SIEMENS

Tutorial

Internet Traffic Measurement and Modelling

Joachim Charzinski
j.charzinski@ieee.org / joachim.charzinski@icn.siemens.de

Siemens
Information and Communication Networks
Munich, Germany

SIEMENS

Acknowledgements Thanks to...

- Siemens
  - Paul Schlüter
  - Michael Schopp
  - Manfred N. Huber
- University of Stuttgart
  - Prof. P.J. Kühn
  - Stefan Bodamer
  - Johannes Färber
  - Klaus Dolzer
- University of Würzburg
  - Prof. P. Tran-Gia
  - Norbert Vicari
  - Stefan Köhler
- COST 257

SIEMENS

Outline

- 1. Introduction
  2. User and Application Behavior
  3. Application Behavior Measurements
  4. Self-Similarity
  5. User Behavior Measurements
  6. Backbone Measurements
  7. Web Performance
  8. Models
  9. Implications for Simulation
  10. Implications for Quality of Service

SIEMENS

1. Introduction

- Dimensioning (what for)
- Measurement
- Tools
- CDF Graphs

SIEMENS

THE Internet?

The Internet versus other IP Networks

- The Internet
  - publicly accessible network
  - worldwide connectivity
  - looks different to different users
  - multiple worldwide ISPs
  - no single backbone
  - connectivity is changing
- Other IP networks
  - company networks, Intranets
  - local area networks
  - special purpose networks
    - VoIP, telephony signalling (future)
    - network management
    - measurement networks

SIEMENS

Dimensioning for IP Traffic

- non-bottleneck links
  - traffic not to be influenced by the link being dimensioned
  - use traffic patterns occuring on the link if capacity $\rightarrow \infty$
  - dimension to the rate needed to have given small impact on traffic
- bottleneck dimensioning
  - consider TCP behavior
  - dimension to offer a certain rate to every active connection or subscriber
- blocking considerations
  - if access control is performed, the blocking probability is also an important paramter
- Example
  - access network with access lines and common trunk line
  - dimension trunk line to fit access lines capacity
  - including the statistical multiplexing gain

SIEMENS

Dimensioning Applications

- Evaluation of user perceived quality of service
  - for SLAs
  - for advertising
- Matching of capacities within network element
  - processor vs. line capacity
  - line card vs. backplane bit rates
  - link bandwidth between modules
- buffer sizes
  - for a TCP connection
  - for a subscriber / port
  - for rate adaptation
- compute / predict system performance
  - number of subscribers to be served (depending on subscriber model)
  - fitness for future traffic demands

SIEMENS

Traffic Engineering vs. Traffic Management

There are different definitions around!

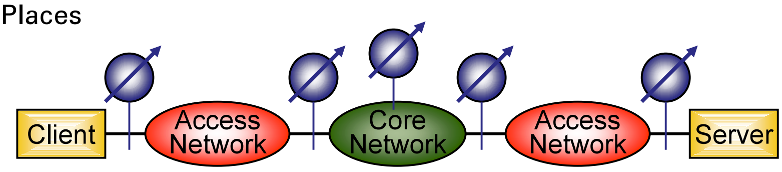
- The MPLS-related term “Traffic Engineering” (e.g. in MPLS) corresponds to traditional “Traffic Management”
- Traffic Engineering (trad.)
  - network planning
  - traffic forecast
  - link and node dimensioning
  - capacity planning
  - long term task
- Traffic Management (trad.) / Traffic Engineering (IETF)
  - managing dynamic traffic load
  - route adaptation
  - capacity management
  - shorter term task

  ⇒ a network management task

SIEMENS

Internet Measurement

Places



Methods

- Packet trace evaluated offline
- Online pre-processed packet trace
- Access Logging (e.g. in proxies)
- Cyclic reading of counter values
- (ping) delay and bandwidth tests
- application layer tests

(first four: passive; last two: active)

SIEMENS

Internet Measurement Data Collection

- Packet trace
- Flow trace: one record per
  - TCP connection
  - other flow levels
- Preprocessed data
  - average values
  - Wavelet coefficient sets

Access Log

- one or two record(s) per dial-in session

Active Measurements

- Measure delay, loss, bandwidth, application performance between two points

SIEMENS

Internet Measurement
Some Tools

- Packet Trace Tools
  - tcpdump (Paxson)
  - Ethereal, argus, etc
- Flow detecting / higher layer tracing tools
  - OC3mon (Apisdorf/Claffy/Thompson) + CoralReef
  - NeTraMet (Brownlee)
  - BLT (Feldmann)
  - tcpanaly (Paxson)
- Active test tools
  - visit NIMI at http://www.ncne.nlanr.net/nimi/
- Network management tools
  - use SNMP to retrieve counter values from network elements
  - RMON probes

SIEMENS

Most results will be given as
Complementary Distribution Functions

- Complementary distribution function

$$F_X^C(x) = P\{X > x\} = 1 - P\{X \le x\}$$

Logarithmic scaling!

