

Shanghai, P.R. China  
November 4<sup>th</sup>, 2003

Hakone, Japan  
November 7<sup>th</sup>, 2003

# Backbone Traffic Management

Thomas Telkamp  
telkamp@gblix.net

# Network Build

# Network Configuration

Capacity Planning

Routing

Scheduling, etc...

Events &  
State

Statistics

Simulation &  
Optimization

BW Allocation

Per class QoS criteria

Characterization

Demands

Time-scale behavior

Measurements and/or Inference

TRAFFIC

Data

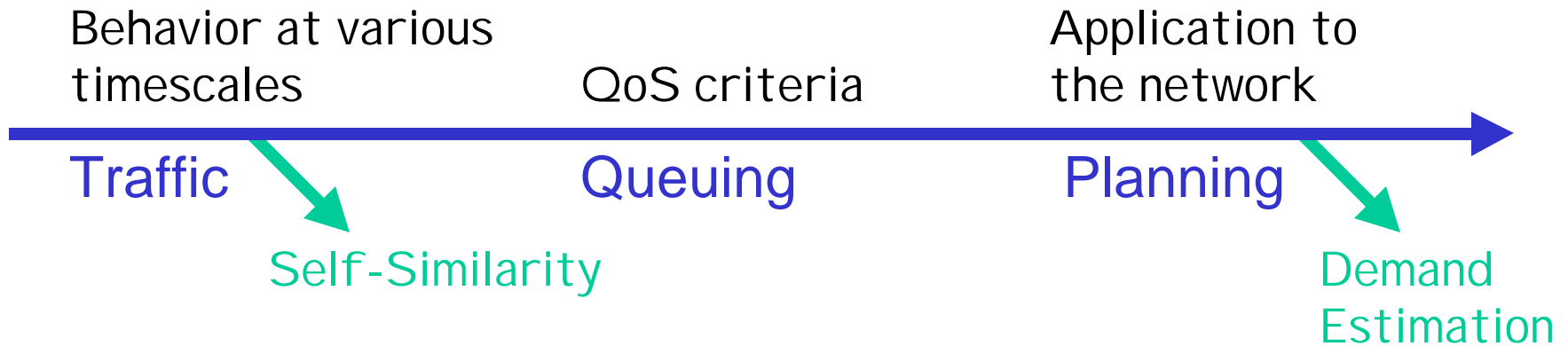
USERS

Voice

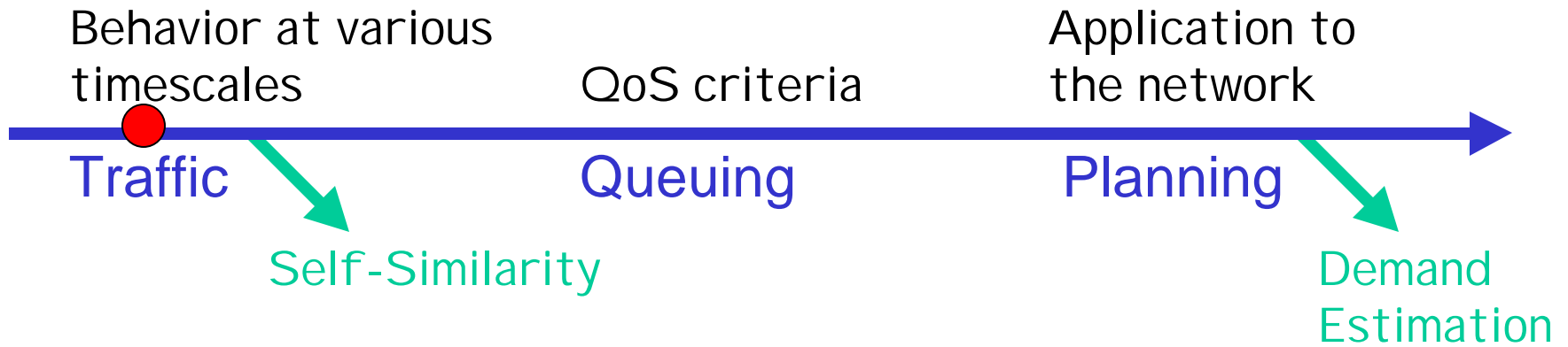
TDM to VoIP

Silence Suppression

# Agenda

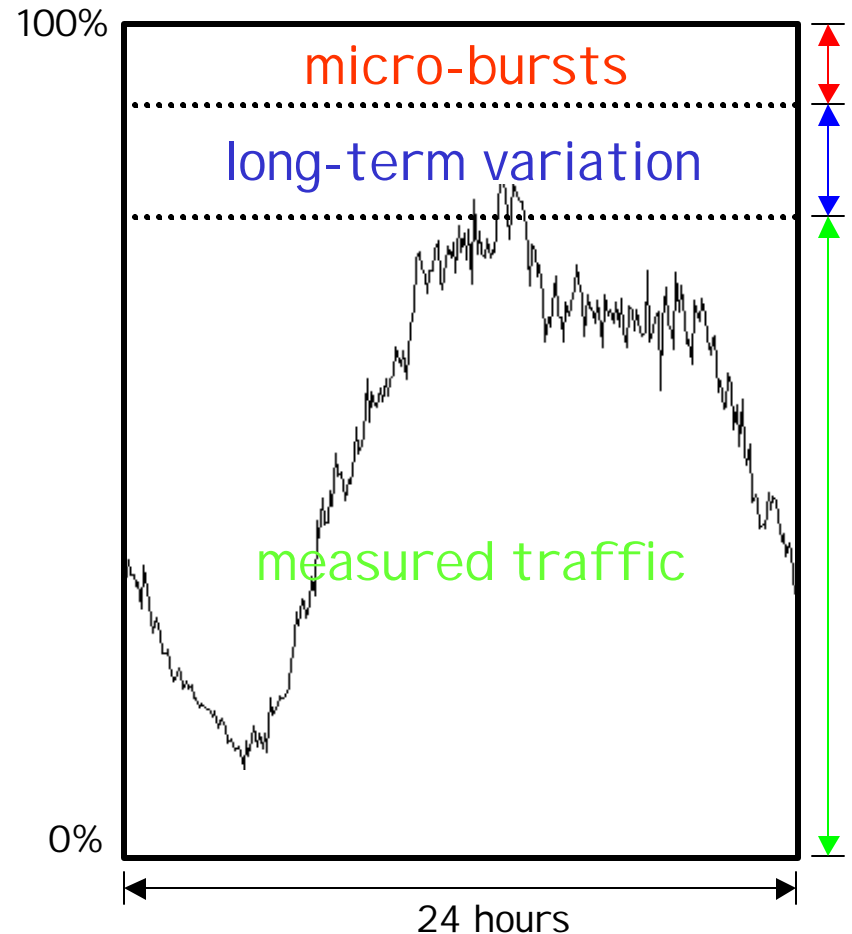


# TRAFFIC



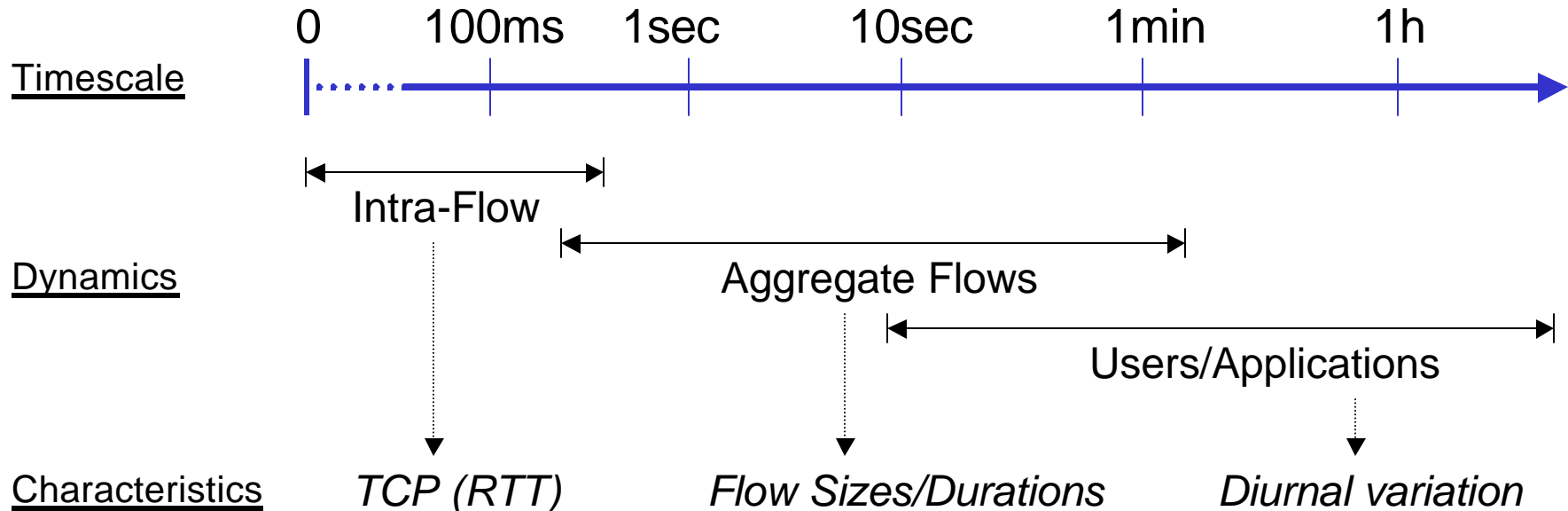
# Demand Characterization

- Long-Term
  - Measured Traffic
  - E.g. P95 (day/week)
  - “unforeseen” events and growth
- Short-Term
  - Critical scale for queuing
  - Determine over-provisioning factor that will prevent queue buildup against micro-bursts



# Relevant Timescales

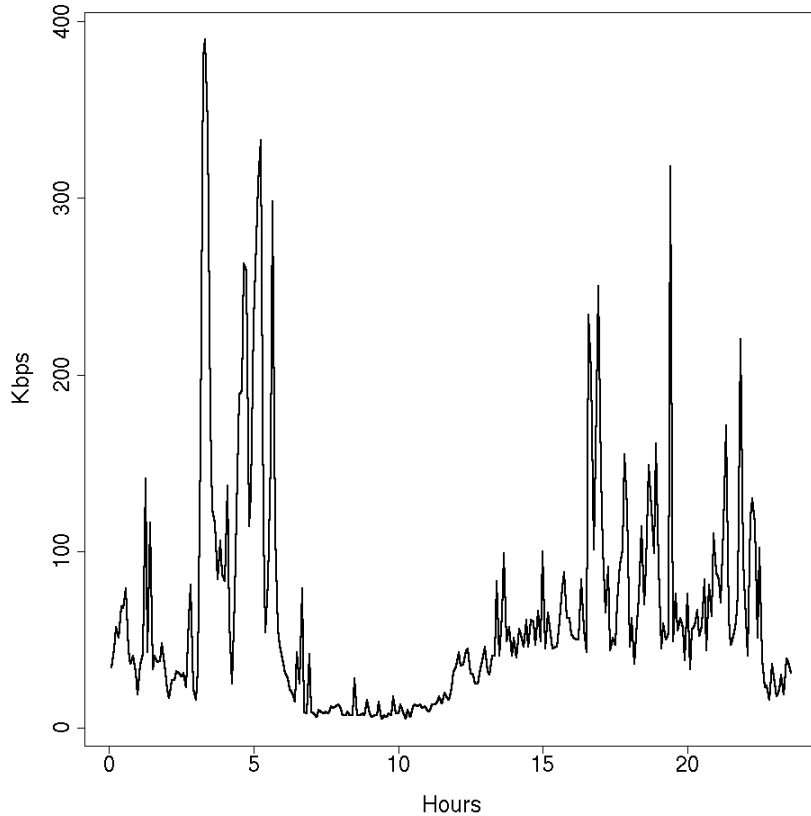
- Long-Term: > 5 minutes
- Short-Term: < 5 minutes



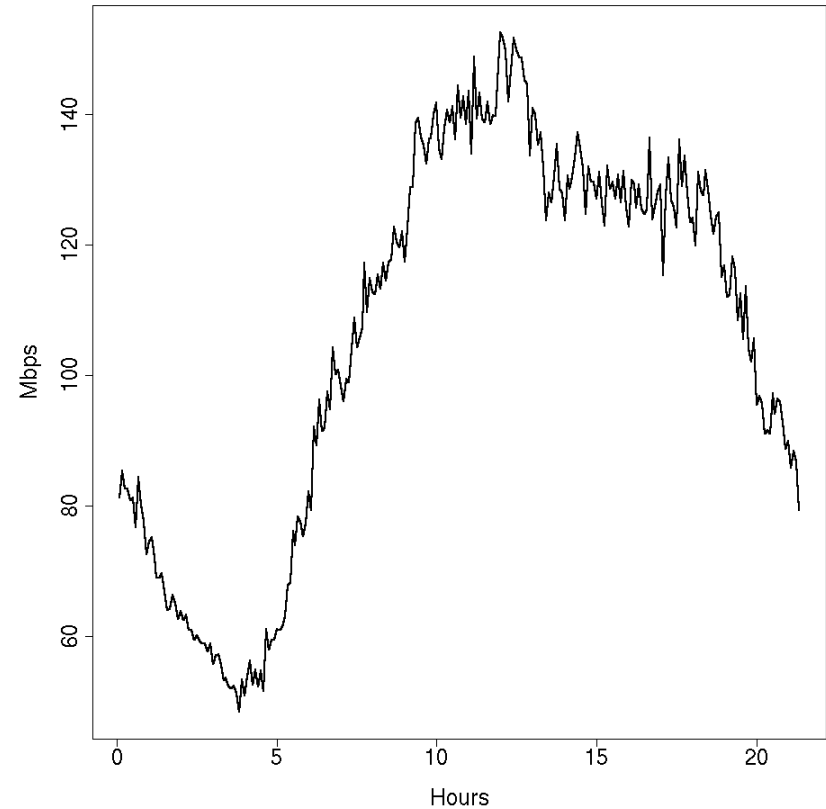
# Long-term Traffic Characterization

- Investigate burstiness of 5-min measurements over days/weeks
- Bursty traffic: peaks are very large compared to average
  - I.e. the distribution is Heavy-Tailed
  - Mean and 95-percentile do not represent the traffic very well
  - Planning becomes very difficult

