On Network Dimensioning Approach for the Internet

Masayuki Murata
Advanced Networked Environment Research Division
Cybermedia Center, Osaka University
(also, Graduate School of Engineering Science,
Osaka University)
e-mail: murata@ics.es.osaka-u.ac.jp
http://www-ana.ics.es.osaka-u.ac.jp/

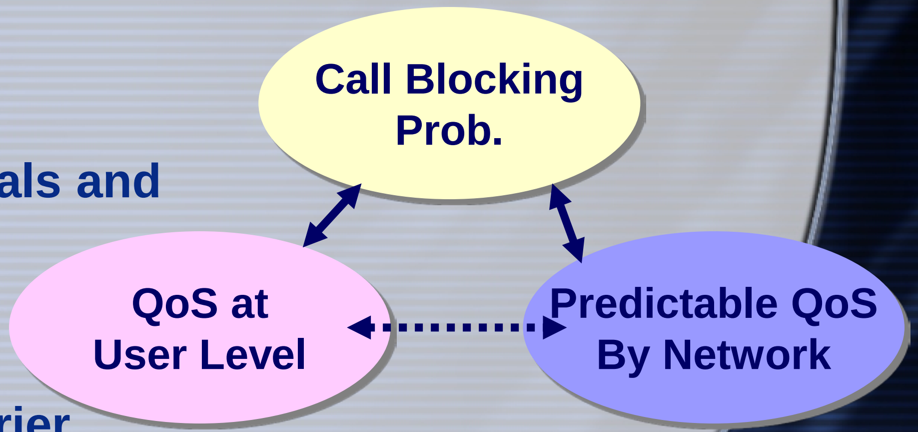
Contents

1. *What is Data QoS?*
2. *Active Traffic Measurement Approach for Measuring End-User's QoS*
3. *Network Support for End-User's QoS*
4. *Another Issue that Network Should Support: Fairness*

QoS in Telecommunications?

- ***Target applications in telecommunications?***
  - **Real-time Applications; Voice and Video**
    - **Require bandwidth guarantees, and that's all**

1. ***Past statistics***
   - **Traffic characteristics is well known**
2. ***Single carrier, single network***
3. ***Erlang loss formula***
   - **Robust (allowing Poisson arrivals and general service times)**
4. ***QoS measurement***
   = call blocking prob.
   - **Can be easily measured by carrier**



What is QoS in Data Applications?

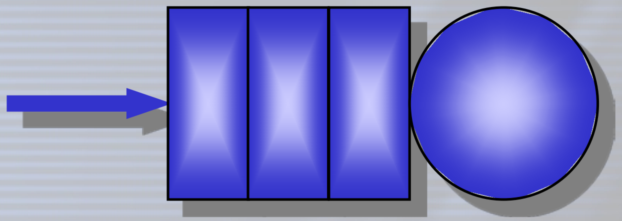
- *The current Internet provides*
  - QoS guarantee mechanisms for real-time applications by int-serv
  - QoS discriminations for aggregated flow by diff-serv
  - No QoS guarantees for data (even in the future)
- *For real-time applications, bandwidth guarantee with RSVP can be engineered by Erlang loss formula*
  - But RSVP has a scalability problem in the # of flows/intermediate routers
- *Distribution service for real-time multimedia (streaming)*
  - playout control can improve QoS
- *How about data?*
  - Data is essentially greedy for bandwidth
  - Some ISP offers the bandwidth-guaranteed service to end users
    - More than 64Kbps is not allowed
    - It implies “call blocking” due to lack of the modem lines
    - No guarantee in the backbone

Fundamental Principles for Data Application QoS

1. ***Data applications always try to use the bandwidth as much as possible.***
   - 8 **High-performance end systems**
   - 8 **High-speed access lines**
2. ***Neither bandwidth nor delay guarantees should be expected.***
   - 8 **Packet level metrics (e.g., packet loss prob. and packet delay) are never QoS metrics for data applications.**
   - 8 **It is inherently impossible to guarantee performance metrics in data networks.**
   - 8 **Network provider can never measure QoS**
3. ***Competed bandwidth should be fairly shared among active users.***
   - 8 **Transport layer protocols (TCP and UDP) is not fair.**
   - 8 **Network support is necessary by, e.g., active queue management at routers.**

Fundamental Theory for the Internet?

- *M/M/1 Paradigm (Queueing Theory) is Useful?*
  - Only provides packet queueing delay and loss probabilities at the node (router's buffer at one output port)



  - Data QoS is not queueing delay at the packet level
  - In Erlang loss formula, call blocking = user level QoS
- *Behavior at the Router?*
  - TCP is inherently a feedback system
- *User level QoS for Data?*
  - Application level QoS such as Web document transfer time
- *Control theory?*