- Example: negative exponential distribution, $x_0 = E[X]$

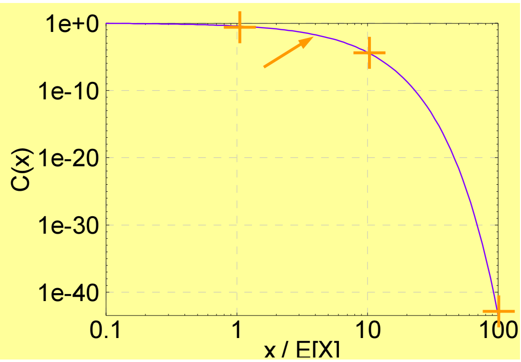
$$F_{\exp}^C(x) = e^{-x/x_0}$$

$P\{X > E[X]\} = 0.368$

$P\{X > 10E[X]\} = 4.5 \cdot 10^{-5}$

$P\{X > 100E[X]\} = 4 \cdot 10^{-44}$

$P\{X > 4.61E[X]\} \approx 0.01$



SIEMENS

Outline

Information and Communications

SIEMENS

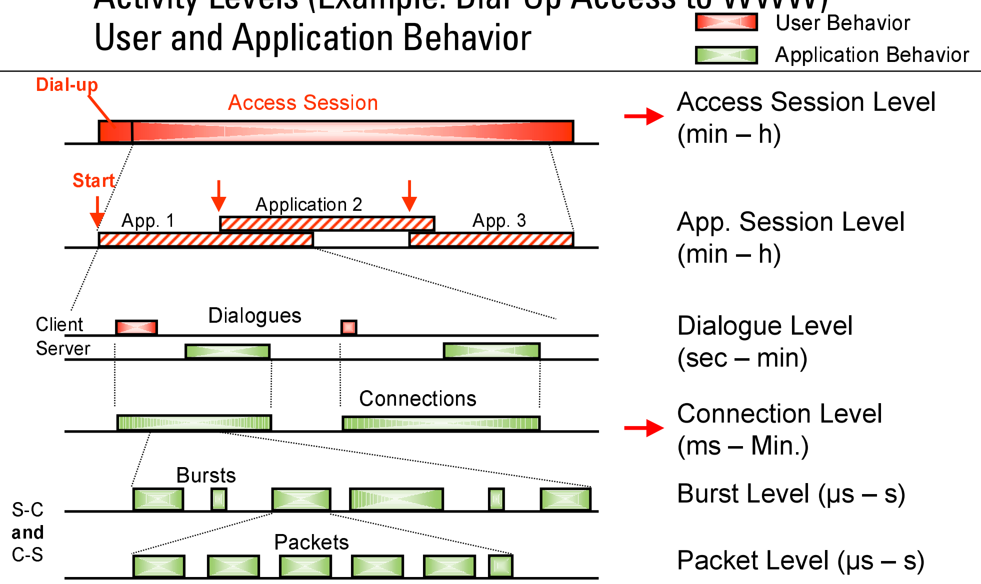
2. User and Application Behaviour

- Activity Levels
- Separation of User and Applications Behaviour
- TCP Effects
- Flow Definitions

Information and Communications

SIEMENS

Activity Levels (Example: Dial-Up Access to WWW) User and Application Behavior



SIEMENS

Separation of User and Applications Behaviour

User Behaviour

- Access session start
- Choice and usage of Applications

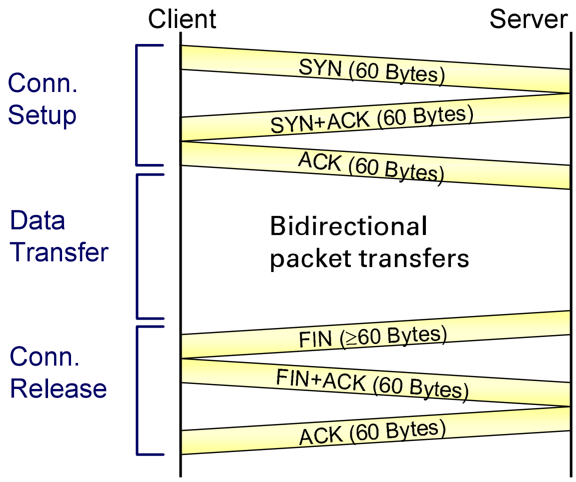
Applications Behaviour

- size distribtution of E-Mail, HTTP page, FTP download
- packet traffic due to
  - TCP connection management
  - TCP acknowledgements
  - higher layer protocols (authentication, GET/PUT, etc)
  - stream flow control
- also: (in parts): application usage duration

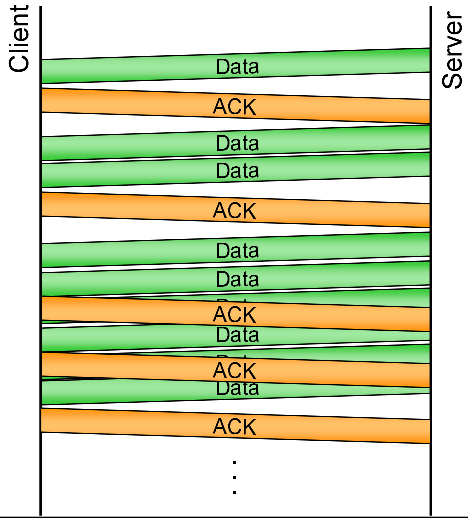
TCP
Effects

- TCP introduces traffic patterns / dependencies (see later)
- lightweight connection management
  - 3 packets to open a connection
  - 3 packets to close a connection
  - connections only seen by end points
  - applications can use short-lived TCP connections
  - application behavior differs from “typical ATM” applications
- Reliable Transport
  - Acknowledgements in reverse direction
- Flow Control
  - TCP adapts its rate to the “fair share” on a bottleneck link
  - influence of delay and loss
  - different implementations around

TCP
Connection Management



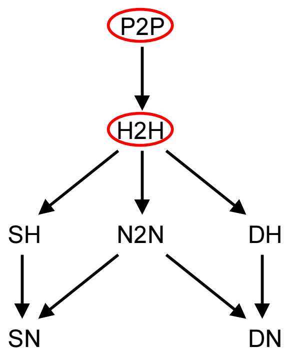
TCP
Bulk Data Transfer



- Acknowledgements in reverse direction
- “Delayed ACK” option: only acknowledge every 2nd packet (+timeout)
- Slow start
  - exponential growth of sender’s window

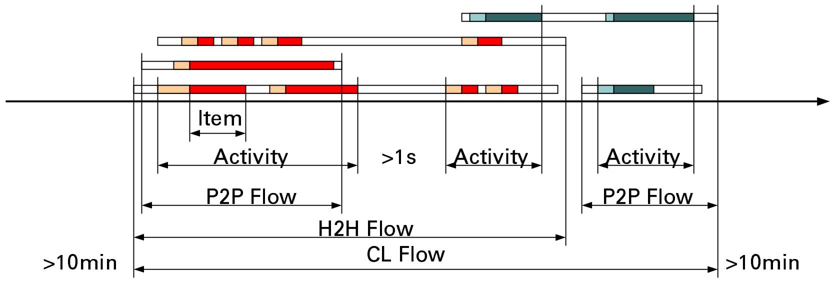
Flow Definitions
Traffic Aggregation Relations

- Port to Port (P2P)
  - single TCP connection or UDP relation
  - fixed IP addr., TOS, protocol and port numbers at both ends
- Host to Host (H2H)
  - same pair of IP addresses
- Network to Network (N2N)
  - same pair of network addresses
- Source Host (SH)
- Source Network (SN)
- Destination Host (DH)
- Destination Network (DN)
- others (e.g. application specific, bidirectional)
  - single HTTP item
  - HTTP download activity in reaction to user click
  - all Client traffic (CL)