# High- vs Low-Bandwidth Demands



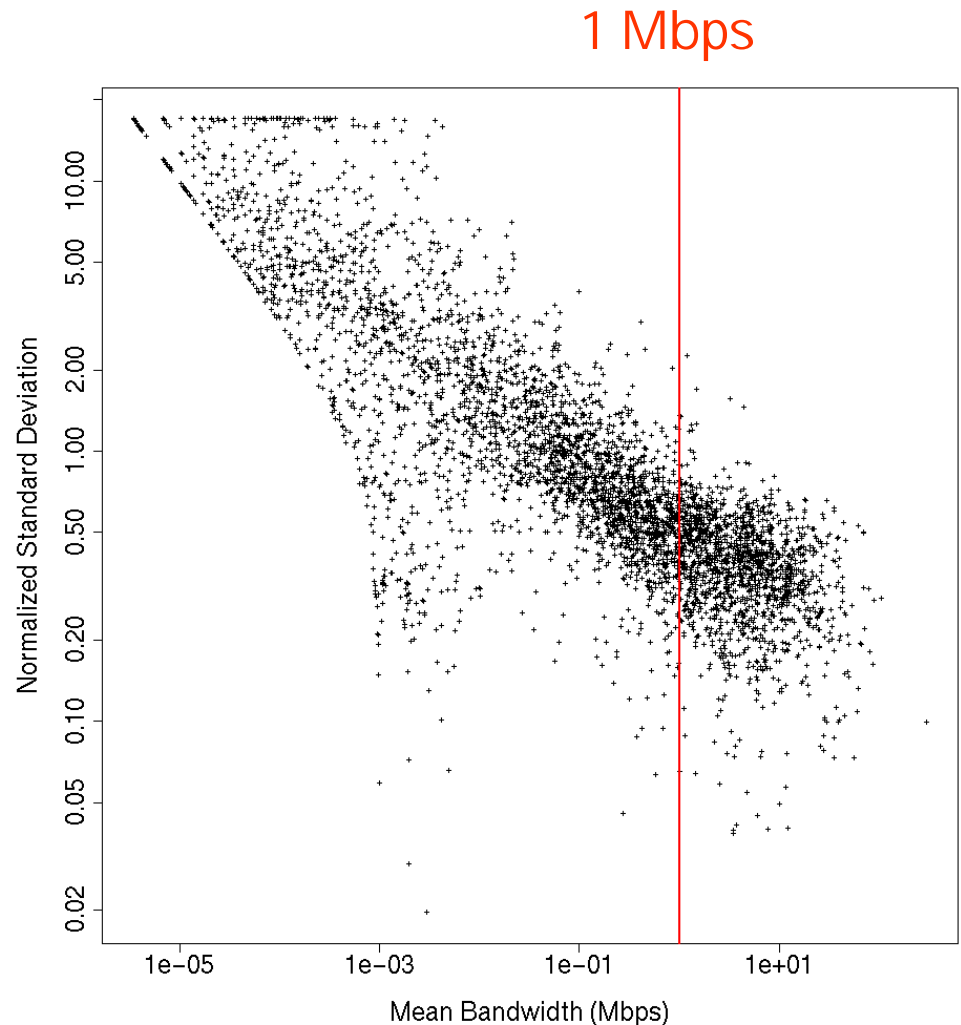
Cleveland -> Denver  
Mean=64Kbps, Max=380Kbps  
P95=201Kbps, Std. dev.=66Kbps  
 $\alpha=1.8$  (tail index)



Washington D.C. -> Copenhagen  
Mean=106Mbps, Max=152Mbps  
P95=144Mbps, Std. dev.=30Mbps  
 $\alpha=21$  (tail index)

# Variance vs Bandwidth

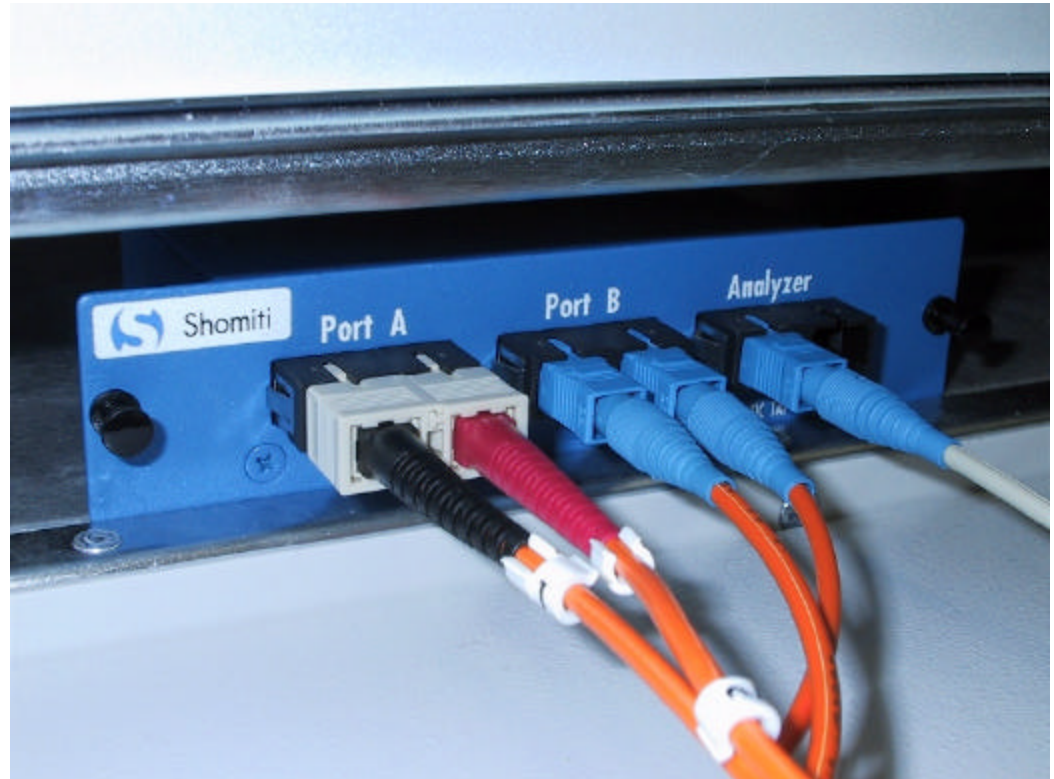
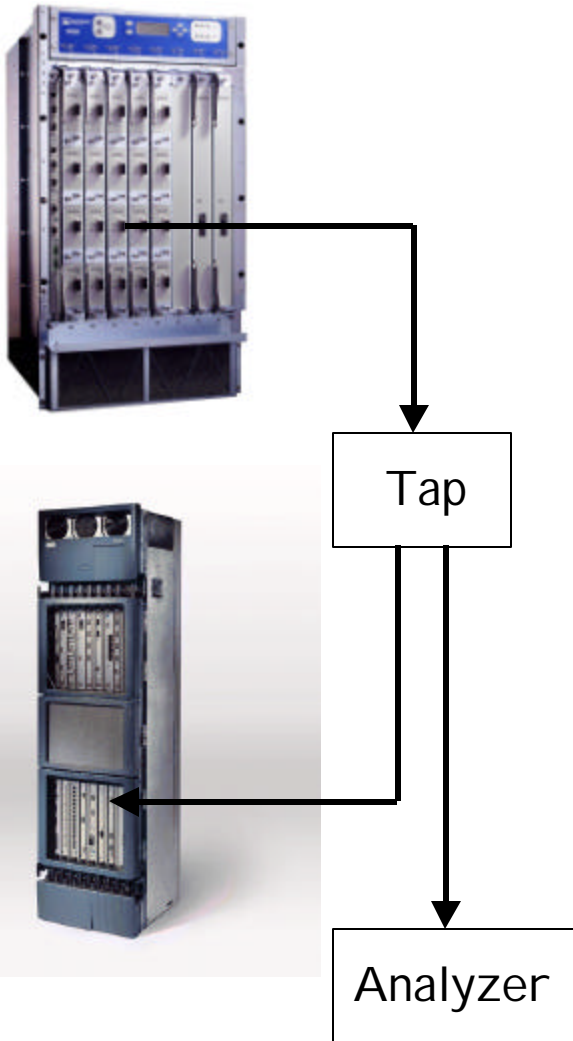
- Around 8000 demands between core routers
- Relative variance decreases with increasing bandwidth [5]
- High-bandwidth demands seem well-behaved
- 97% of traffic is carried by the demands larger than 1 Mbps (20% of the demands!)



# Short-term Traffic Characterization

- Investigate burstiness within 5-min intervals
- Measurements at critical timescale for queuing, like 1ms or 10ms
- Only at specific locations
  - Complex setup
  - A lot of data
- Analyze statistical properties

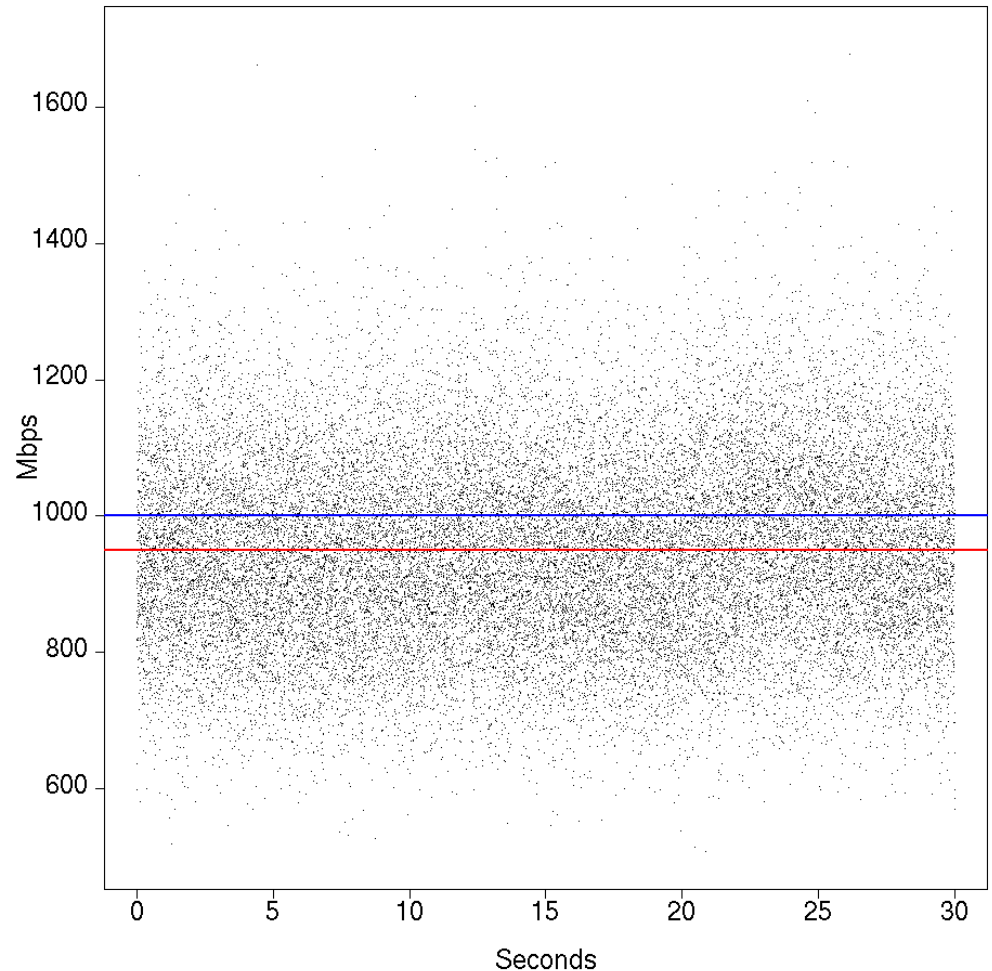
# Fiber Tap (Gigabit Ethernet)



# Raw Results

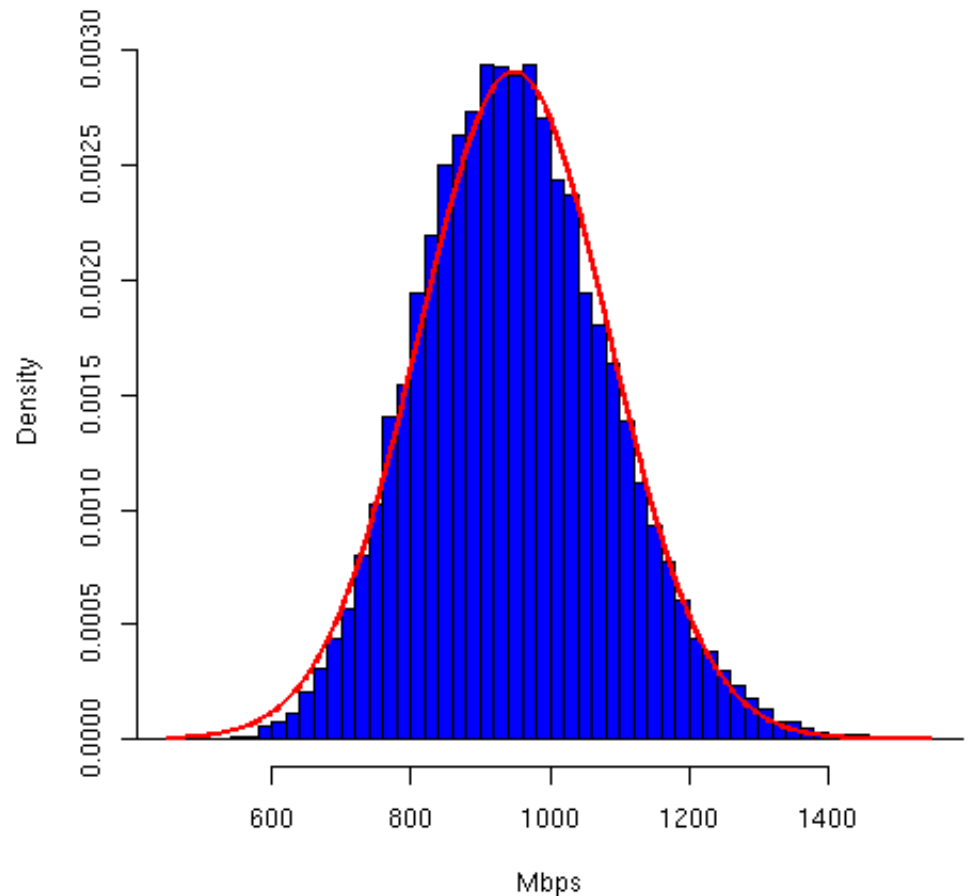
## 30 sec of data, 1ms scale

- Mean = 950 Mbps
- Max. = 2033 Mbps
- Min. = 509 Mbps
  
- 95-percentile: 1183 Mbps
- 5-percentile: 737 Mbps
  
- (around 250 packets per 1ms interval)