Challenges for Network Provisioning

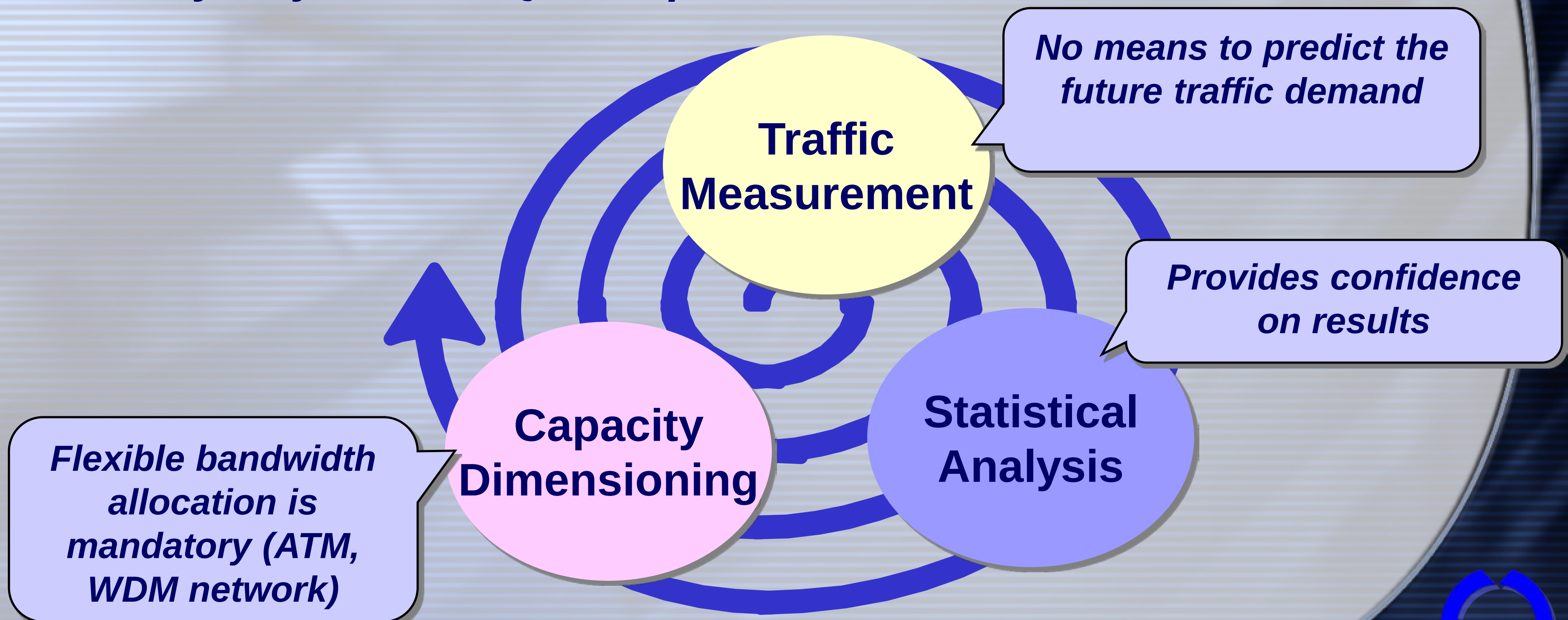
- ***New problems we have no experiences in data networks***
  - **What is QoS?**
  - **How can we measure QoS?**
  - **How can we charge for the service?**
  - **Can we predict the traffic characteristics in the era of information technology?**
  - **End-to-end performance can be known only by end users**
- ***QoS prediction at least at the network provisioning level***

Traffic Measurement Approaches

- ***Passive Measurements (OC3MON, OC12MON,,,)***
  - **Only provides point observations**
  - **Actual traffic demands cannot be known**
  - **QoS at the user level cannot be known**
- ***Active Measurements (Pchar, Netperf, bprobe,,,)***
  - **Provides end-to-end performance**
  - **Measurement itself may affect the results**
  - **Not directly related to network dimensioning (The Internet is connectionless!)**
- ***How can we pick up meaningful statistics?***
  - **Routing instability due to routing control**
  - **Segment retransmissions due to TCP error control**
  - **Rate adaptation by streaming media**
  - **Low utilization is because of**
    - **Congestion control?**
    - **Limited access speed of end users?**
    - **Low-power end host?**

Spiral Approach for Network Dimensioning

- ***Incremental network dimensioning by feedback loop is an only way to meet QoS requirements of users***



Traffic Measurement Approaches

- ***Traffic measurement projects***
  - "Cooperative Association for Internet Data Analysis," http://www.caida.org/
  - "Internet Performance Measurement and Analysis Project," http://www.merit.edu/ipma/
- ***How can we pick up meaningful statistics?***
  - Routing instability due to routing control
  - Segment retransmissions due to TCP error control
  - Rate adaptation by streaming media
  - Low utilization is because of
    - Congestion control?
    - Limited access speed of end users?
    - Low-power end host?

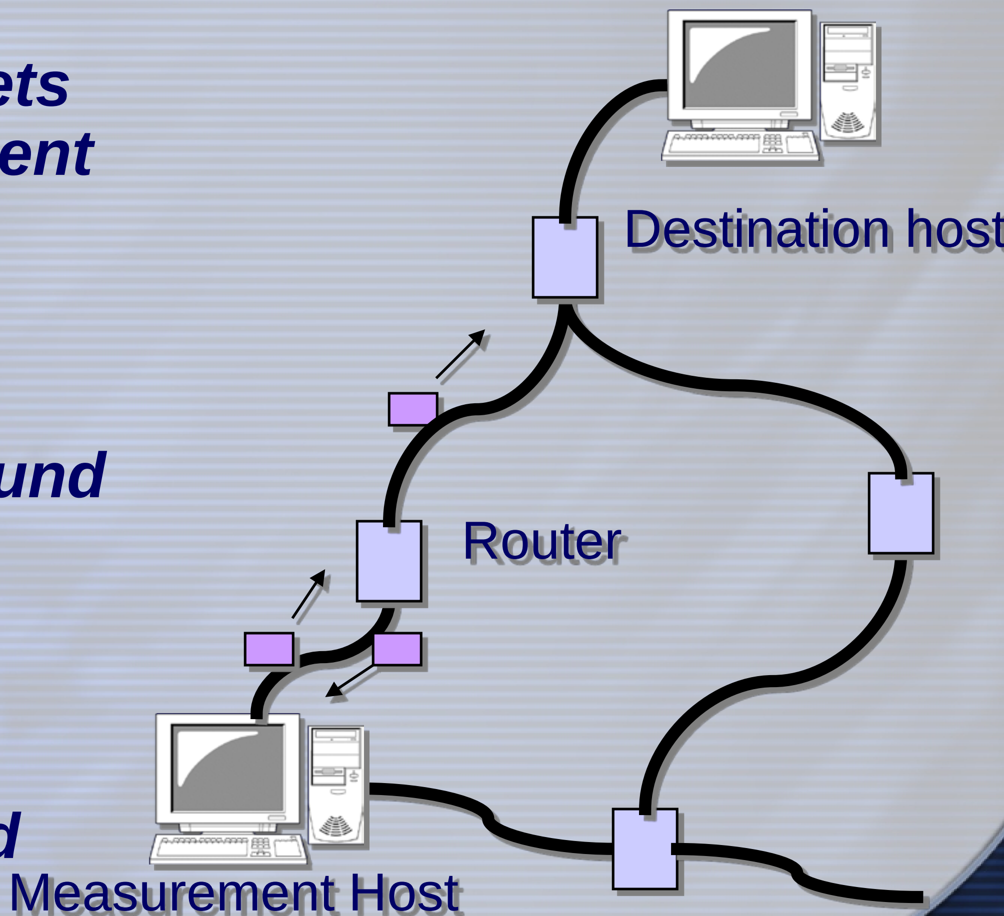
Passive and Active Traffic Measurements

- ***Passive Measurements***
  - OC3MON, OC12MON, . . .
  - Only provides point observations
  - Actual traffic demands cannot be known
  - QoS at the user level cannot be known
- ***Active Measurements***
  - Pchar, Netperf, bprobe, . . .
  - Provides end-to-end performance
  - Not directly related to network dimensioning (The Internet is connectionless!)