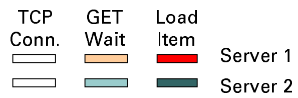


SIEMENS

Measurement
Flow Definitions etc



- P2P Flow (TCP connection)
- H2H Flow (one server)
- CL Flow (one client session)
- Activity
- Element



SIEMENS

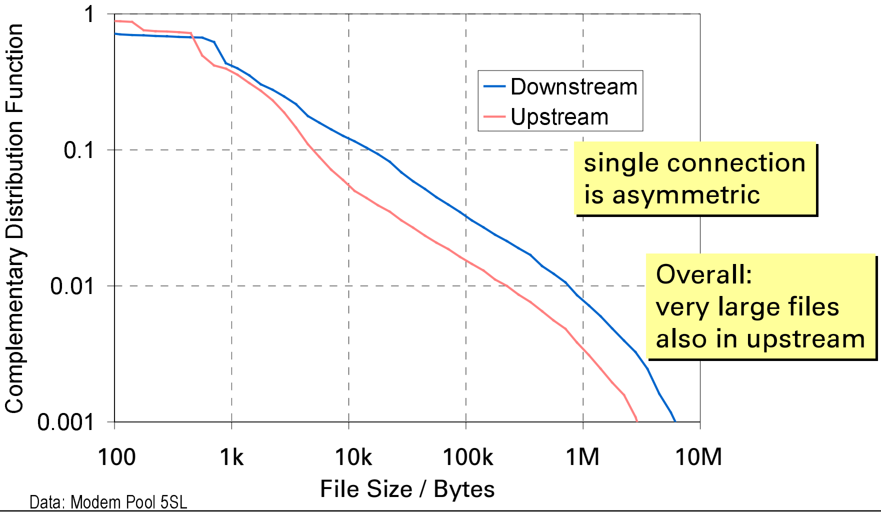
Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

3. Application Behaviour Meas.

- E-Mail
- File Transfer Protocol
- WWW Access
- Multimedia Applications
- Online Games
- Others