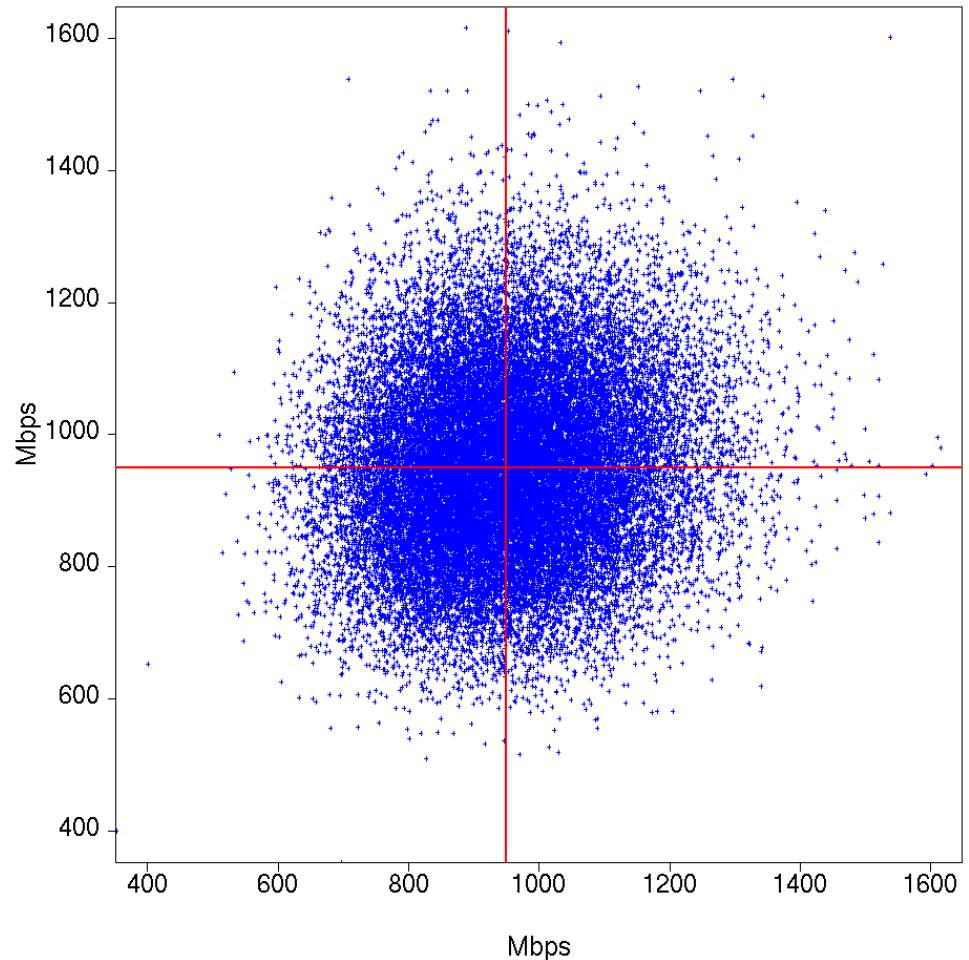
# Traffic Distribution Histogram (1ms scale)

- Fits normal probability distribution very well (Std. dev. = 138 Mbps)
- No Heavy-Tails
- Suggests small overprovisioning factor



# Autocorrelation Lag Plot (1ms scale)

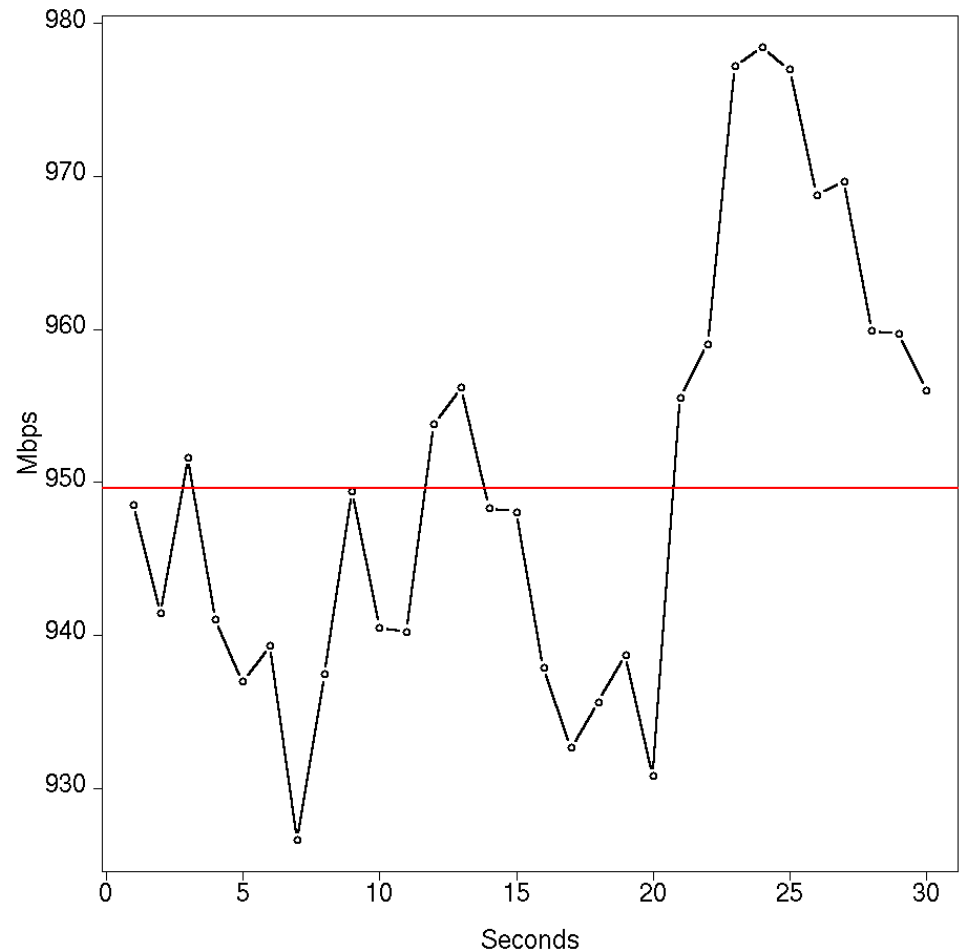
- Scatterplot for consecutive samples
- Are periods of high usage followed by other periods of high usage?
- Autocorrelation at 1ms is 0.13 (=uncorrelated)



# Utilization

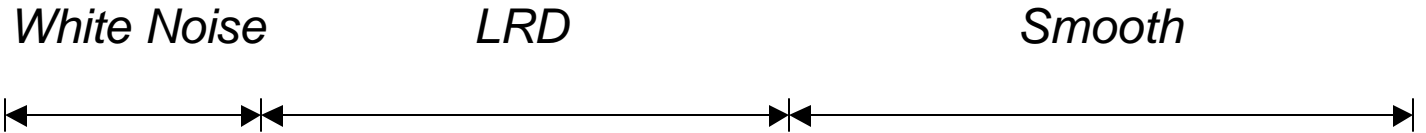
## 30 sec of data, 1 sec scale

- Mean = 950 Mbps
- Max. = 978 Mbps
- Min. = 927 Mbps
  
- Clearly longer derivations from the mean
  
- High autocorrelation at 1 sec. (0.68)



# Relevant Timescales

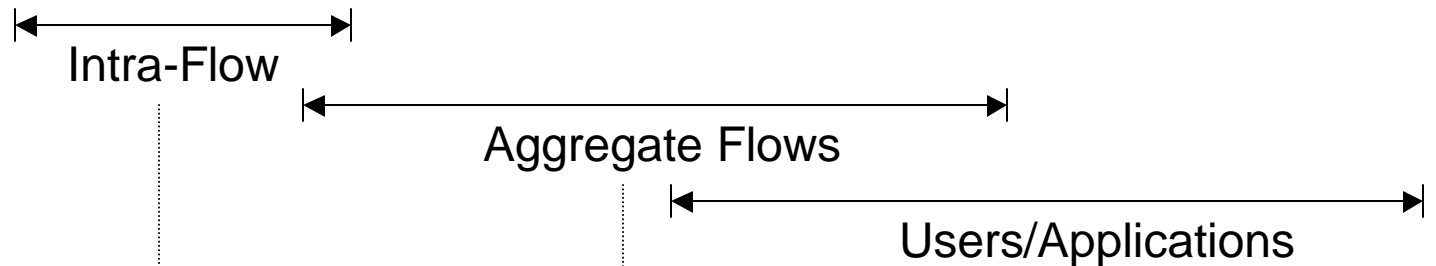
Behavior



Timescale



Dynamics



Characteristics

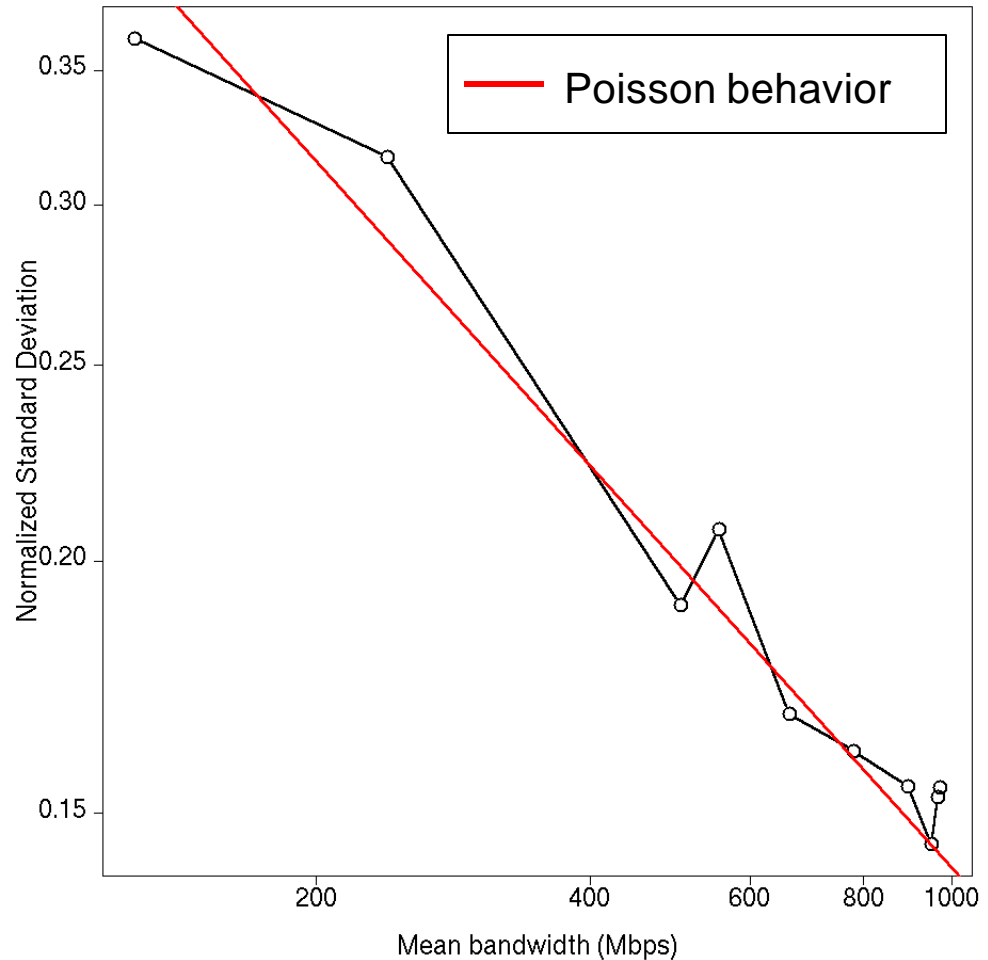
*TCP (<RTT)*      *Flow Sizes/Durations*      *Diurnal variation*

# Traffic Behavior

- < 150ms: Gaussian White Noise
  - I.I.D, Normal distribution, no correlation
  - Short-Range Dependent ( $H=0.5$ )
- 150ms – 30sec: Long-Range Dependence
  - Non-summable autocorrelation
  - Statistically Self-Similar
  - $X =_d m^{1-H} X^{(m)}$ , "Clustering", Fractional Brownian Motion (fBM)
- > 30 sec: Smooth
  - Degenerate case: autocorrelation = 1
  - Almost constant, diurnal variation

# Variance vs. Mean Bandwidth

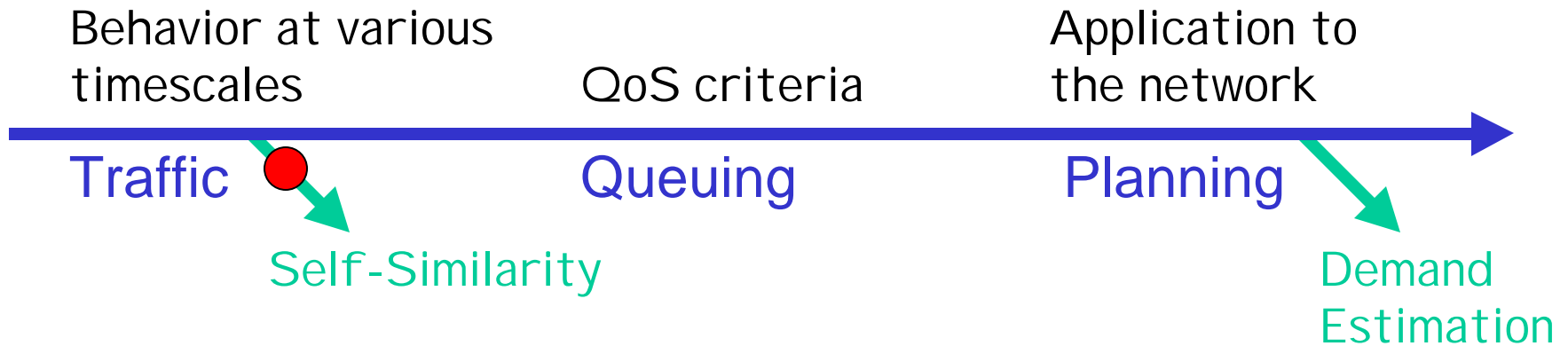
- Standard deviation of bandwidth increases with the root of the average
- Relative variance decreases with increasing bandwidth [5]
- $\left[ \frac{\sigma}{\bar{i}} \sim 1/\sqrt{\bar{i}} \right]$