Active Bandwidth Measurement
Pathchar, Pchar

- ***Send probe packets from a measurement host***

- ***Measure RTT (Round Trip Time)***
- ***Estimate link bandwidth from relation between minimum RTT and packet size***



Relation between Minimum RTT and Packet Size

- *Minimum RTT between source host and nth hop router; proportional to packet size*

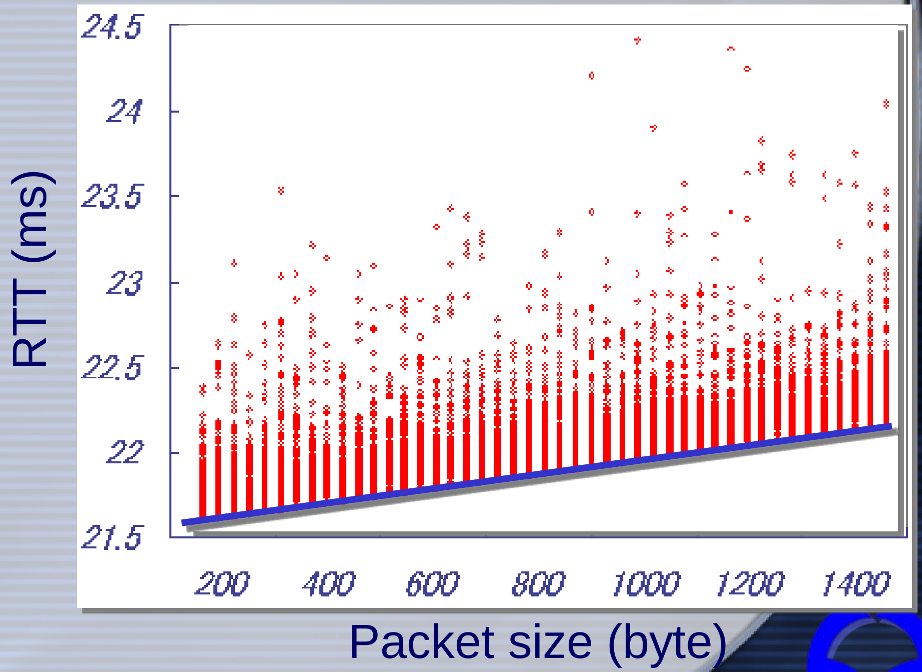
$$\min RTT_s = \sum_{j=1}^{n} \left(\frac{s + s_{ICMP}}{b_j} + 2p_j \right) + \sum_{i=1}^{n} f_i$$

$$= s \sum_{j=1}^{n} \frac{1}{b_j} + \alpha$$

$\min RTT_s$: Minimum RTT

s : Packet Size

b_j : Bandwidth of link **n**



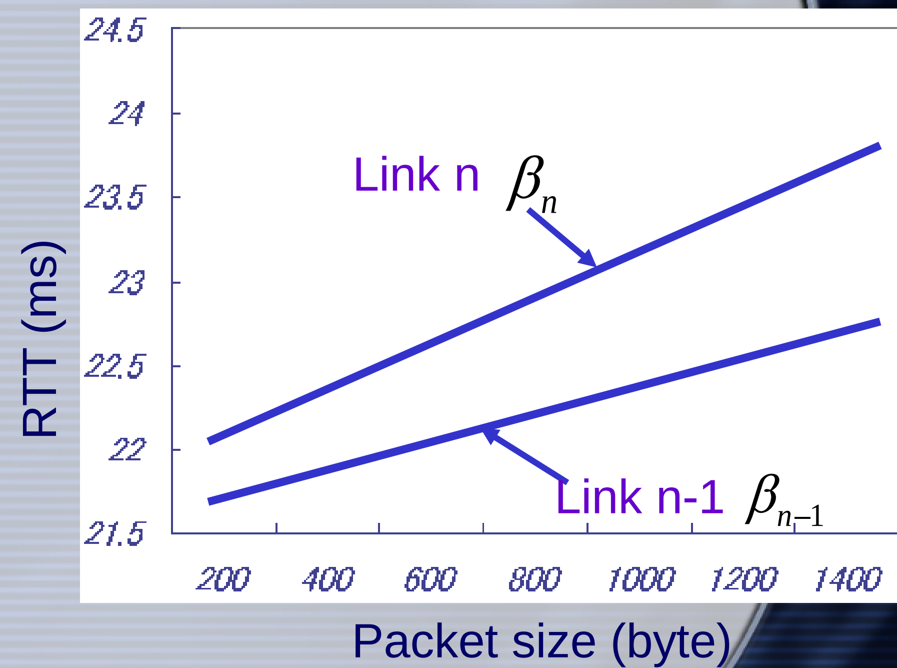
Bandwidth Estimation in Pathchar and Pchar