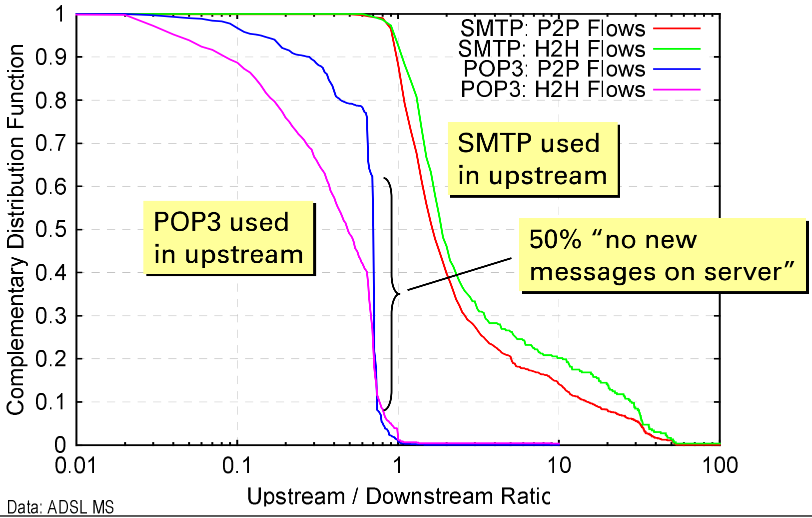
E-Mail
File Size Distributions



Data: Modem Pool 5SL

SIEMENS

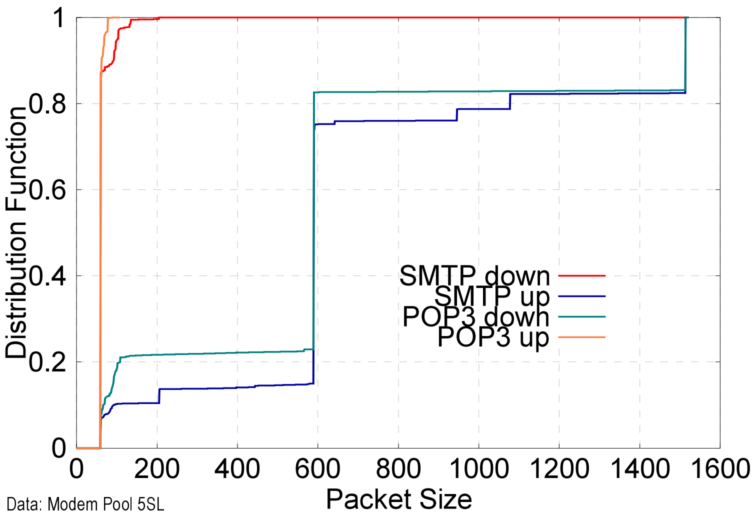
E-Mail
Rate Ratio Distributions



Data: ADSL MS

SIEMENS

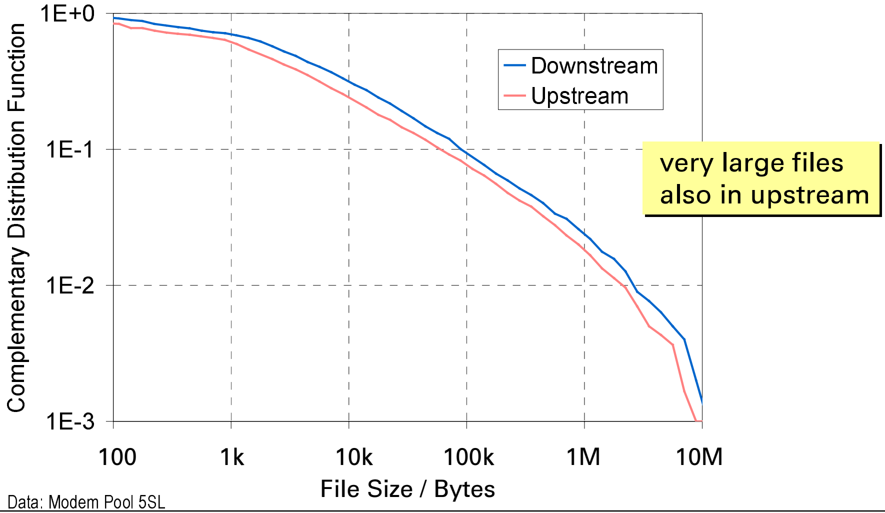
E-Mail
Packet Size Distributions



Data: Modem Pool 5SL

SIEMENS

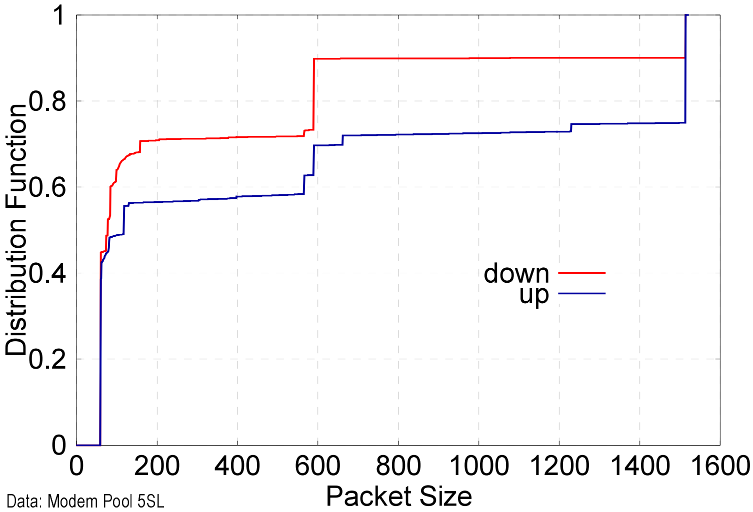
FTP
File Size Distributions



Data: Modem Pool 5SL

SIEMENS

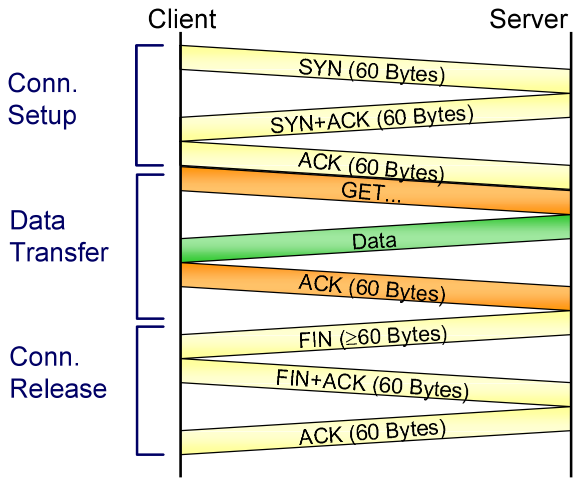
FTP
Packet Size Distributions



Data: Modem Pool 5SL

SIEMENS

WWW Access



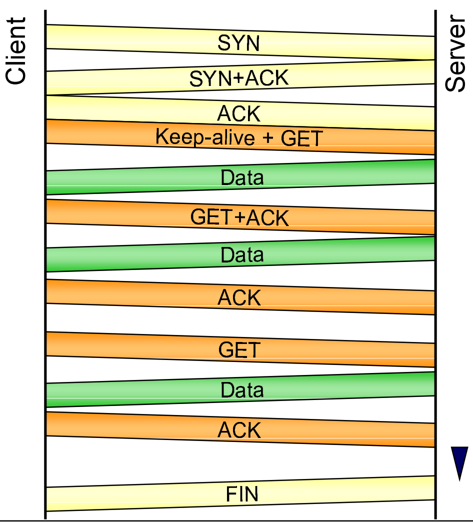
Message	% Pkts	Pkt./ Conn.	% Bytes
SYN (60 Bytes)	4	1	0.5–0.7
SYN+ACK (60 Bytes)	4	1	0.5–0.7
ACK (60 Bytes)	4	1	0.5–0.7
GET...	6	1.5	5–6
Data	40	10	83–88
ACK (60 Bytes)	30	7.5	3–5
FIN (≥60 Bytes)	3	0.75	<1
FIN+ACK (60 Bytes)	3	0.75	0.4
ACK (60 Bytes)	2	0.5	0.4

Up/down ratios:
1:1.05 1:7.9

Measurements: ADSL MS and Modem Pool 5SL: >2 mio. connections

SIEMENS

WWW Access
Keep-Alive / Persistent Connections



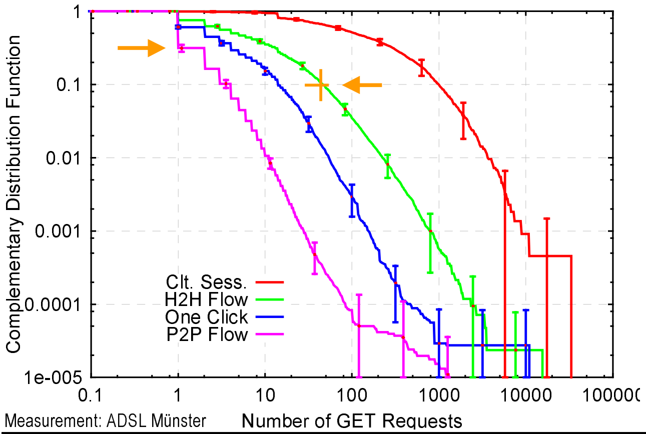
HTTP Version 1.0:
- Keep-alive
- to be negotiated

HTTP Version 1.1
- persistent connections
- default

Timeout (e.g. 15sec)

