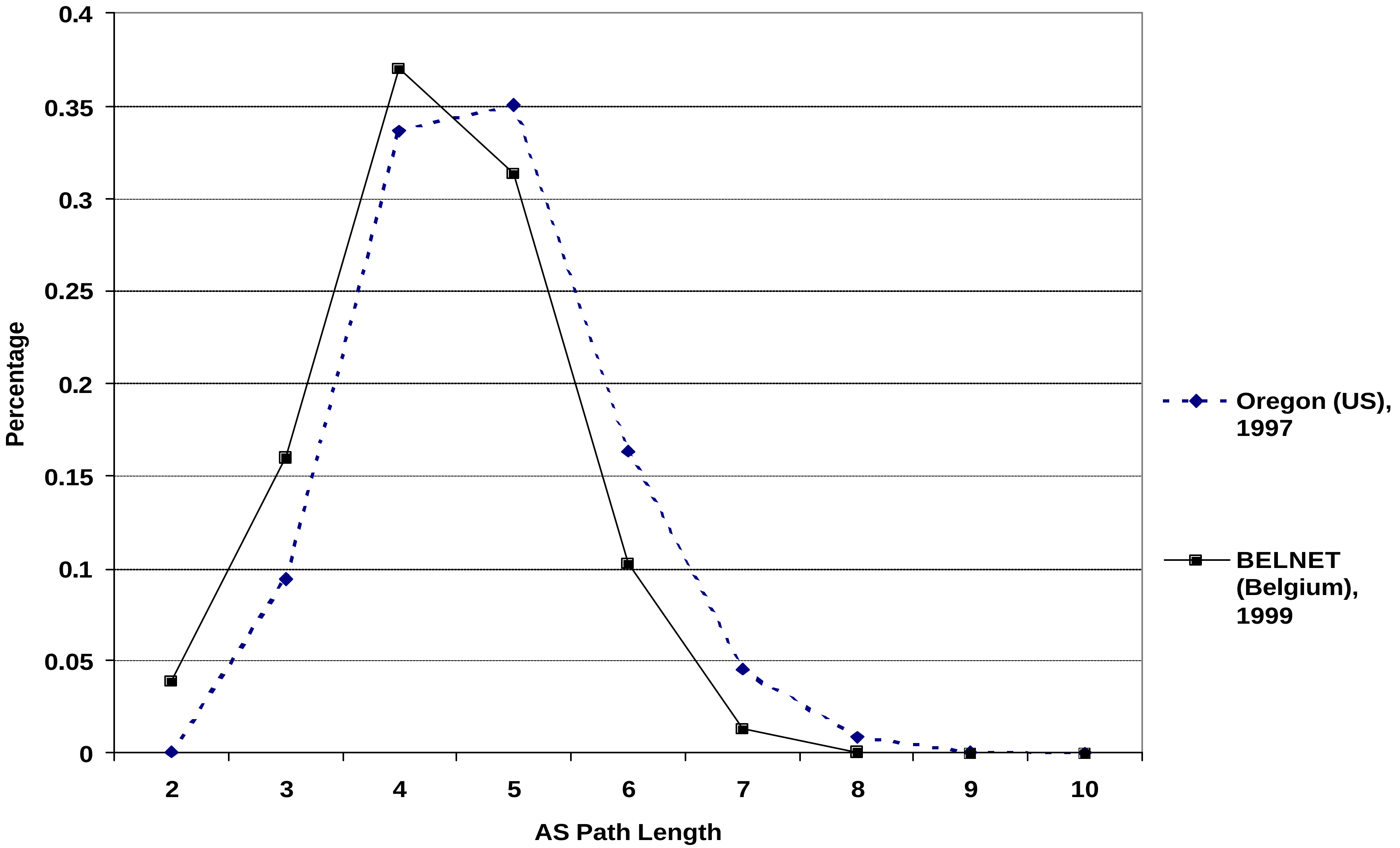
# Bandwidth Pricing (cont.)

- Connecting ISP's is very expensive
- Facts:
  - Transit DS-3 link: $50,000/month between carriers.
  - Transit OC-3 link: $150,000/month between carriers.
  - Chicago NAP:
    - $3,900/month/DS-3,
    - $4,700/month/OC-3.
- Conclusion:
  - It is very expensive to over-provision bandwidth bottleneck links.

# Network Heterogeneity

- Reality: end users go through multiple provider networks to reach each other.
- Facts:
  - A few years ago, 60% of Telstra (largest Australian SP) traffic was with US providers
  - The longest Internet hop-count in 1997: 31
  - The average provider-hop-count (in BGP AS): 3-4
- Conclusion:
  - A large portion of end-user traffic goes through the bottleneck links.

# Network Heterogeneity (cont.)



# Internet Traffic Trend

- **VPN Traffic:**
  - In 1999, ~50% traffic going through private links to the Internet.
  - in 1997, a $10 billion business.
- **Real-Time Traffic:**
  - In 1999, US phone traffic was 40,000 TB/month; Internet traffic was 10,000 – 16,000 TB/month.
  - Will be a fool to count out IP telephony traffic volume.
- **Disruptive Innovations:**
  - WEB: traffic double every 3-4 months in 1995 and 1996
  - Napster: same growth in some school networks in 1999.
- **Conclusion: How can we over-provision network resources for the future?**

# Reservation Scaling Issue:

- **Memory scaling:**
  - RSVP requires 5 MB of memory for 10,000 sessions.
  - In 1991, backbone routers operated with 16 MB
  - In 2000, backbone routers operated with 256 MB (thanks to the cheap memory cost).
  - Conclusion: Will proper software practice, reservation has no memory scaling problem.