- ***Slope of line is determined by minimum RTTs between n*th *router and source host***

$$\sum_{j=1}^{n} \frac{1}{b_j} = \beta_n$$

- ***Estimate the slope of line using the linear least square fitting method***
- ***Determine the bandwidth of n*th *link***

$$b_n = \frac{1}{\beta_n - \beta_{n-1}}$$



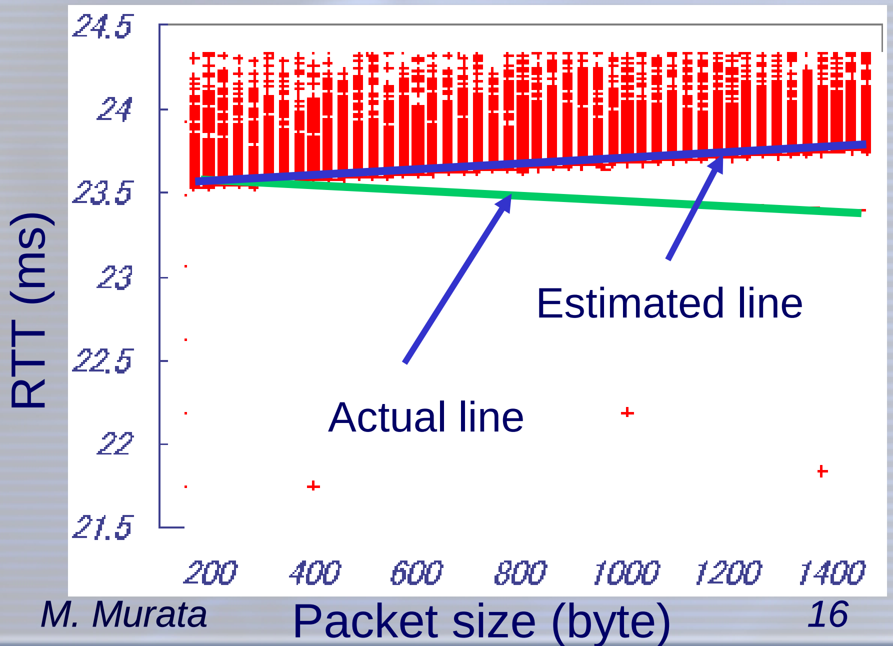
Several Enhancements

1. ***To cope with route alternation***
   - **Clustering approach**
2. ***To give statistical confidence***
   - **Confidence intervals against the results**
3. ***To pose no assumption on the distribution of measurement errors***
   - **Nonparametric approach**
4. ***To reduce the measurement overhead***
   - **Dynamically controls the measurement intervals; stops the measurement when the sufficient confidence interval is obtained**

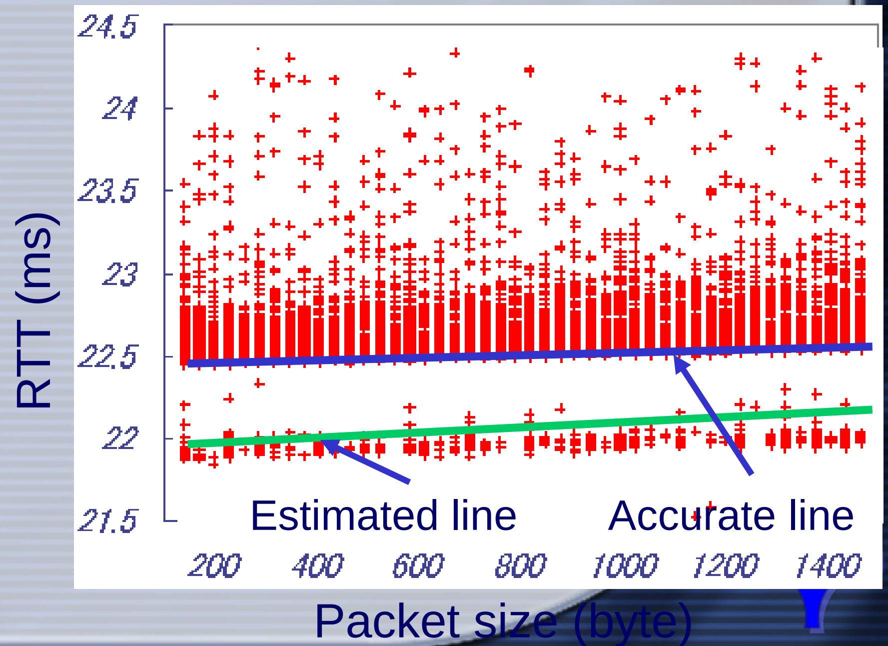
Measurement Errors

- *Assuming the errors of minimum RTT follow the normal distribution*
  - ∑ *Sensitive to exceptional errors*