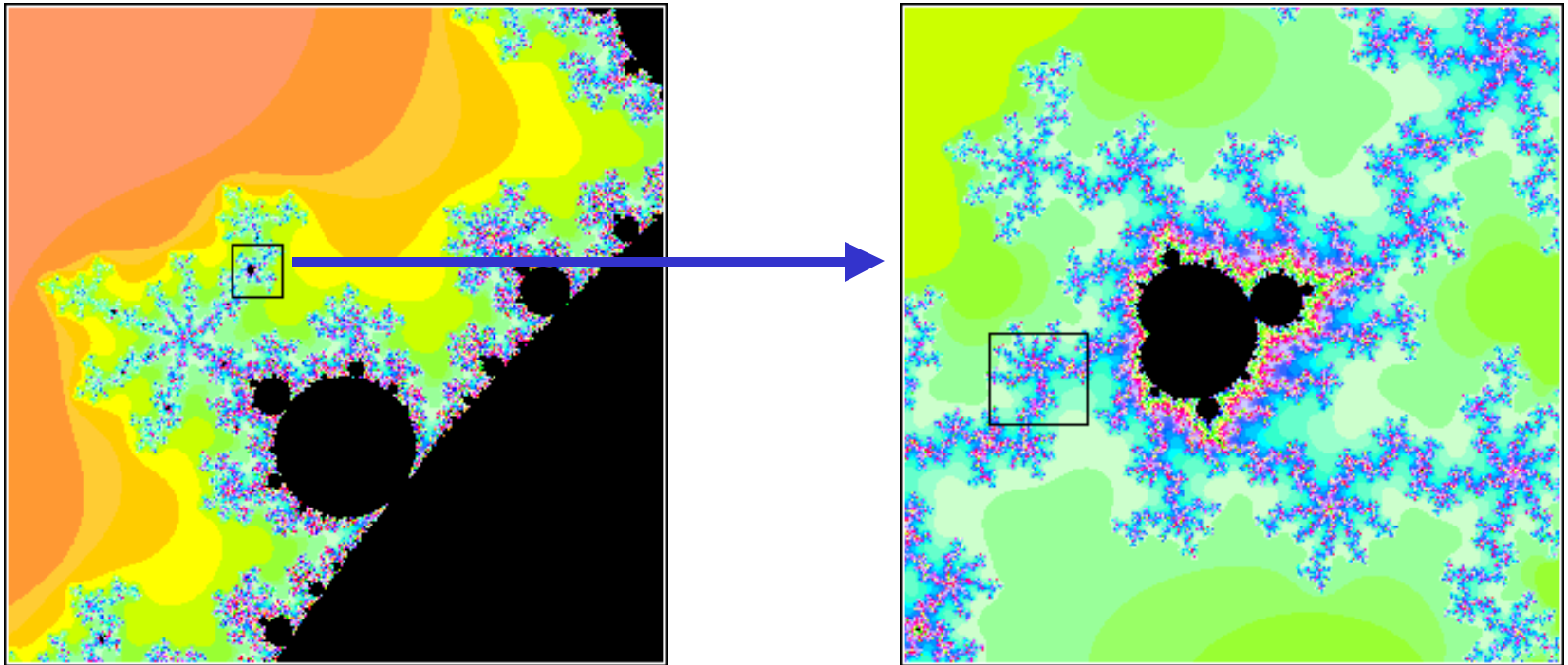
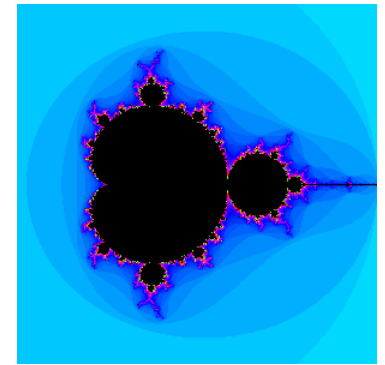
# Traffic: Summary

- Traffic aggregates “smooth” with increasing traffic rate
  - ‘Vertical’ aggregation
- Traffic does not aggregate well with increasing timescales (Long-Range Dependence)
  - ‘Horizontal’ aggregation
- Long-term well-behaved traffic

# Self-Similarity

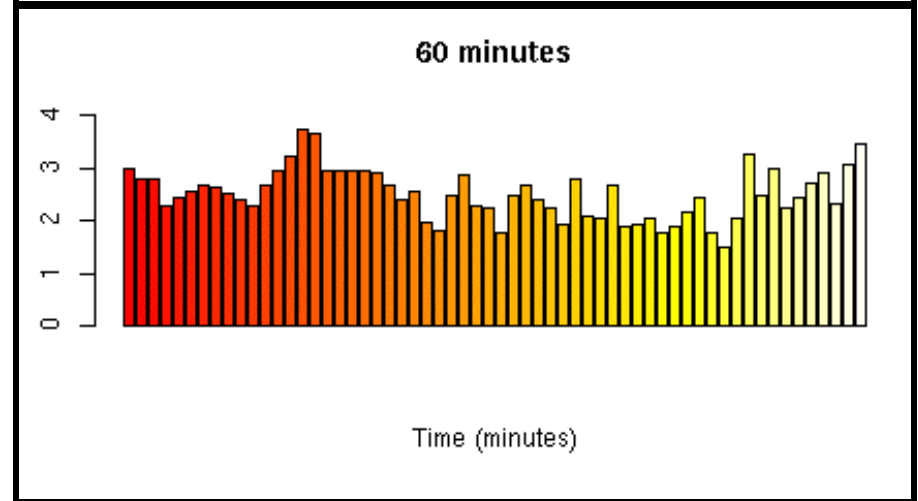
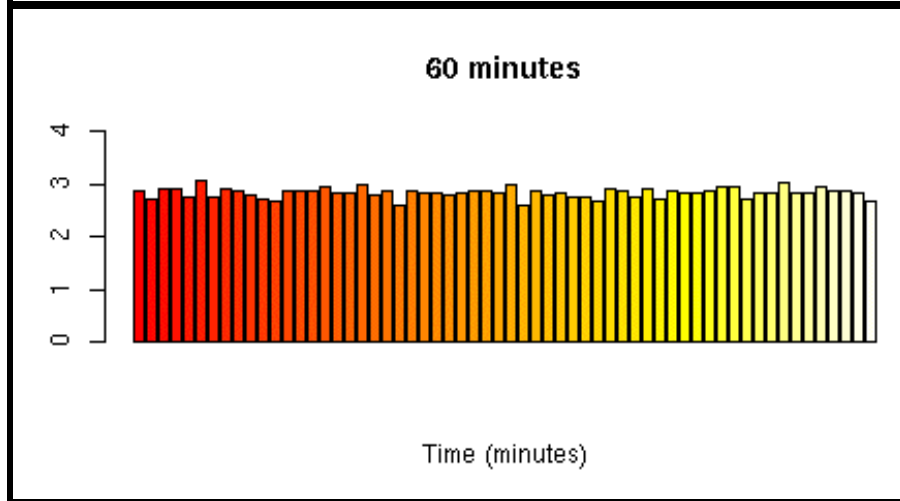
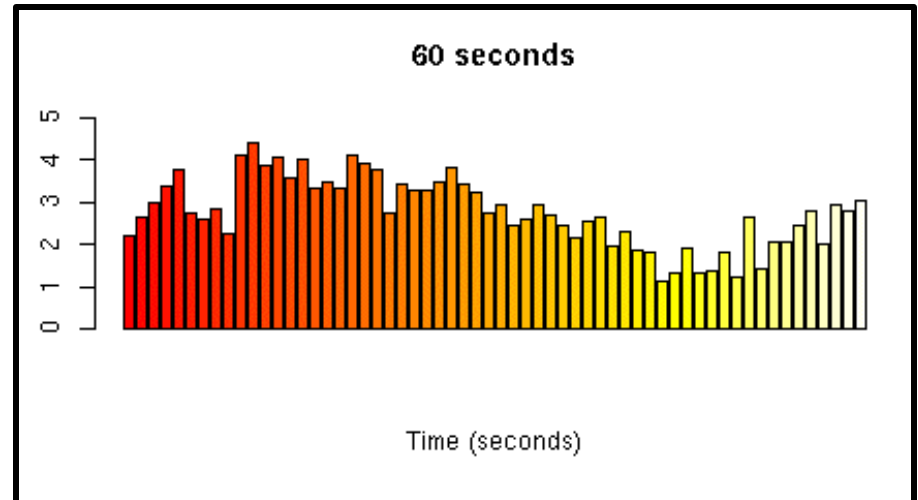
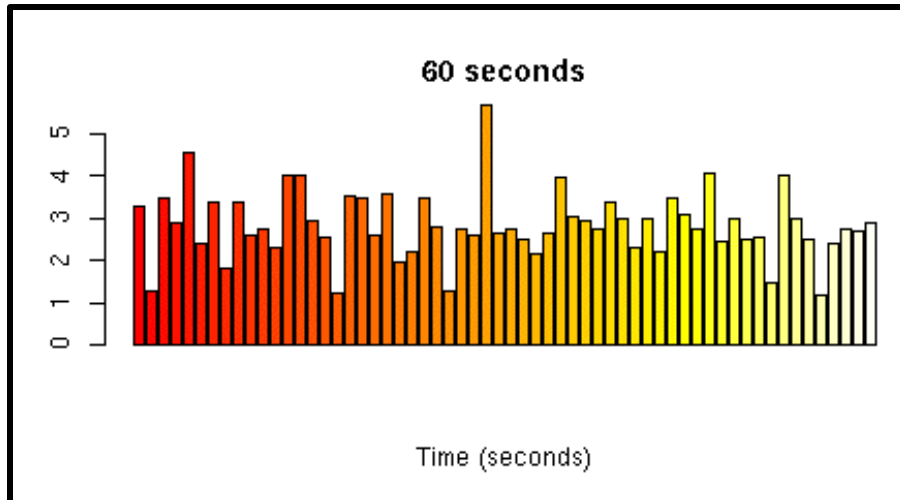


# Mandelbrot Fractal



A self-similar object is one whose component parts resemble the whole: Invariant under changes of scale

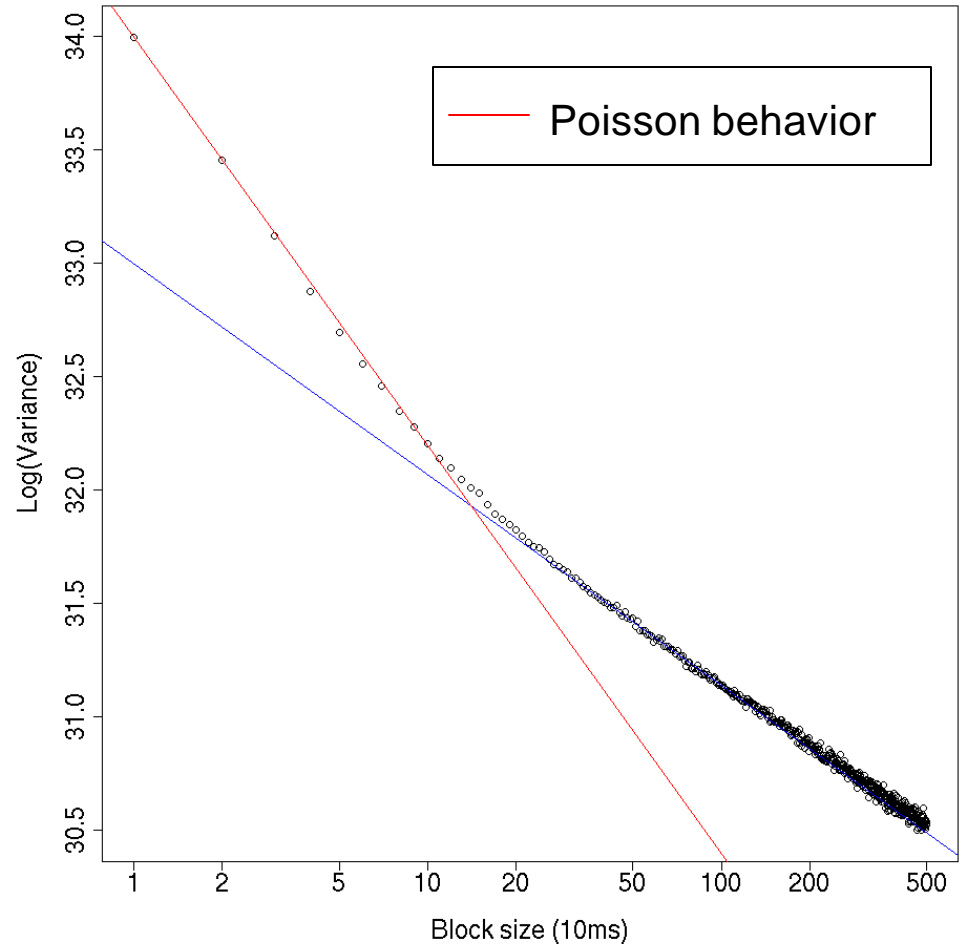
# Poisson vs Self-Similar traffic



Scale Invariant!