SIEMENS

WWW Access
Number of GET Requests per HTTP/TCP Connection



P2P = Port to Port
H2H = Host to Host
CL = Client

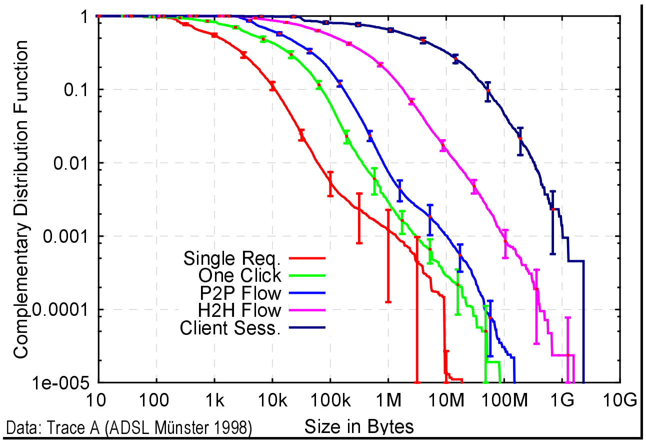
Timeout = 10min

30% of all connections serve more than one request

90% of all flows get less than 40 elements from the same server

SIEMENS

WWW Access
Flow Volumes

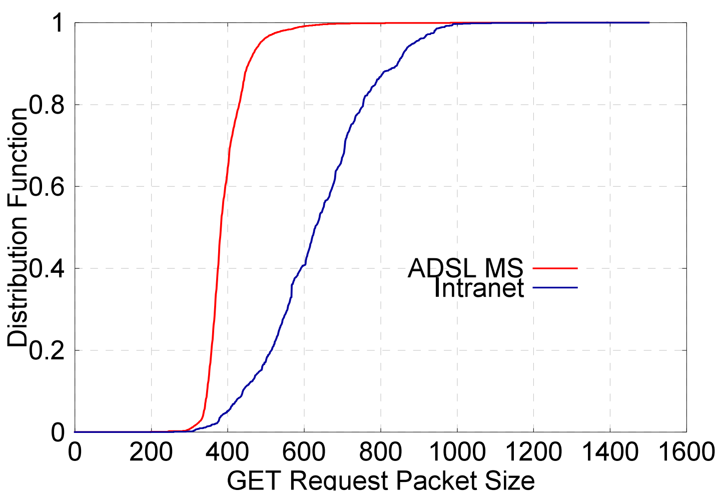


P2P = Port to Port
H2H = Host to Host
CL = Client

Timeout = 10min

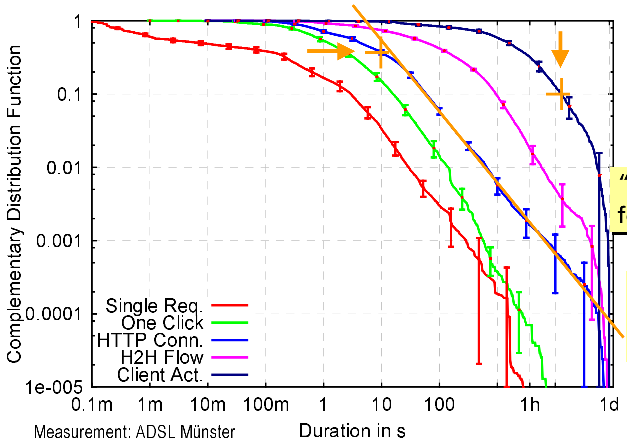
SIEMENS

WWW Access
Distribution of GET request packet sizes



SIEMENS

WWW Access
Connection, Flow and Usage Duration



P2P = Port to Port Flow
H2H = Host to Host Flow
CL = Client Usage
Timeout = 10min

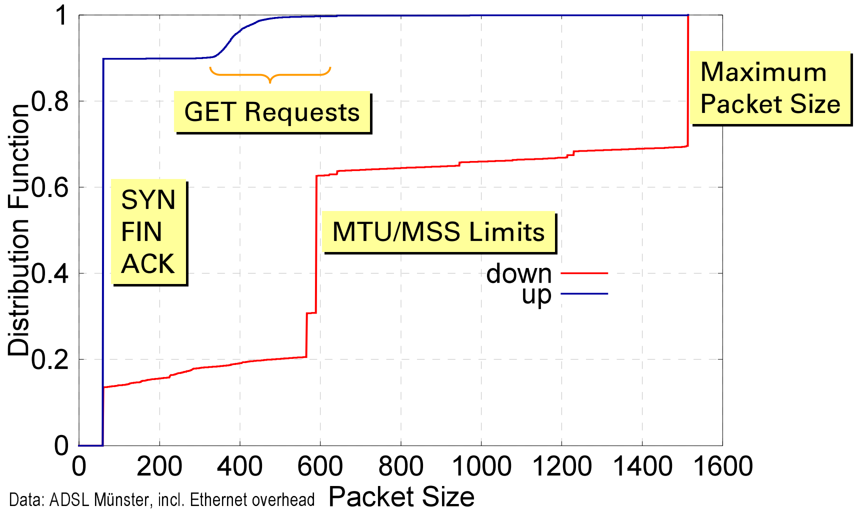
"Power Laws" for nearly four orders of magnitude

$\alpha<1$
→ infinite expectation if continued until ∞

60% of all connections are shorter than 10 sec

10% of all Web sessions are longer than 3 hrs

WWW Access
Packet Size Distribution



Data: ADSL Münster, incl. Ethernet overhead

WWW Access
Locality

- Majority of connections with external servers
- Large percentage of unresolvable network addresses
- Example: ADSL measurement
  - local = within DFN WiN (German Research Network)
  - non-local = outside DFN WiN
  - undefined = no reverse resolution of IP address possible

Locality	% P2P Flows	% H2H Flows
Local	12	7
Non-local	63	63
Undefined	25	30

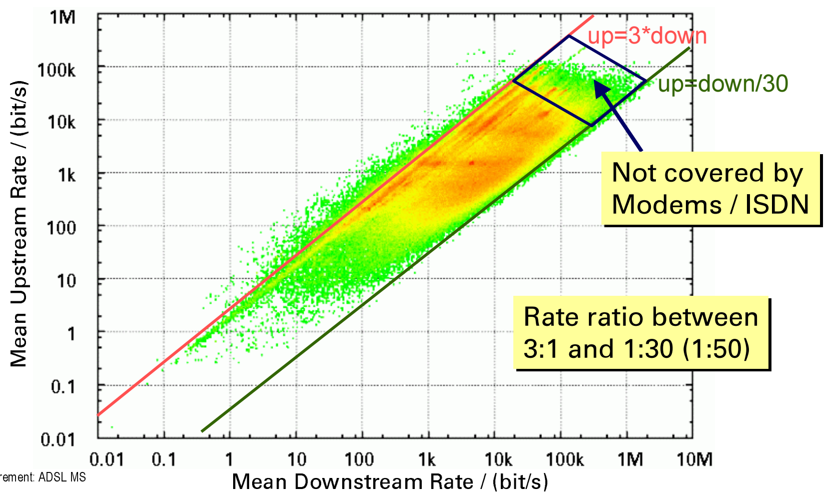
SIEMENS

WWW Access
Distributions Summary

- GET request packet size
  - URL length + HTTP protocol information
  - Normal Distribution around 400 –700, s.d. 50–200 bytes
  - Larger in database-enabled environments
- Number of GET requests
  - Power tail (?), α=2–2.5
- Downstream connection volume
  - Power tail, α=1.1–1.5
  - infinite variance
- Connection duration
  - Power tail, α=0.6–1.2
  - infinite variance and expectation
- Packet sizes depend on configuration
  - minimum TCP+IP=40Byte for SYN, FIN, ACK
  - maximum given by MTU (client and server configuration dependent)