A few exceptional large errors



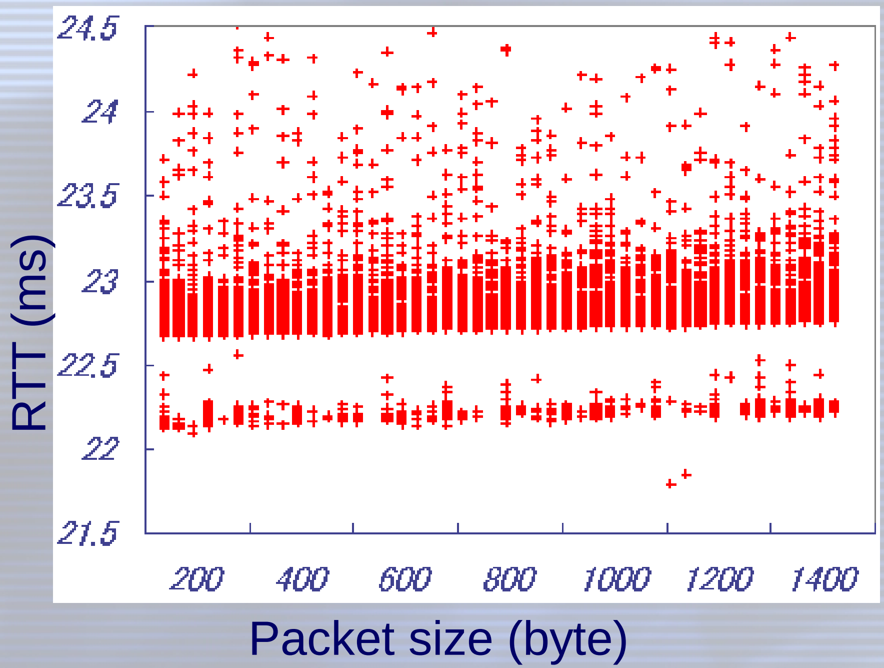
"Errors" caused by route alternation



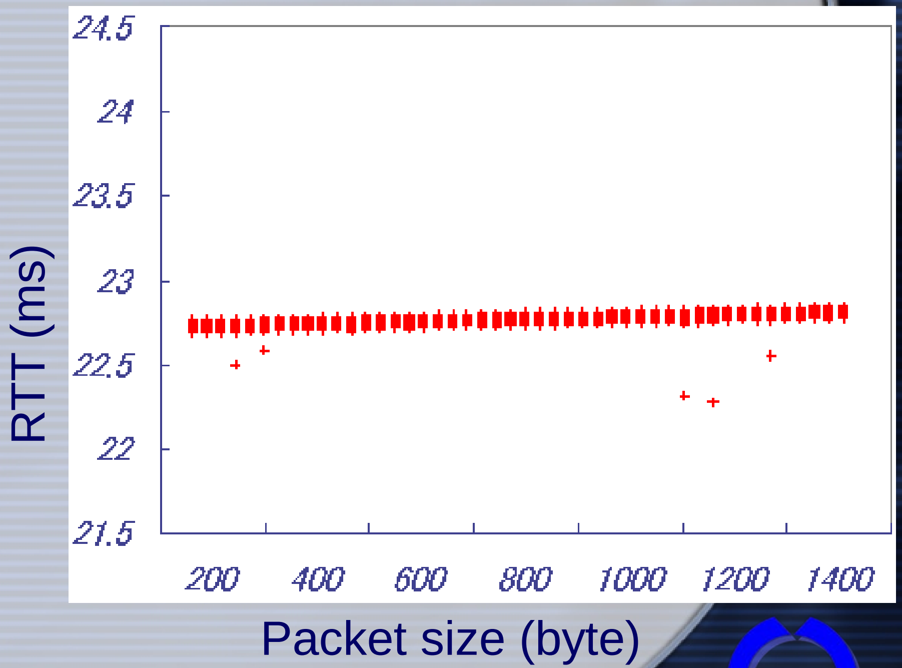
Clustering

- *Remove "errors" caused by route alternation*

Errors due to router alternation

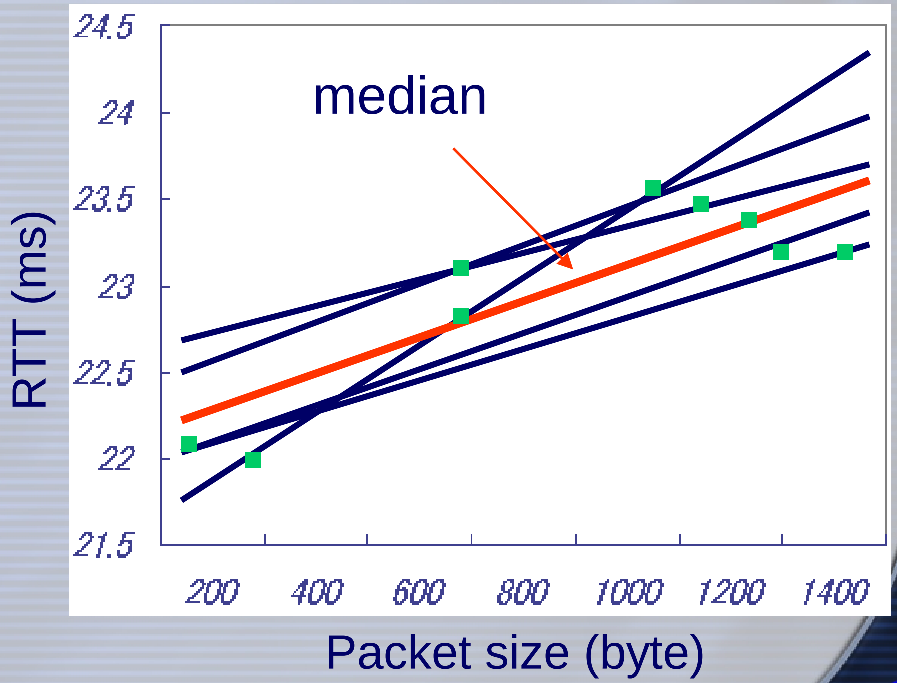


After clustering



Nonparametric Estimation

- ***Measure minimum RTTs***
- ***Choose every combination of two plots, and calculate slopes***
- ***Adopt the median of slopes as a proper one***
- ***Independent of error distributions***



Adaptive Control for Measurement Intervals

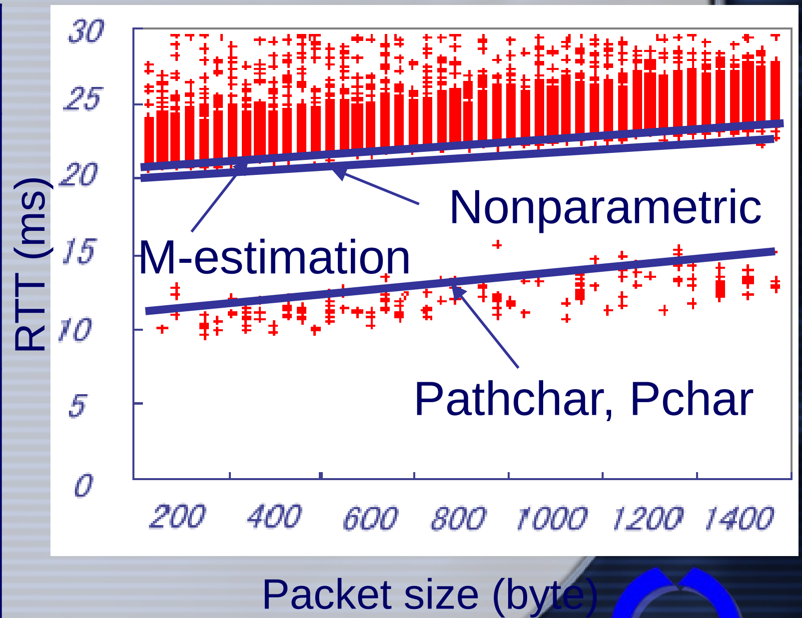
- ***Control the number of probe packets***
  1. **Send the fixed number of packets**
  2. **Calculate the confidence interval**
  3. **Iterate sending an additional set of packets until the confidence interval sufficiently becomes narrow**