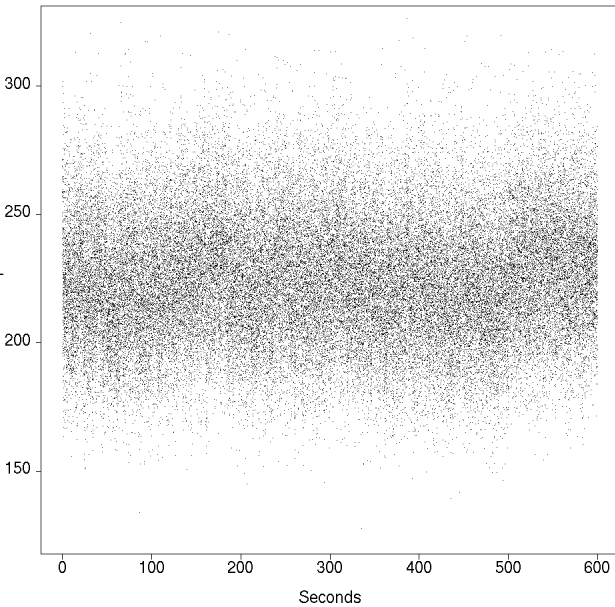
# Variance vs. Timescale: Internet

- $\text{Var}(X^{(m)}) = \sigma^2 m^{2H-2}$
- Random variable  $X$ :  
 $\text{Var}(X^{(m)}) = \sigma^2 m^{-1}$
- A dependency structure will cause  $\text{Var}(X^{(m)})$  to converge to zero slower than the rate  $m^{-1}$

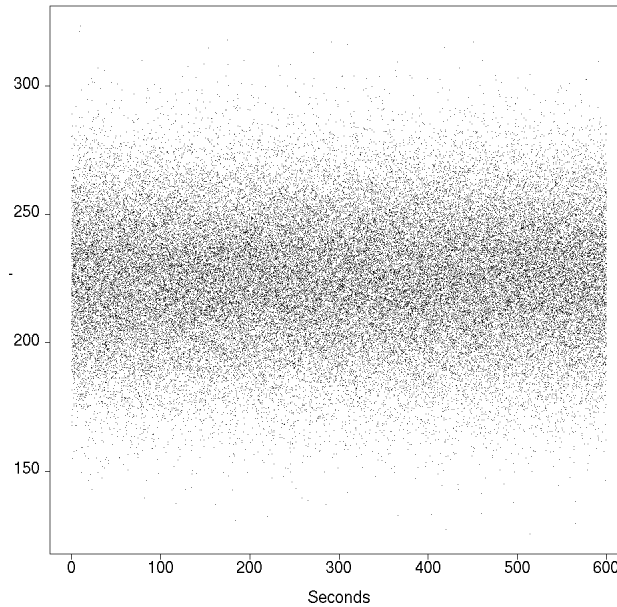


# Traffic Generation

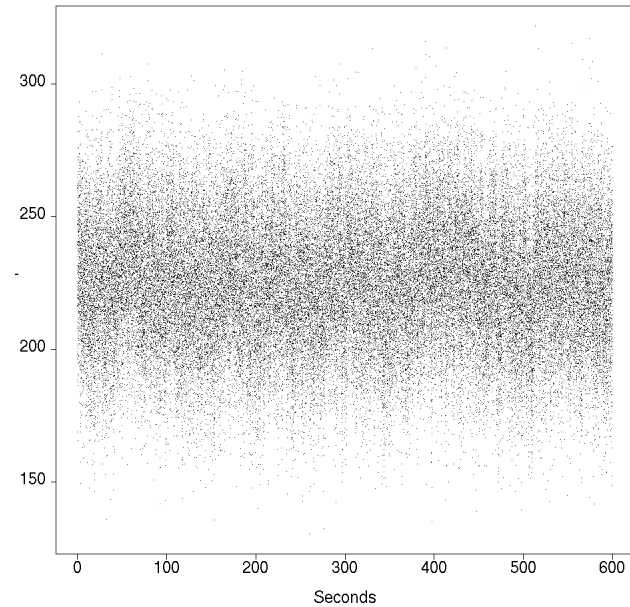
Real traffic



Poisson model

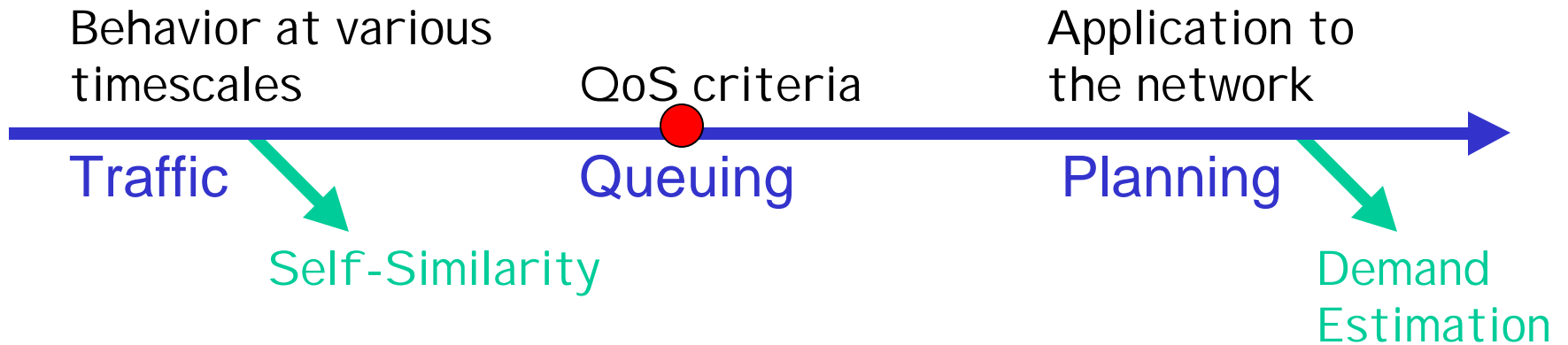


fBM model



Mean = 224.8 Mbps  
Std. Dev. = 24 Mbps

# QUEUING



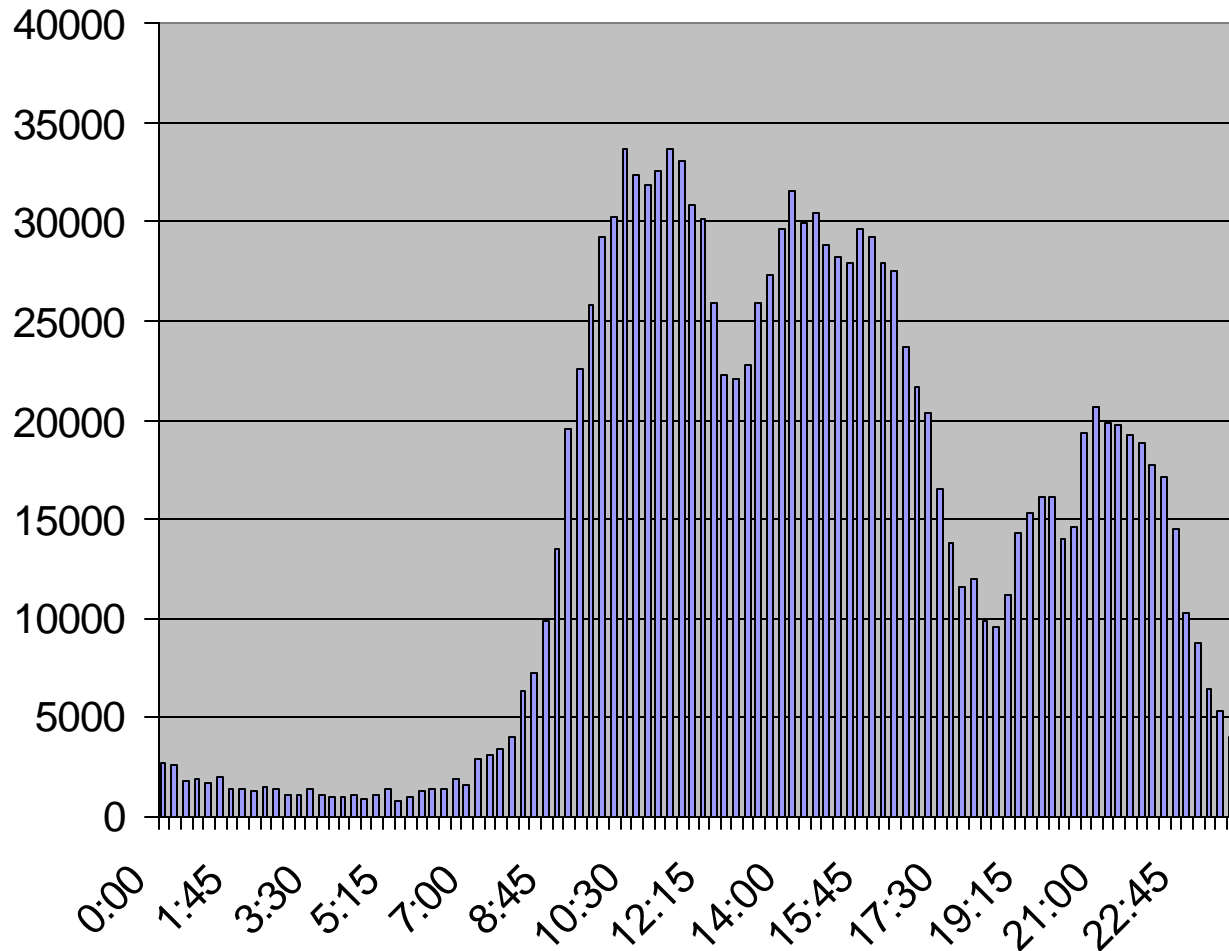
# Queuing Models/Simulation

- **M/M/1 queuing formula**
- Markovian
  - Poisson-process
  - Infinite number of sources
- "Circuits can be operated at over 99% utilization, with delay and jitter well below 1ms" [2] [3]
- **Self-Similarity**
- Traffic is bursty at many or all timescales
- "Scale-invariant burstiness (i.e. self-similarity) introduces new complexities into optimization of network performance and makes the task of providing QoS together with achieving high utilization difficult" [4]

# Telephony Traffic

(inter-city on 6/3/2002)

Centi-Erlang



# Voice Capacity Allocation

- Erlang traffic model(s)
- 1 Erlang = 1 hour of calls
  - Average numbers of calls in an hour
- Busy Hour Traffic: about 330 Erlang
- Erlang B formula (for 330 Erlang):
  - Blocking 1% -> 354 lines required
  - Blocking 0.1% -> 376 lines required
- Reverse calculation:
  - 100 lines available, blocking 0.1%:
    - 75 Erlang (=average of 75 calls in an hour)

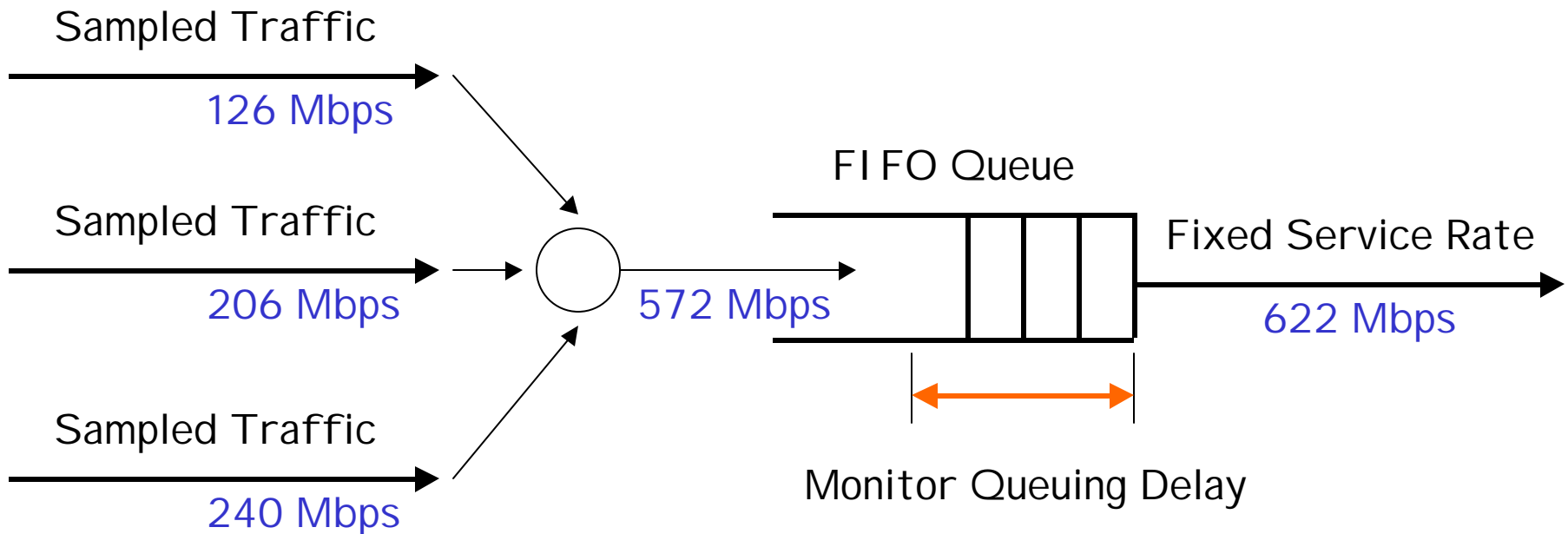
# IP Capacity Allocation

- Measurement data
  - E.g. 5-min average utilization
- Performance objectives
  - E.g. packet loss = 0%, jitter < 10ms
  - End-to-end: convert to per-hop objective
- But we don't have an "Erlang formula"...
- Two paths towards a solution:
  - 1) Model the traffic, and fit parameters
  - 2) Empirically derive guidelines