SIEMENS

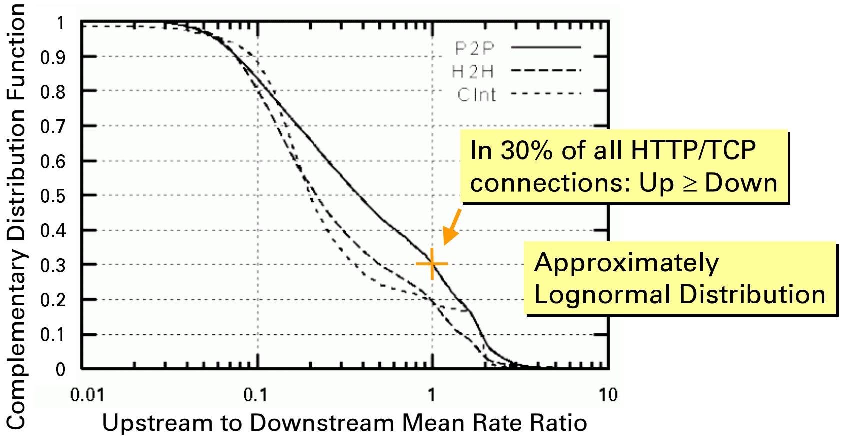
WWW Access
Bit Rate Symmetry



Measurement: ADSL MS

SIEMENS

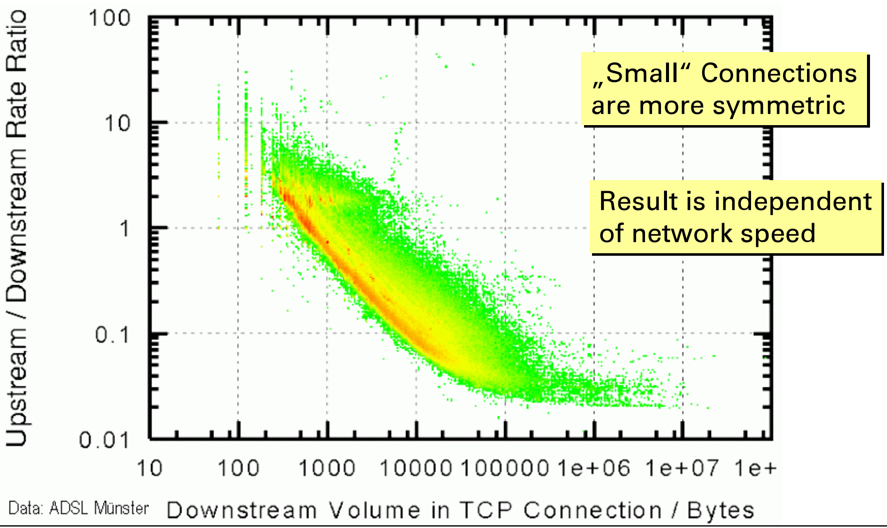
WWW Access
Bit Rate Symmetry Distributions