∑ ***Can reduce the measurement period and the number of packets with desired confidence intervals***

Estimation Results against Route Alternation

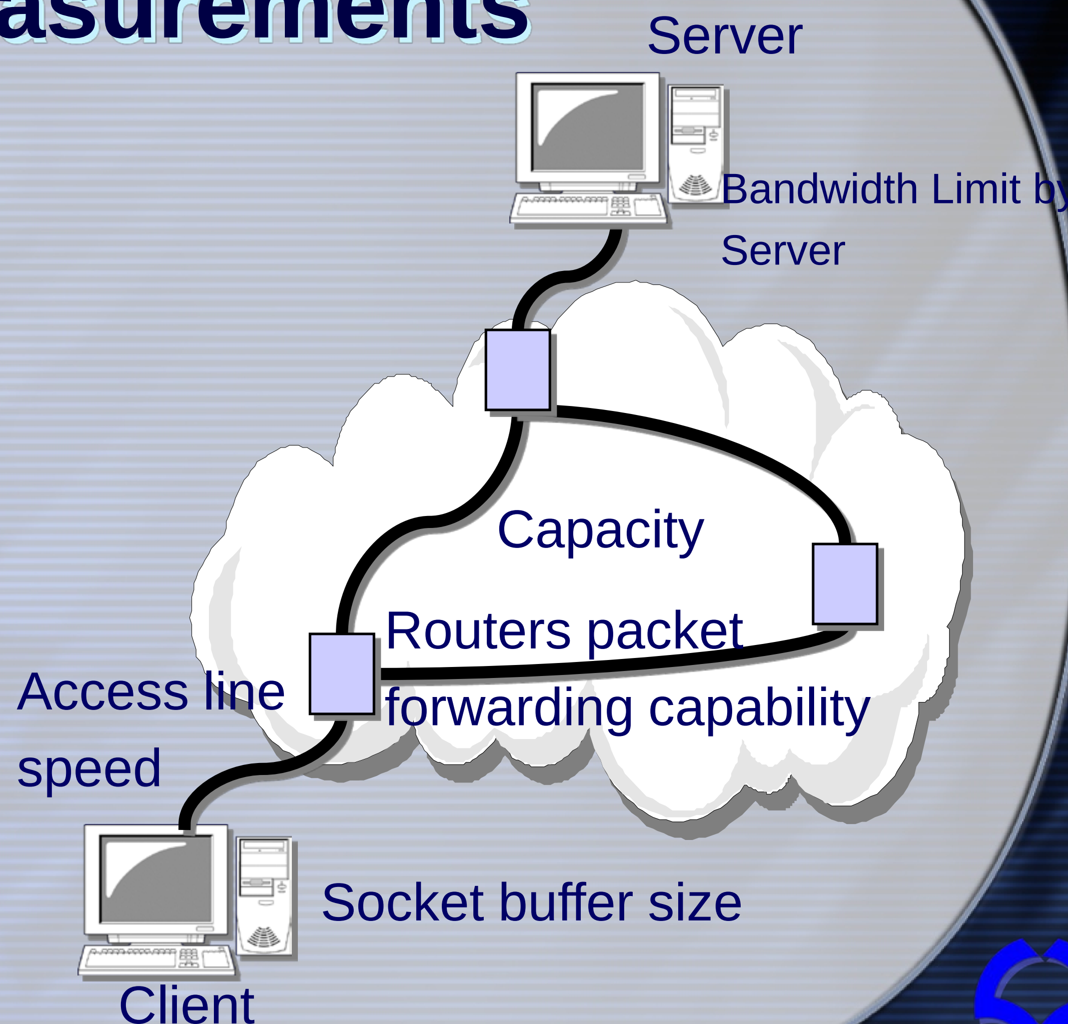
- ***Effects of Clustering***
  - **Can estimate correct line by the proposed method**

Bandwidth	Method	Estimation Results	The # of packets
10 Mbps	*Pathchar* *M-estimation* *Wilcoxon* *Kendall*	*-22.6* *10.1<12.4<16.1* *16.6<17.0<24.1* *14.2<17.0<25.3*	*200* *200* *200* *200*
12 Mbps	*Pathchar* *M-estimation* *Wilcoxon* *Kendall*	*8.25* *9.79<9.94<10.1* *13.3<13.8<14.4* *13.6<13.8<14.1*	*200* *20* *90* *90*



Bottleneck Identification by Traffic Measurements

- ***Network dimensioning by the carriers solely is impossible***
- ***End users (or edge routers) only can know QoS level***
  - **Bottleneck identification**
  - **Feedback to capacity updates**



Comparisons between Theoretical and Measured TCP Throughputs

- *Expected Value of TCP Connection's Window Size E[W]*

$$E[W] = \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2}$$

If $E[W] \geq W_{max}$, the socket buffer at the receiver is bottleneck

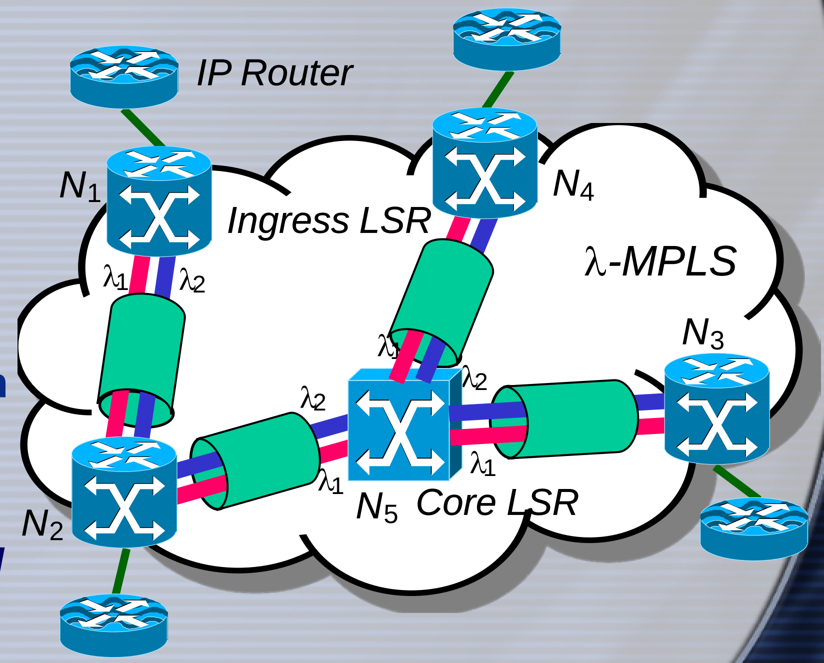
- *Expected TCP Throughput can be determined by RTT, RTT、Packet Loss Probability, Maximum Widnow Size, and Timeout Values*