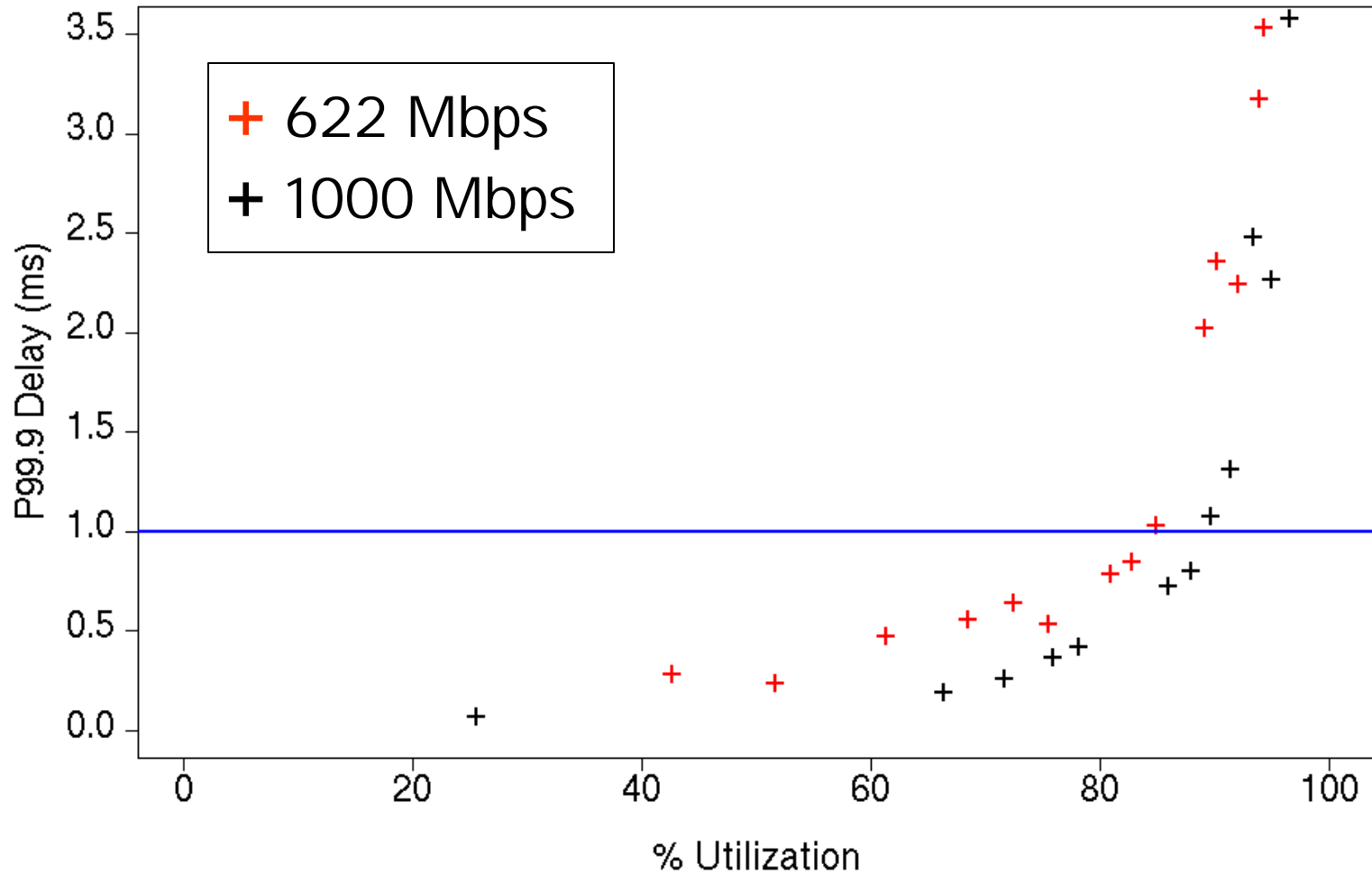
# Queuing Simulation

- Feed multiplexed sampled traffic data into FIFO queue
- Measure amount of traffic that violates the delay bound

Example: 92% Utilization



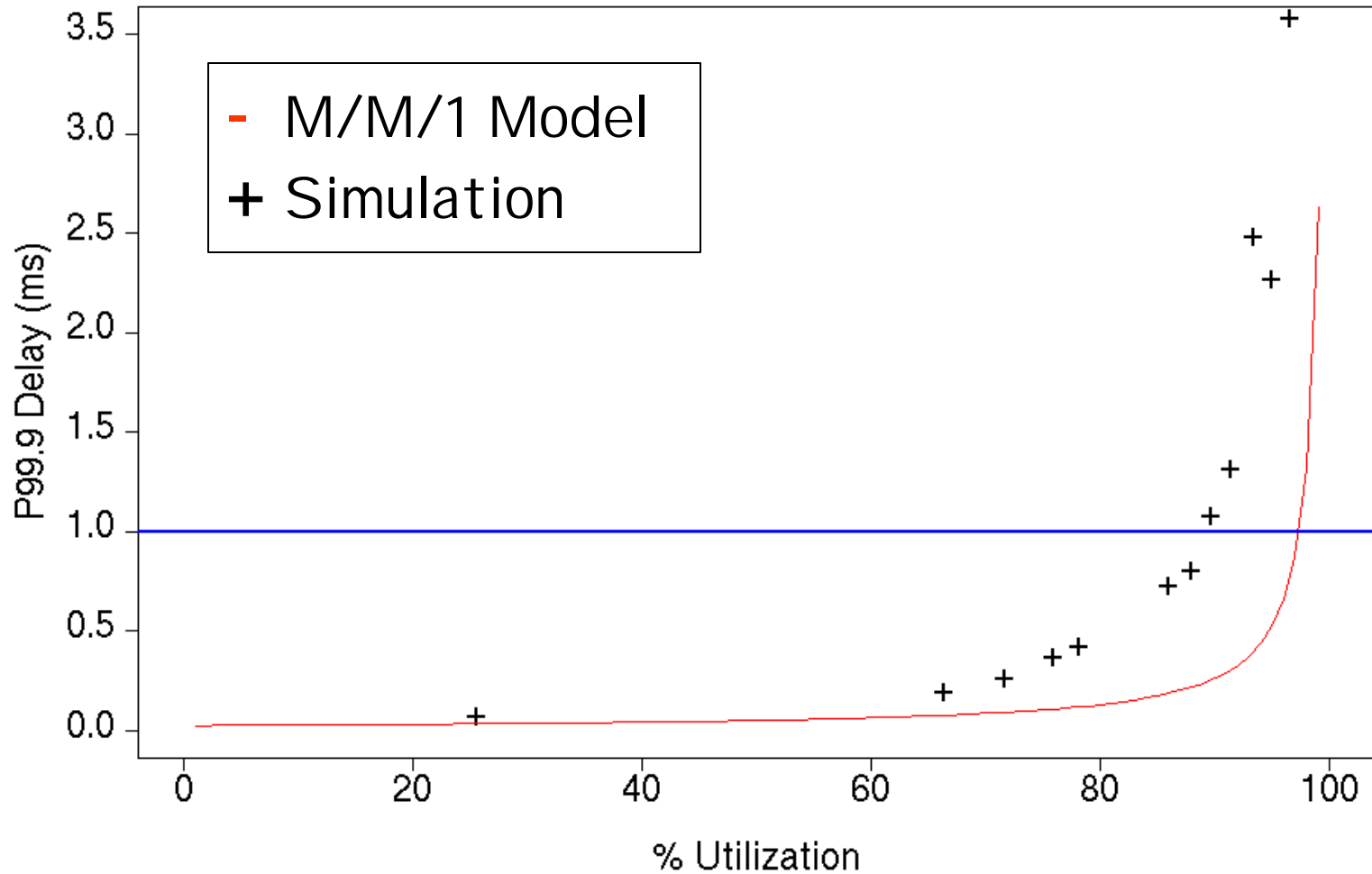
# Queuing Simulation: Results



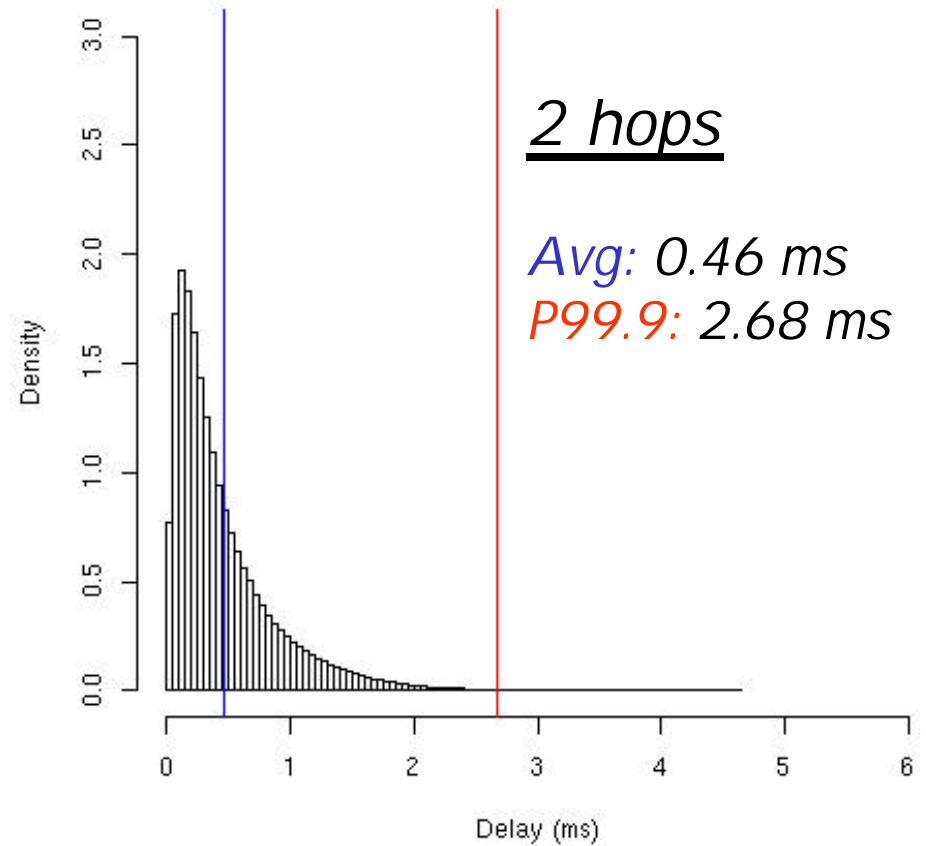
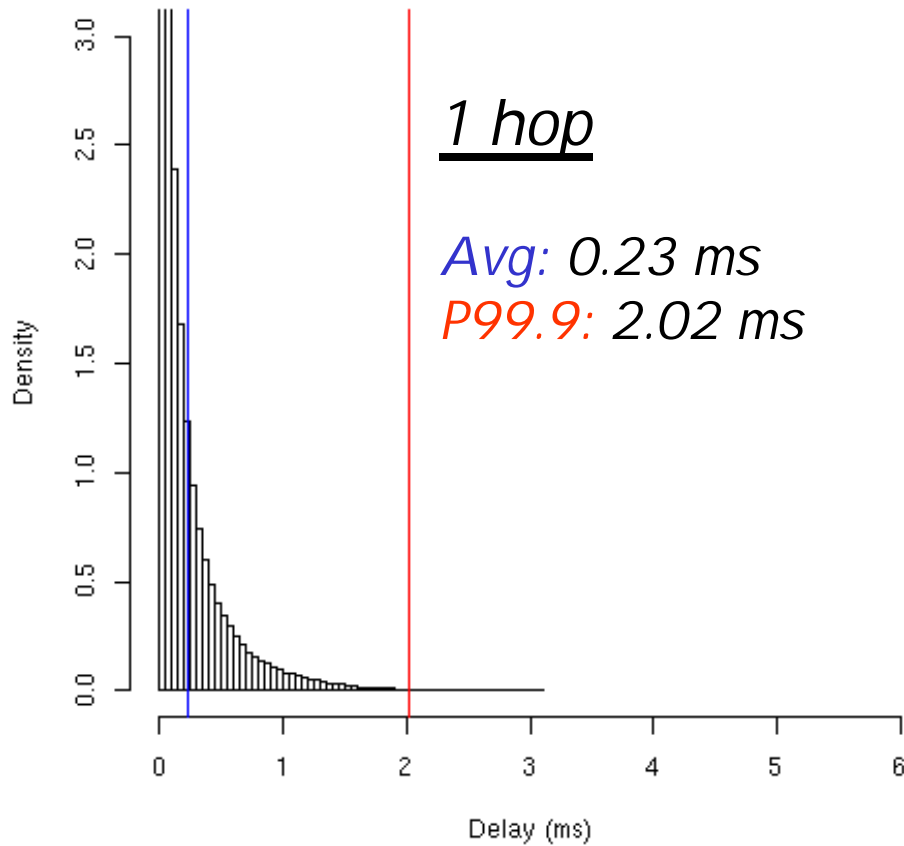
# Queuing Simulation: Numeric Results

- 1 Gbps (Gigabit Ethernet)
  - 1 ms delay bound for 999 out of 1000 packets (99.9-percentile):
    - 90% maximum utilization
- 622 Mbps (STM-4c/OC-12c)
  - 1 ms delay bound for 999 out of 1000 packets (99.9-percentile):
    - 85% maximum utilization

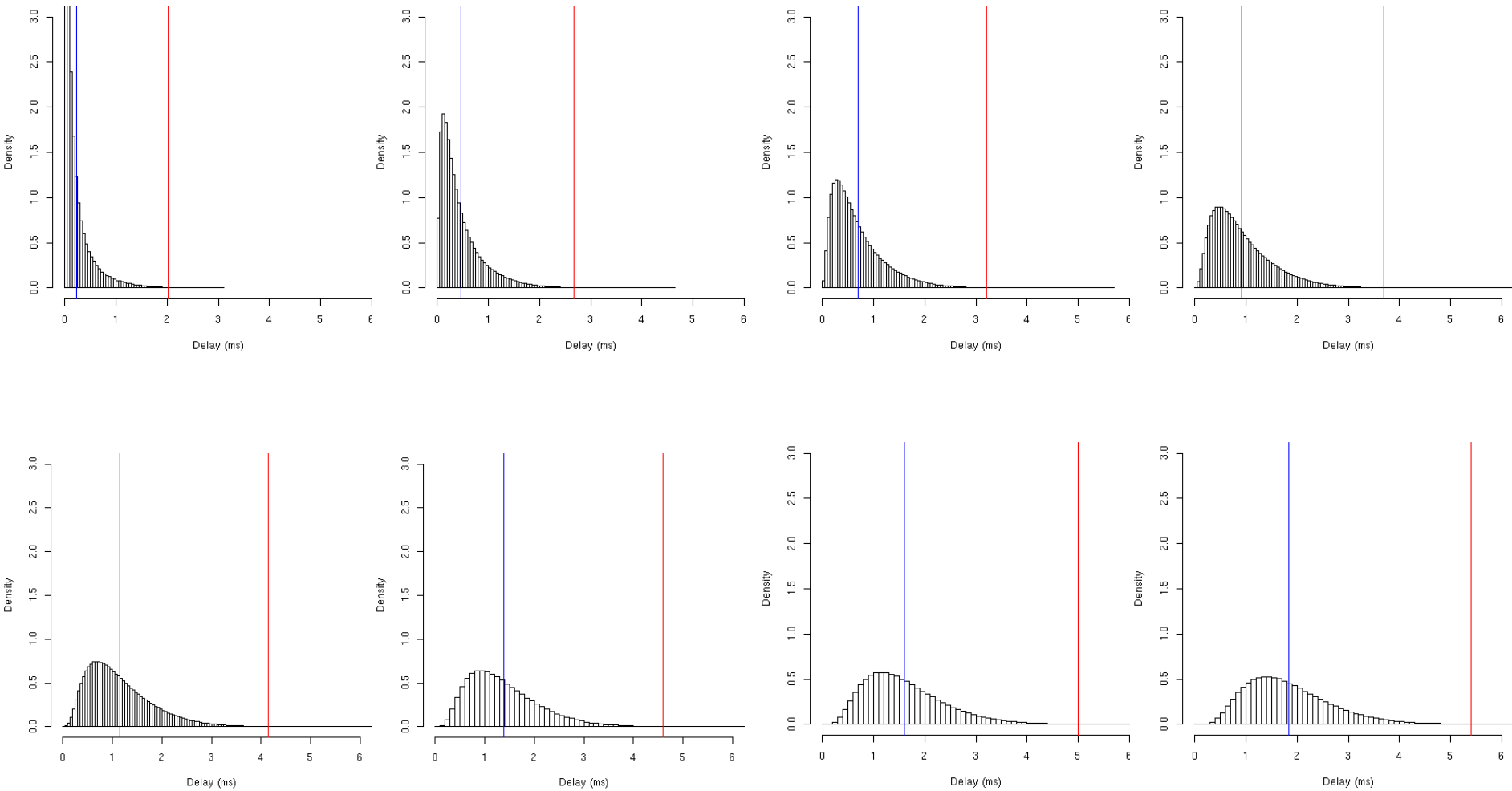
# Theory vs. Simulation (1Gbps)