Data: ADSL Münster

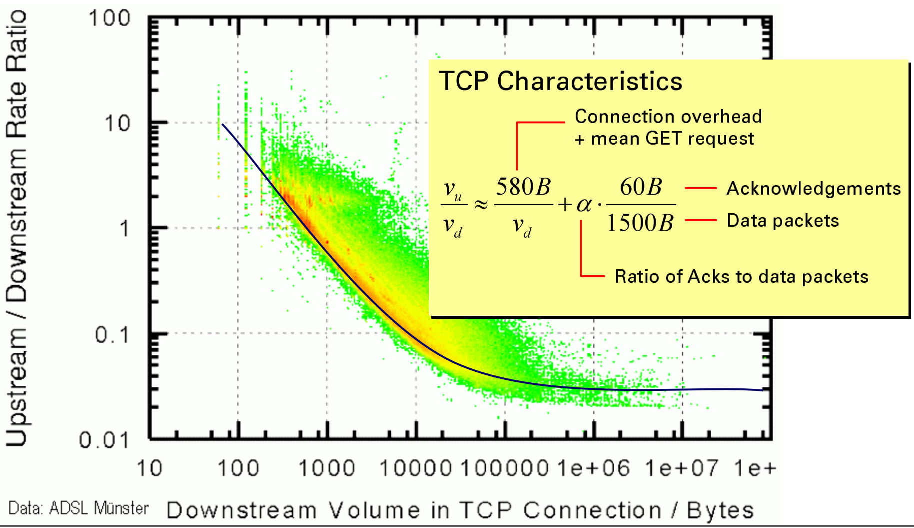
SIEMENS

WWW Access
Bit Rate Symmetry

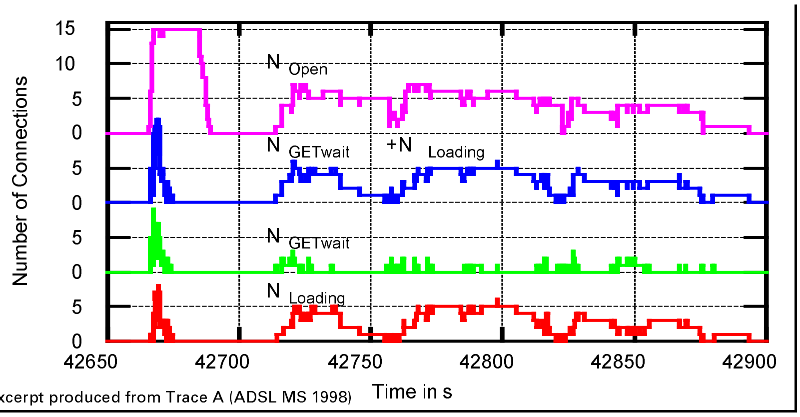


Data: ADSL Münster

WWW Access
Bit Rate Symmetry

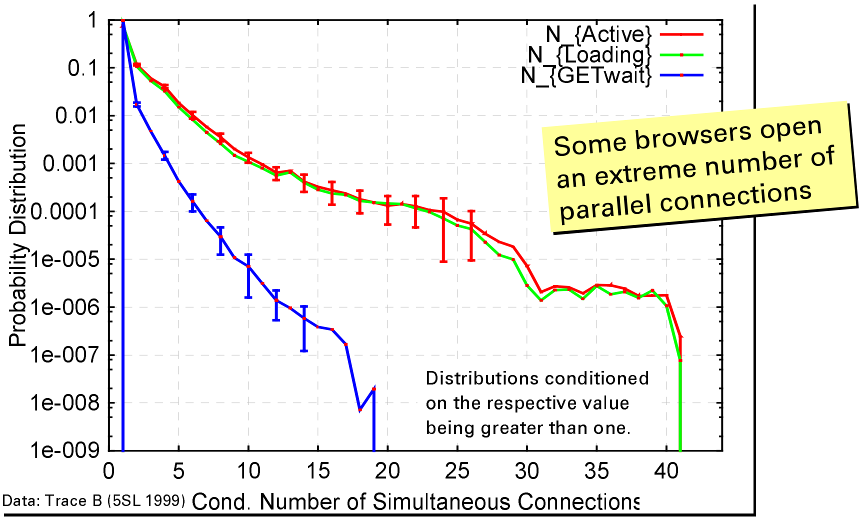


WWW Access
Parallel Connections



- Parallel connections are either idle or loading
- only a few wait for data

WWW Access
Number of Parallel Connections



Data: Trace B (5SL 1999)

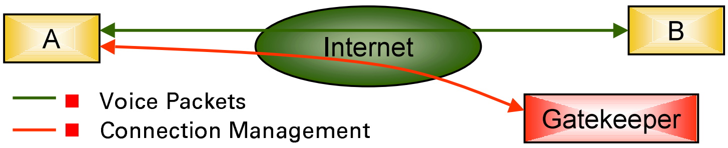
Multimedia Applications

- Real-time streams
  - Internet Telephony
  - Video Conferencing
- Non-real-time streams
  - audio/video streaming
  - e.g. real audio/video (see next slides)
- Stream traffic + additional control traffic
- examples here:
  - traffic characteristics VoIP
  - real audio measured traffic

SIEMENS

Internet Telephony

- packetized voice, e.g.
  - ITU G.723 (10–20kbit/s, with silence suppression -> nearly ON-OFF)
  - G711 (64kbit/s uncompressed)
  - GSM standard
- transmission via RTP/UDP
- connection management using ITU H.323 or SIP
- Gatekeeper for
  - address translation
  - connection admission control (if applicable)



SIEMENS

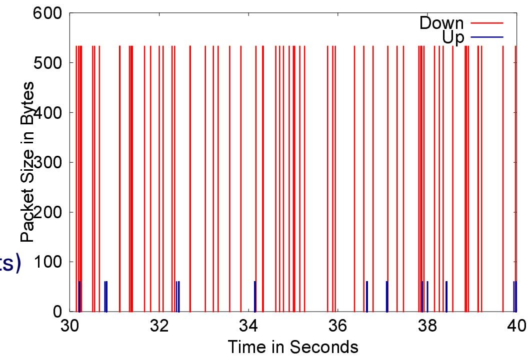
Real Audio

- Transfer of audio clips in the Internet
  - CD clips
  - live concerts
  - speech (newscasts)
- Ports
  - control
  - data
- rate can be adapted to local connection
  - choice e.g. between 14/28/56/112k modems / T1 / LAN
  - large buffers used to compensate for delay variation and packet loss

SIEMENS

Real Audio
Traffic Characterization

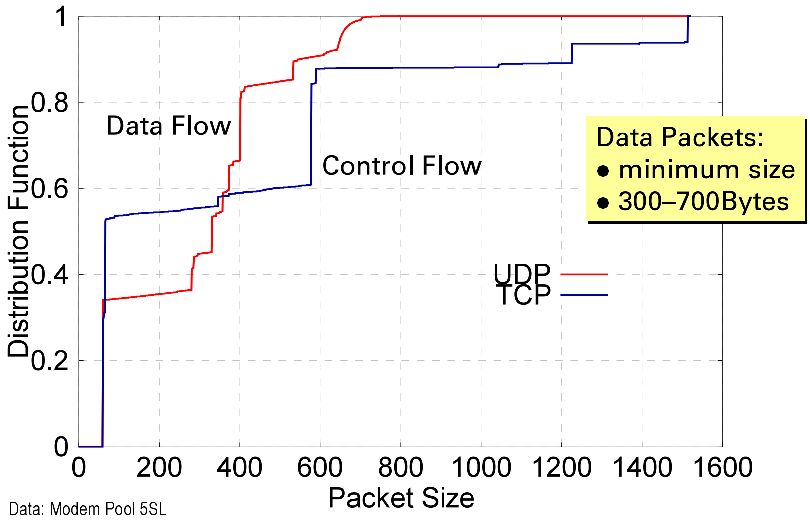
- Data stream
  - UDP
  - relatively constant packet size of 280–550 Bytes
- Data Control stream
  - UDP
  - intra-stream control (during UDP stream)
- Media Selection
  - TCP
  - selection (before UDP stream starts)



Data: Modem Pool 5SL

SIEMENS

Real Audio
Packet Size Distribution



Data: Modem Pool 5SL

SIEMENS

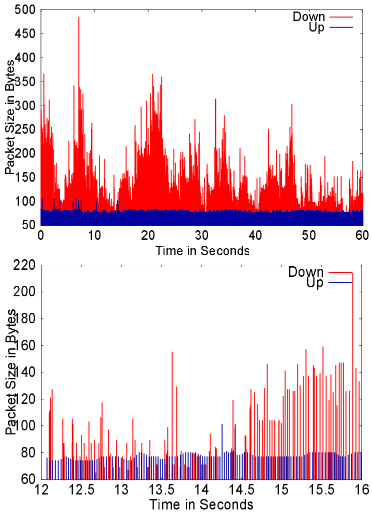
Online Games

- different concepts
  - complete download of executable game (e.g. Java applet)
  - user plays against server
  - direct interaction between multiple users
- different games and players
  - periodic state updates vs. transmission of state changes
  - action vs. thinking games
  - activity patterns of individual users (“hunt” vs. “wait”)
- interactive games: UDP traffic
  - extreme delay requirements (few ms)
  - high background load (small packets at high rate)
  - additional traffic peaks
  - often very long usage (many hours)
  - mixed distributions of packet interarrival time and packet length