$$\frac{\frac{1-p}{p} + W_{max} + \hat{Q}(W_{max})\frac{1}{1-p}}{RTT\left(\frac{b}{8}W_{max} + \frac{1-p}{pW_{max}} + 2\right) + \hat{Q}(W_{max})T_0\frac{f(p)}{1-p}} \qquad E[W] \geq W_{max}$$

- **Difference indicates the bottleneck within the network**

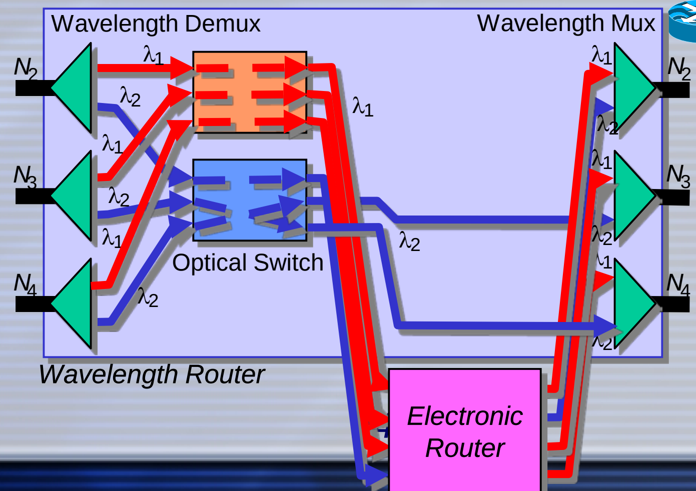
Past Researches on WDM Networks

- ***Routing and Wavelength Assignment (RWA) Problem***
  - **Static assignment**
    - **Optimization problem for given traffic demand**
  - **Dynamic assignment**
    - **Natural extension of call routing**
    - **Call blocking is primary concern**
    - **No wavelength conversion makes the problem difficult**
- ***Optical Packet Switches for ATM***
  - **Fixed packets and synchronous transmission**



Logical Topology by Wavelength Routing

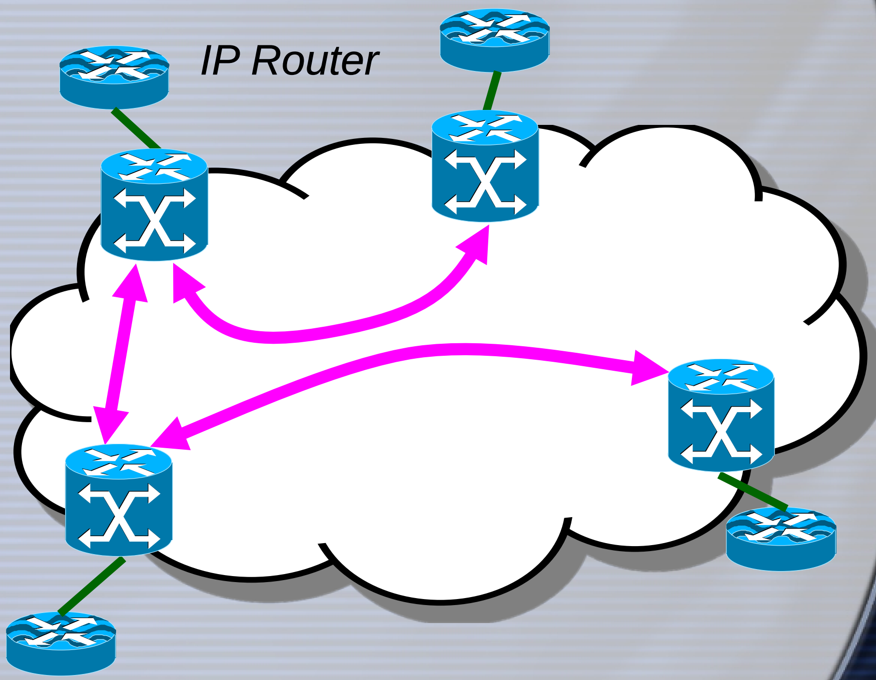
- *Physical Topology*
- *Logical Topology*



M. Murata

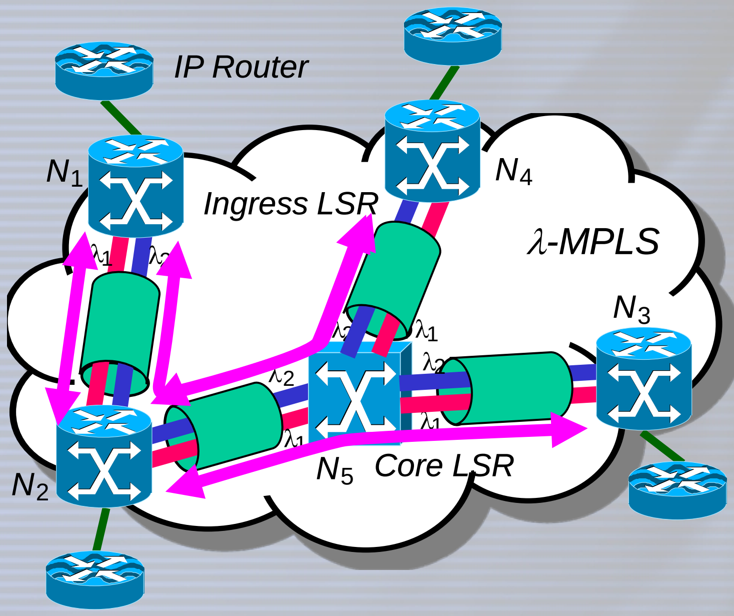
Logical View Provided to IP

- ***Redundant Network with Large Degrees***
  - ***Smaller number of hop-counts between end-nodes***
  - ***Decrease load for packet forwarding at the router***
  - ***Relief bottleneck at the router***