# Multi-hop Queueing



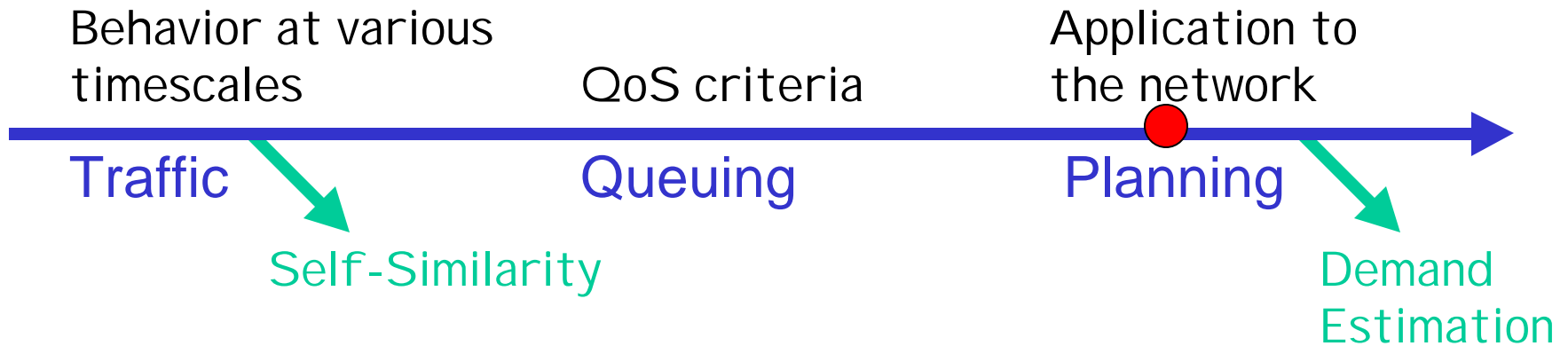
# Multi-hop Queueing (1-8 hops)



# Queueing: Summary

- Queueing Simulation:
  - Gigabit Ethernet (backbone) link
    - overprovisioning percentage in the order of 10% is required to bound delay/jitter to less than 1 ms
  - Lower speeds (<1G)
    - overprovisioning factor is significant,
  - Higher speeds (2.5G/10G)
    - overprovisioning factor becomes very small
- P99.9 multi-hop delay/jitter is not additive

# PLANNING



# QoS in Backbone Networks

- Requirements are:
  - low delay
  - low jitter
  - low packet loss
- Common practice in backbone networks is overprovisioning:
  - Enough capacity in the network to meet demands
  - In peak times, and under failure conditions
- Prevent significant queue buildup

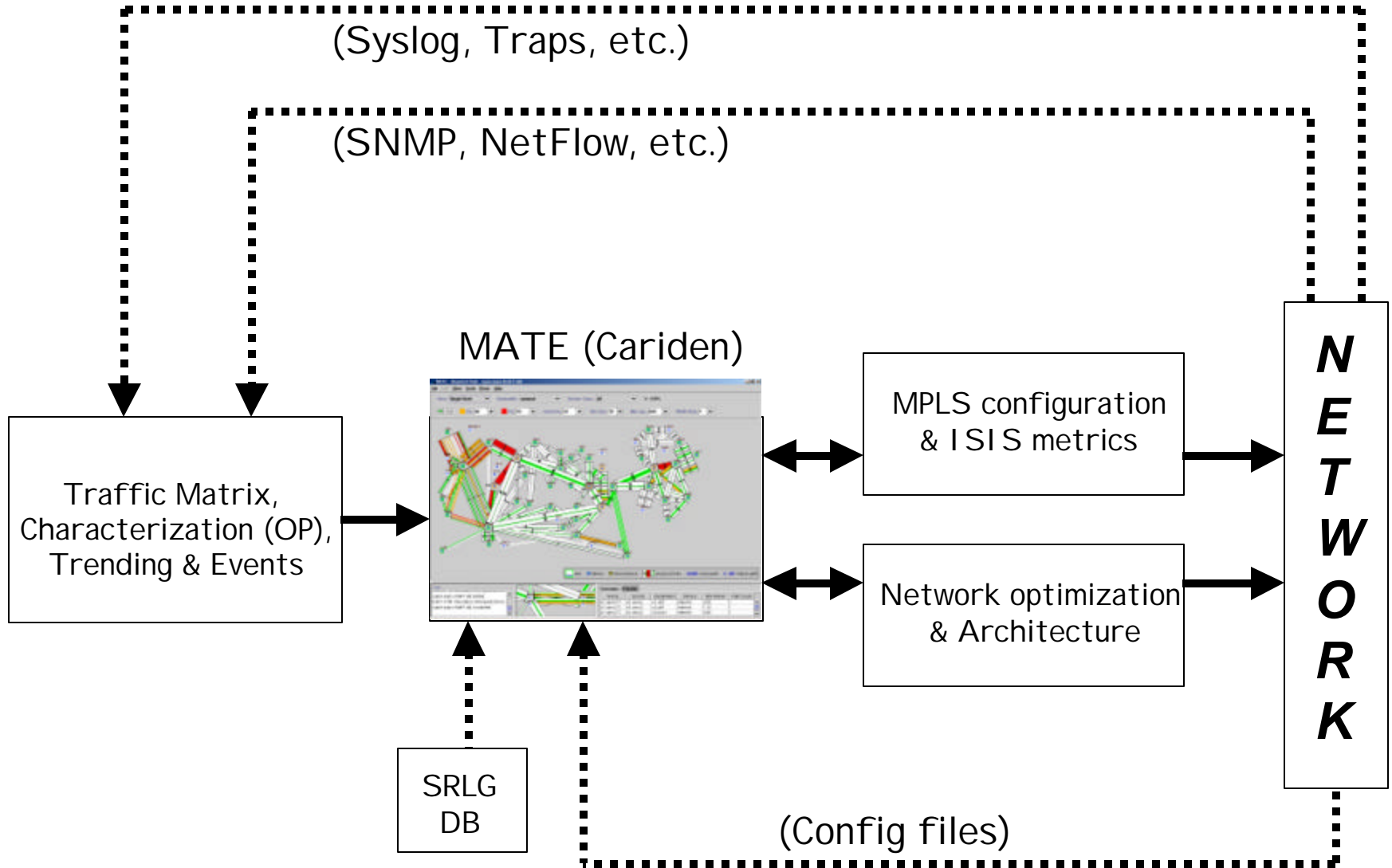
# QoS in Backbone Networks

- The overprovisioning approach is effective
  - See Packet Design presentation at NANOG 22 [1]
- But capital is limited today...
- Can we do better than the rules-of-thumb:
  - "upgrade at 40% or 50% utilization"
  - "maximum 75% utilization under (single) failure"
- Determine minimum capacity deployment to meet objectives under normal and failure conditions

# Network Planning Framework

- Demand Characterization
  - Long-term well-behaved traffic
  - Overprovisioning for short-term bursts can be experimentally derived
- How to use this for planning purposes?
- Failure Analysis
  - Determine failure scenarios
    - E.g. single link failures, routers, SRLG's, etc...
- Input for simulation

# Planning & Routing Framework



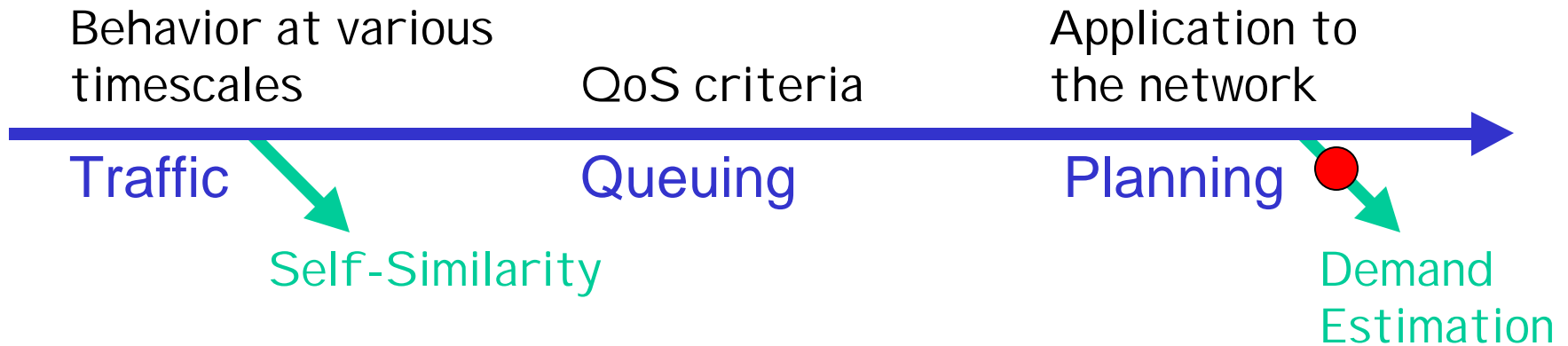
# Traffic Matrix

- How to build a site-to-site traffic matrix?
- MPLS:
  - Simple, using SNMP
- No MPLS:
  - NetFlow (aggregated)
  - Estimation (statistical inference)
    - see next slides...
  - Other?

# Planning: Summary

- IP Network Planning Process:
  - QoS targets -> "headroom"
  - Traffic Matrix
    - MPLS/NetFlow/Estimation
  - Trending
  - Simulation (Failure analysis)
  - Optimization
    - MPLS
    - IGP

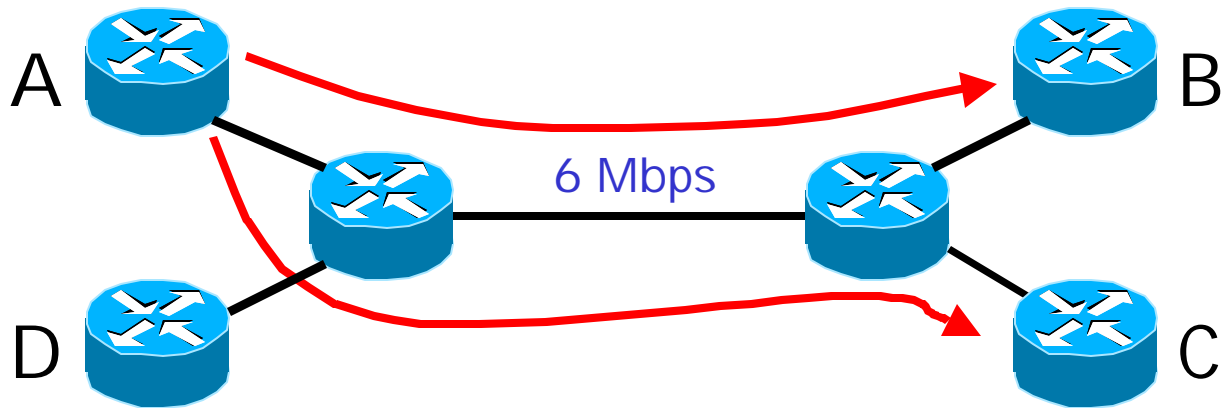
# Demand Estimation



# Demand Estimation

- Problem:
  - Estimate point-to-point demands from measured link loads
- Network Tomography
  - Y. Vardi, 1996
  - Similar to: Seismology, MRI scan, etc.
- Underdetermined system:
  - $N$  nodes in the network
  - $O(N)$  links utilizations (*known*)
  - $O(N^2)$  demands (*unknown*)