# Reservation Scaling Issue: (cont.)

- **Bandwidth scaling:**
  - Reservation is to reserve bandwidth for data flows.
  - Control/data bandwidth ratio:

    ratio = message-size / (refresh-period * data rate)

    in RSVP:

    message-size = 350 Bytes/(PATH-RESV pair)

    refresh-period = 30 second

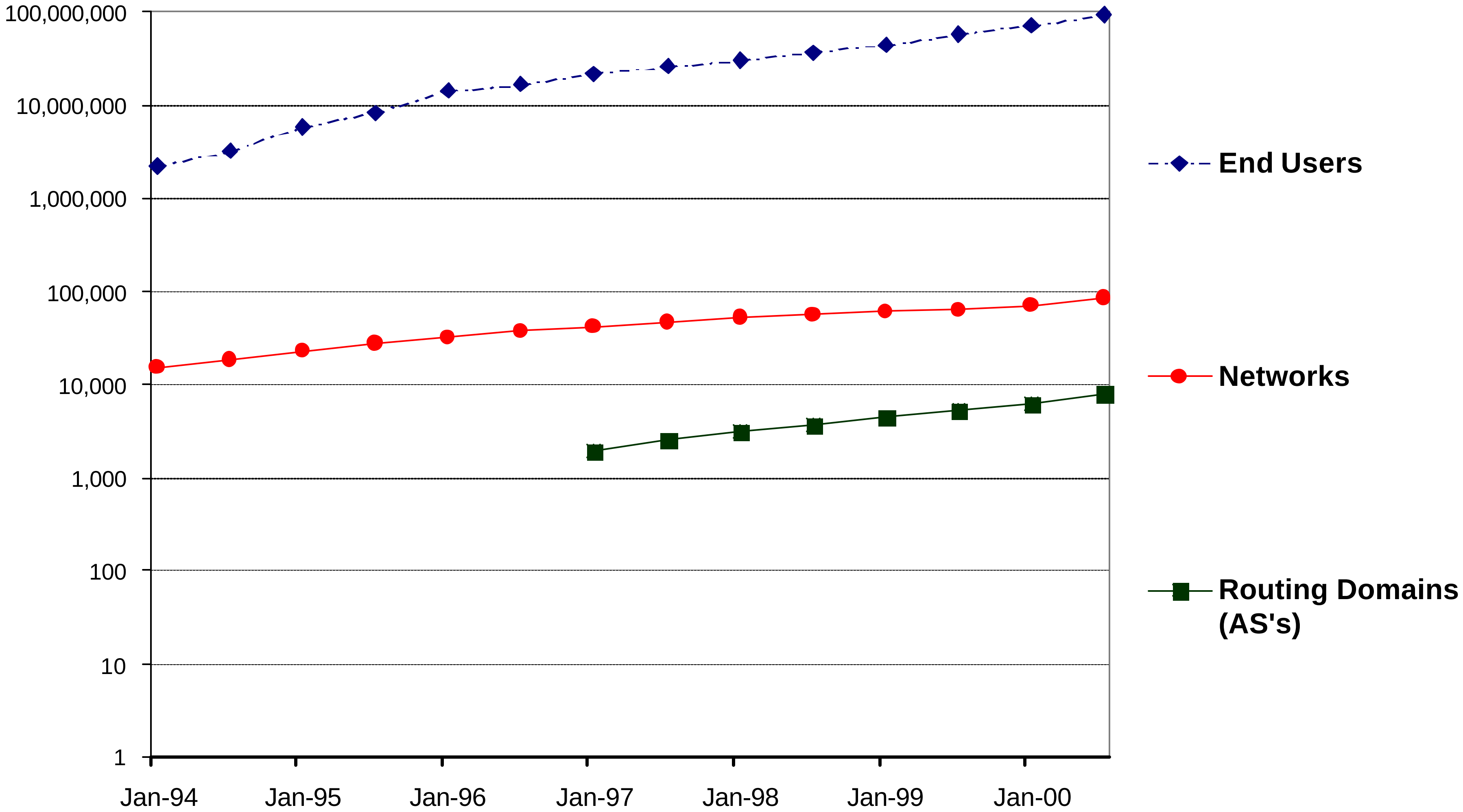
    data rate = 56 kbps

    So, ratio = 0.17%
  - For MPLS setup, data rate can be zero.
  - Refresh reduction proposals, such as staged refresh timer, can reduce the control message bandwidth consumption.
  - Conclusion: This is a manageable problem.

# Reservation Scaling Issue: (cont.)

- **Message Processing Cost:**
  - A stand-alone RSVP implementation with small IBM routers:
    - 1.1 msec per new RSVP session
    - 0.64 msec per refresh RSVP session
  - After simplified reservation process (YESSIR)
    - 70% improvement on the same IBM router platform
    - 10,000 flows/second on a 600 MHz Pentium PC running FreeBSD.
  - Conclusion: Routers can setup and maintain large number of reservations. With simplification and optimization on the protocols, routers can support even more reservations.

# Reservation Scaling Issue: (cont.)



# Reservation Scaling Issue: (cont.)

- **State Management Cost:**
  - This is the real scaling problem.
  - SP's cannot manage too many flows
    - Maximum capacity to date: 70K routing policy entries
    - Forget about per-user reservation management in the backbone.
  - Reservation management's better be simple, secure, robust and scalable.
  - Robust requests the reservations to be able to adapt to network changes.
  - Scalable means to provide admission control at coarse level.
  - This is the focus on reservation research!

# Conclusion

- Over-provisioning may not work for inter-networking communication.
  - Bandwidth may be cheap, but not free.
- There is a trade-off between the cost of adding bandwidth and the cost of traffic management.
- Reservation does not have scaling problem at routers, if implemented correctly.
- Some form of coarse service differentiation is likely to be needed.
- Finally, we do need resource reservation, the question is how.