Incremental Network Dimensioning

- ***Initial Step***
  - **Topology design for given traffic demand**
- ***Adjustment Step***
  - **Adds a wavelength path based on measurements of traffic by passive approach, and end-users' quality by active approach**
  - **Only adds/deletes wavelength paths; not re-design an entire topology**
  - **Backup paths can be used for assigning the primary paths**
- ***Entirely Re-design Step***
  - **Topology re-design by considering an effective usage of entire wavelengths**
  - **Only a single path is changed at a time for traffic continuity**



Re-partitioning of Network Functionalities

- ***Too much rely on the end host***
  - **Congestion control by TCP**
    - **Congestion control is a network function**
  - **Inhibits the fair service**
    - **Host intentionally or unintentionally does not perform adequate congestion control (S/W bug, code change)**
    - **Obstacles against charged service**
- ***What should be reallocated to the network?***
  - **Flow control, error control, *congestion control*, routing**
  - **RED, DRR, ECN, diff-serv, int-serv (RSVP), policy routing**
  - **We must remember too much revolution lose the merit of Internet.**

Fairness among Connections

- ***Data application is always greedy***
- ***We cannot rely on TCP for fairshare of the network resources***
  - **Short-term unfainess due to window size throttle**
  - **Even long-term unfairness due to different RTT, bandwidth**

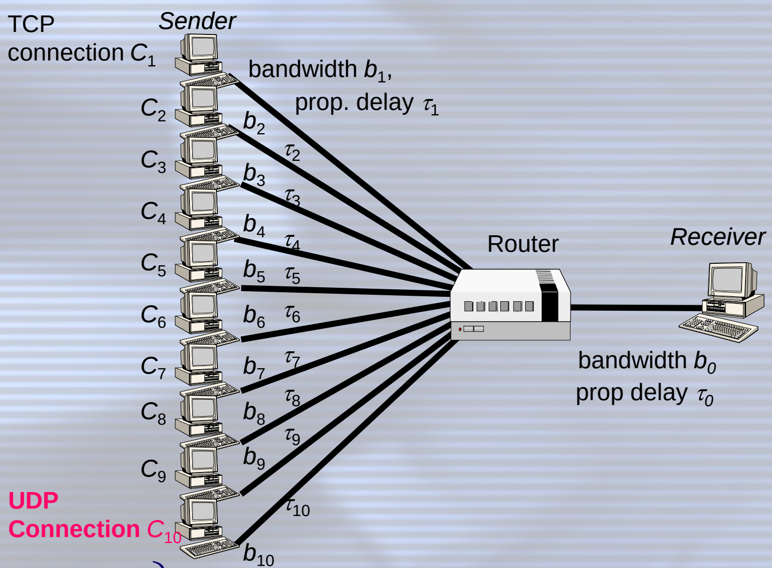
∑ ***Fair treatment at the router; RED, DRR***

∑ ***Fairshare between real-time application and data application***
  - **TCP-Friendly congestion control**

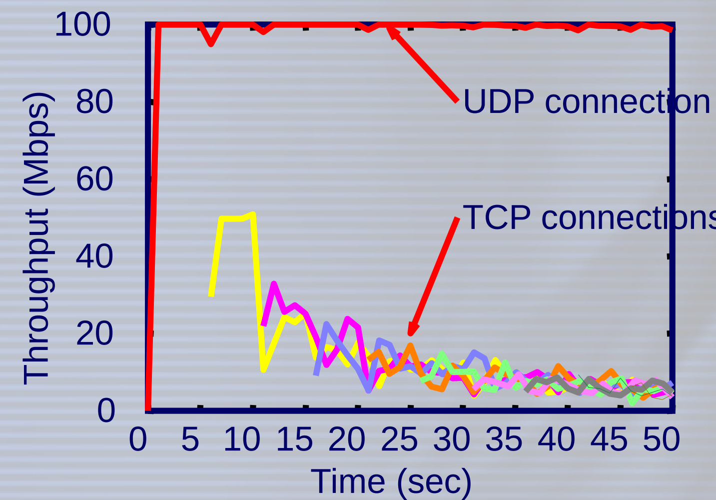
Layer-Four Switching Techniques for Fairshare

- *Flow classification in stateless?*
- *Maintain per-flow information (not per-queueing)*
  - FRED
    - Counts the # of arriving/departing packets of each active flow, and calculates its buffer occupancy, which is used to differentiate RED's packet dropping probabilities.
  - Stabilized RED
  - Core Stateless Fair Queueing
    - At the edge router, calculates the rate of the flow and put it in the packet header. Core router determines to accept the packet according to the fairshare rate of flows by the weight obtained from the packet header. DRR-like scheduling can be used, but no need to maintain per-flow queueing.

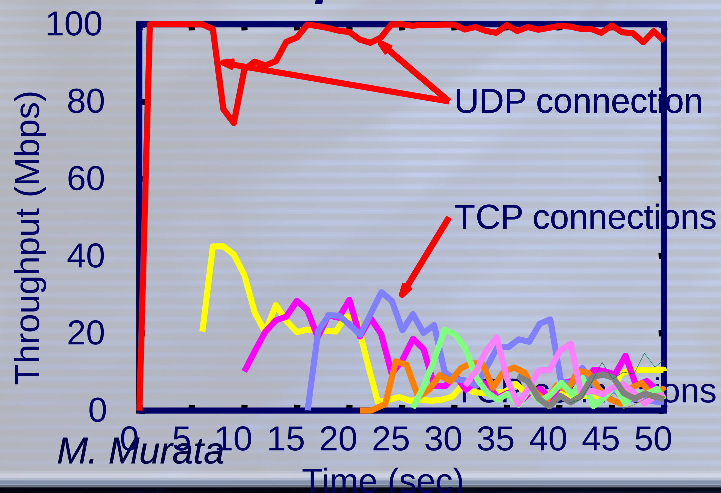
Effects of FRED



RED Router



Drop-Tail Router



FRED Router