# Example



$y$ : link utilizations

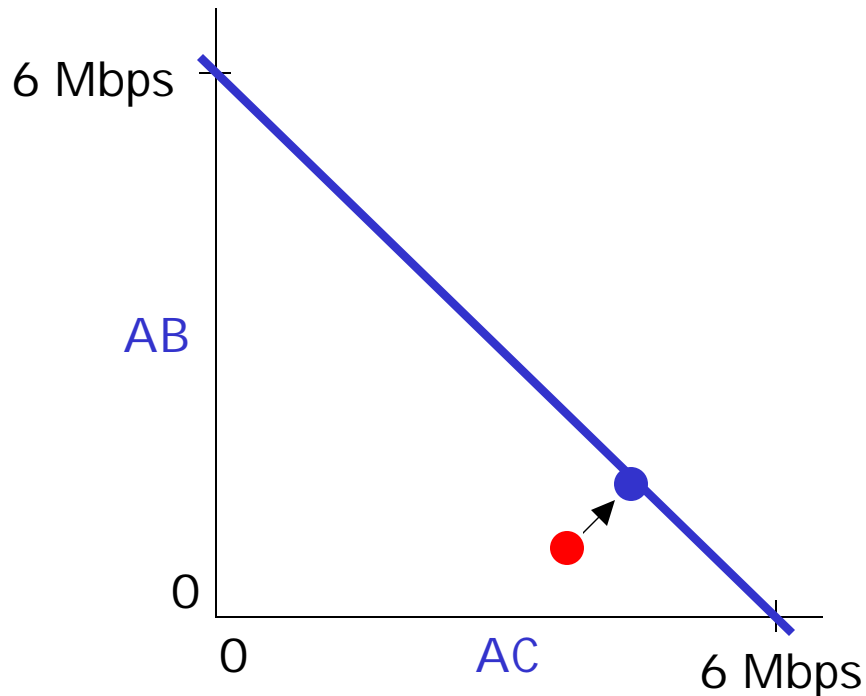
$A$ : routing matrix

$x$ : point-to-point demands

Solve:  $y = Ax$      $\rightarrow$  In this example:  $6 = AB + AC$

# Example

Solve:  $y = Ax$  -> In this example:  $6 = AB + AC$



## Additional information

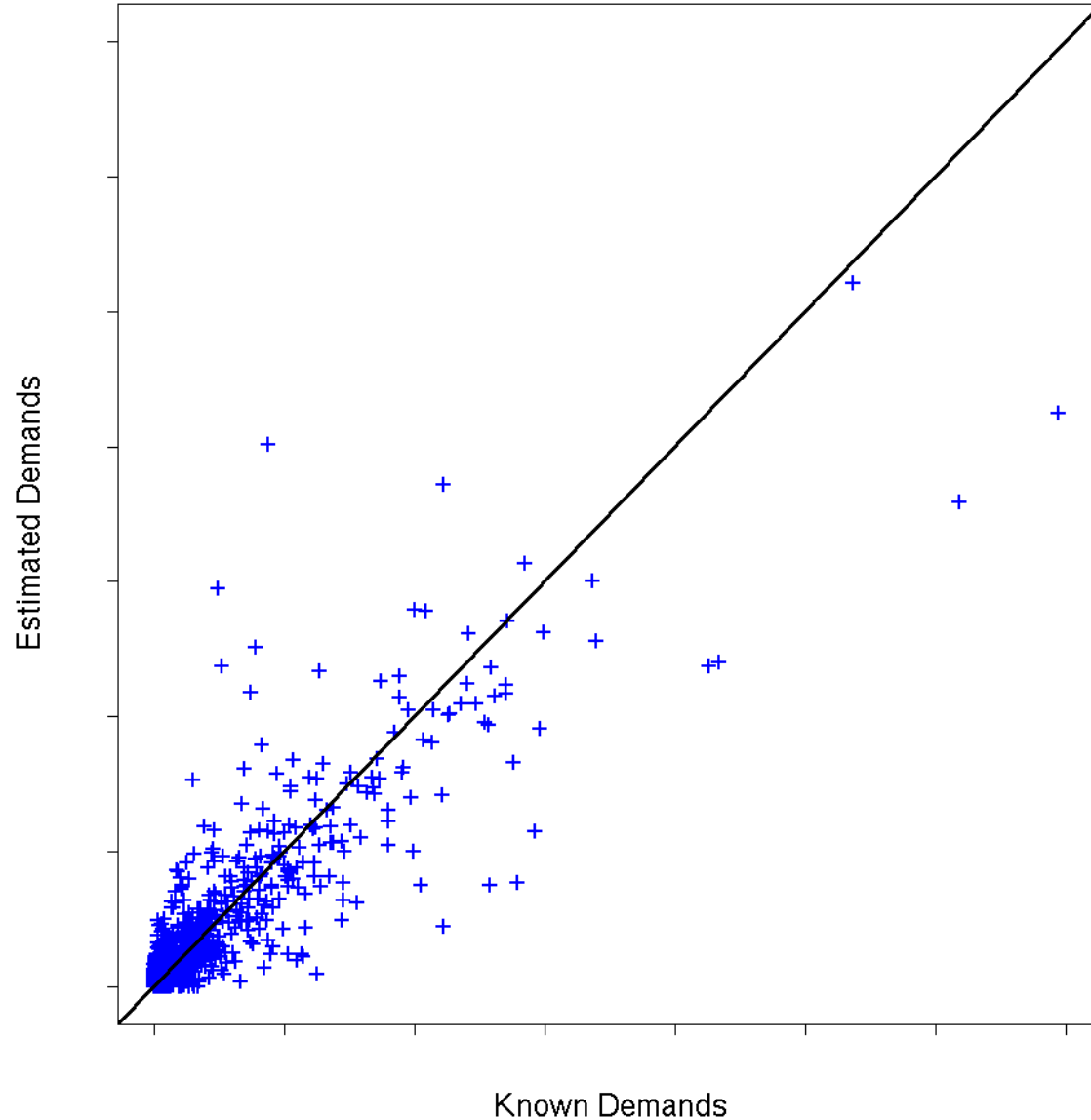
E.g. Gravity Model (every source sends the same percentage as all other sources of it's total traffic to a certain destination)

Example: Total traffic sourced at Site A is *50Mbps*. Site B sinks 2% of total network traffic, C sinks 8%.

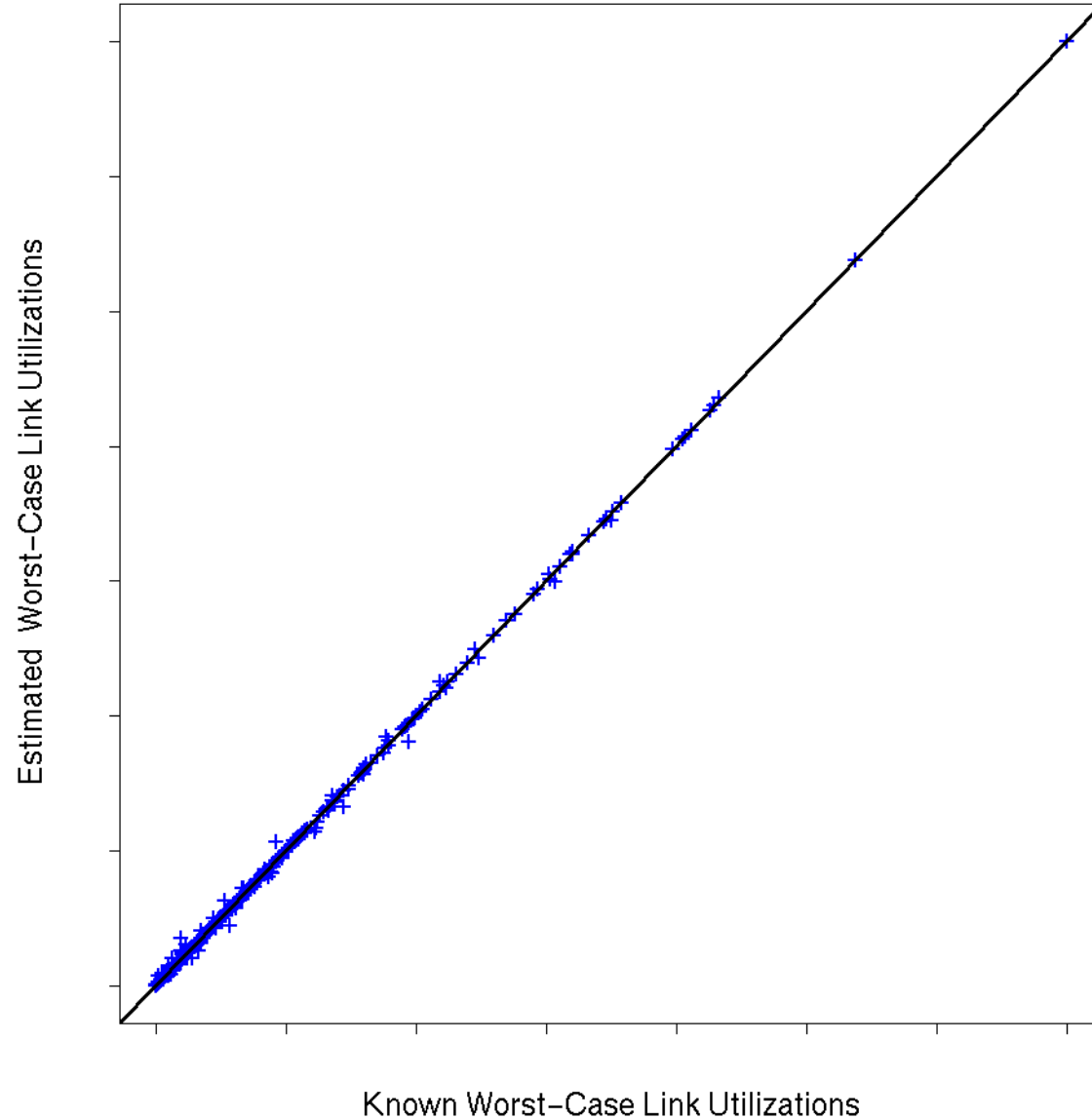
*AB = 1 Mbps and AC = 4 Mbps*

Final Estimate: AB = 1.5 Mbps and AC = 4.5 Mbps

# Real Network: Estimated Demands



# Estimated Link Utilizations!



# Demand Estimation Results

- Individual demands:
  - Inaccurate estimates...
- Estimated worst-case link utilizations:
  - Accurate!
- Explanation:
  - Multiple demands on the same path indistinguishable, but their sum is known
  - If these demands fail-over to the same alternative path, the resulting link utilizations will be correct

# Backbone Traffic Management

- Internet backbone traffic analysis
  - “well-behaved” at relevant timescales
- Queueing simulations
  - small overprovisioning factor for quite strict QoS criteria, on high-speed links
- Traffic Matrix
  - NetFlow/MPLS/Estimation
- IP Network Planning Process:
  - QoS targets -> “headroom”
  - Simulate and/or Optimize

# Acknowledgements

- Arman Maghbouleh (Cariden)
- Vishal Sharma (Metanoia)
- Haobo Yu (Packet Design)
- Alan Gous (Cariden)
- Richard Rensman (KPN)
- Bill Cleveland (Bell Labs)
- Jin Cao (Bell Labs)
- Yuan Gao (Bell Labs)
- Michael Johansson (KTH)

# References

- [1] Steve Casner, Cengiz Alaettinoglu and Chia-Chee Kuan, *A Fine-Grained View of High-Performance Networking*, NANOG 22 <http://www.nanog.org/mtg-0105/casner.html>
- [2] Chris Liljenstolpe, *Design Issues in Next Generation Carrier Networks*, MPLS 2001 Conference
- [3] Peter Lothberg, *A View of the Future: The IP-Only Internet*, NANOG 22, <http://www.nanog.org/mtg-0105/lothberg.html>
- [4] Zafer Sahinoglu and Sirin Tekinay, *On Multimedia Networks: Self-Similar Traffic and Network Performance*, IEEE Communications Magazine, January 1999
- [5] Robert Morris and Dong Lin, *Variance of Aggregated WebTraffic*, IEEE INFOCOM 2000, Tel Aviv, March 2000, pages 360-366

# Related Work

- [1] Jin Cao, William S. Cleveland, Don X. Sun,  
*A Statistical Model for Allocating Bandwidth to Best-Effort Internet Traffic*, to appear in *Statistical Science*, 2004
- [2] Chuck Fraleigh, Fouad Tobagi, Christophe Diot,  
*Provisioning IP Backbone Networks to Support Latency Sensitive Traffic*, Proc. IEEE INFOCOM 2003, April 2003
- [3] Cao, J., W.S. Cleveland, D. Lin, D.X. Sun,  
*Internet Traffic Tends Towards Poisson and Independent as the Load Increases*. In *Nonlinear Estimation and Classification*, New York, Springer-Verlag, 2002

# Further Reading

## Statistics (traffic analysis):

- [1] W.N. Venables B.D. Ripley,  
*Modern Applied Statistics with S-PLUS (3<sup>rd</sup> edition)*,  
New York, Springer-Verlag, 1999
- [2] Peter Dalgaard,  
*Introductory Statistics with R*,  
New York, Springer-Verlag, 2002
- [3] W.N. Venables B.D. Ripley,  
*S Programming*,  
New York, Springer-Verlag, 2002

# Further Reading

## Self-Similarity:

- [4] Benoit B. Mandelbrot,  
*The Fractal Geometry of Nature*,  
New York, W.H. Freedman and Company, 1982
- [5] Kihong Park, Walter Willinger (ed.),  
*Self-Similar Network Traffic and Performance Evaluation*,  
New York, John Wiley & Sons, 2000
- [6] D. Sornette,  
*Critical Phenomena in Natural Sciences*,  
New York, Springer-Verlag, 2000
- [7] Paul Doukhan, George Oppenheim, Murad S. Taqqu (ed.),  
*Long-Range Dependence*,  
Boston, Birkhauser, 2003
- [8] Paul Embrechts, Makoto Maejima,  
*Selfsimilar Processes*,  
Princeton, Princeton University Press, 2002

# Further Reading

## Queueing:

- [9] Leonard Kleinrock,  
*Queueing Systems, Volume I: Theory*,  
New York , John Wiley & Sons, 1975
- [10] Leonard Kleinrock,  
*Queueing Systems, Volume II: Computer Applications*,  
New York, John Wiley & Sons, 1976
- [11] Robert B. Cooper,  
*Introduction to Queueing Theory, Second Edition*,  
New York, North Holland, 1981
- [12] Jean-Yves Le Boudec, Patrick Thiran,  
*Network Calculus: a theory of deterministic queueing systems  
for the Internet*,  
New York, Springer-Verlag, 2001

# Further Reading

## Network Optimization (planning):

- [13] Ravindra Ahuja, Thomas Magnanti, James Orlin,  
*Network Flows: Theory, Algorithms, and Applications*,  
New Jersey, Prentice-Hall, 1993
- [14] Dorit Hochbaum (ed.),  
*Approximation algorithms for NP-hard problems*,  
Boston, PWS Publishing Company, 1995

## Demand Estimation:

- [15] Paul Garthwaite, Ian Jolliffe, Byron Jones,  
*Statistical Inference (2<sup>nd</sup> edition)*,  
Oxford, Oxford University Press, 2002

## Thomas Telkamp



Croeselaan 148  
3521 CG Utrecht  
The Netherlands

Tel: +31 30 238 1250

Fax: +31 30 238 1555

E-mail: [telkamp@gblix.net](mailto:telkamp@gblix.net) / [telkamp@ltdcm.net](mailto:telkamp@ltdcm.net)