SIEMENS

Online Games
Example

- Data measured in ADSL field trial on 29.Aug. 1998
- 00:00 until 01:45
- upstream
  - 20kbit/s rather constant rate
  - 30–35 packets per second
- downstream
  - 30–40kbit/s variable
  - 30–35 packets per second
  - packet size and bit rate peaks



SIEMENS

Other Applications

Wide Area Networks
- Domain Name System (DNS)
- Network News (nntp)
- Routing Protocols

Corporate Networks
- File Services
  - e.g. Network File System (NFS)
- Printer services
- Database access
- Specific Corporate Applications

- Network Management
- Backup Services

Information and Communications

SIEMENS

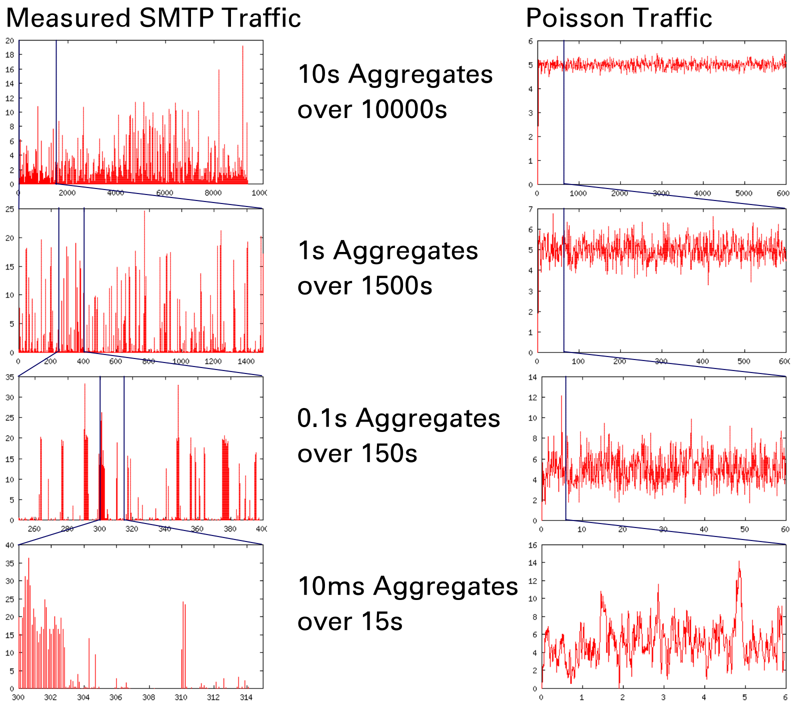
Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

Information and Communications

4. Self-Similarity

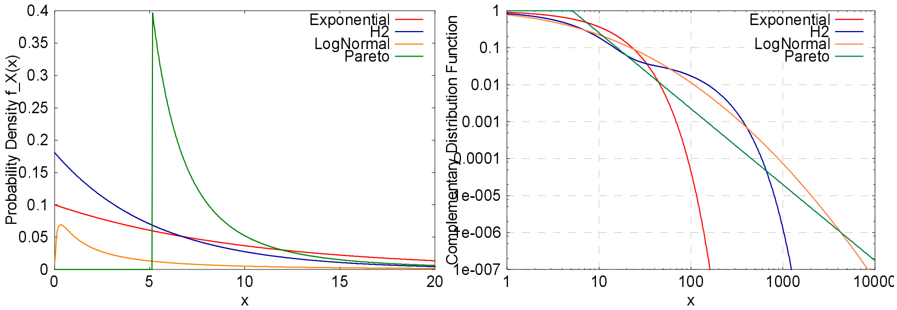
- Scaling Example
- Heavy-tailed distributions, power tails
- Hurst Parameter
- Construction
- Influences
- Consequences
- Multifractals



SIEMENS

Probability Distributions
Examples

- Same mean = 10
- Coefficient of variation c_V=3 (except for exponential distribution)
  - Exponential: λ=0.1
  - Hyperexponential: p_1=0.947; λ_1=0.189; p_2=0.053; λ_2=0.0106
  - Lognormal: σ^2=2.3; μ=1.15
  - Pareto: x_0=5.132; α=2.054



SIEMENS

Power Tails

- The complementary distribution function C(x)=P{X>x} decays like

  $$C(x) \sim x^{-\alpha} \cdot L(x)$$

  for a slowly varying L(x) as x $\to \infty$

Alpha	Distribution	Mean	Variance
≤ 0	Invalid	–	–
$\in (0,1]$	Valid	∞	∞
$\in (1,2]$	Valid	Finite	∞
>2	Valid	Finite	Finite

- very large values of X occur with a much higher probability than expected from "usual" distributions

Power Tails
Example

- Negative exponential distribution

  $C_X(x) = P\{X > x\} = e^{-\lambda x}, \quad x \geq 0$

- Pareto distribution

  $C_X(x) = P\{X > x\} = (x_0/x)^{\alpha}, \quad x \geq x_0$

- Common mean value = 10

Distribution Type	Neg. Exp.	Pareto	Pareto
Parameters	λ=0.1	α=4; x_0=7.5	α=1.5,x_0=3.3
P{X>10}	0.37	0.32	0.07
P{X>100}	5e-5	3e-5	6e-3
P{X>1000}	4e-44	3e-9	2e-4

Self-Similarity

- Variance does not decrease as fast as expected
  - on time scale aggregation
  - on ensemble aggregation (multiple sources)
- "fractal" or "self-similar" characteristics
  - mean bit rates over time
  - mean packet rates over time
- due to heavy-tailed distributions of ON phases
  - causing long-range dependence

Large File Size Variance → Long-Range Dependence → Self-Similarity

- Limits
  - packet level time resolution
  - instationarity

  -> check time dependence of traffic parameters

SIEMENS

Self-Similarity
Different Views

- Dependence of variance on aggregation time
  - Hurst Parameter H

$$Y_t^m = \frac{1}{m} \sum_{s=m(t-1)+1}^{mt} X_s \longrightarrow VAR(Y_t^{(m)}) \sim m^{-2(1-H)}$$

- Long-Range Dependence
  - autocorrelation function decays with $k^{2(H-1)}$
  - Hyperbolic instead of exponential decay of autocorrelation

$$\rho(k) = Cov(X_t, X_{t+k}) \sim k^{2(H-1)} \qquad k \to \infty$$

- Spectral Density
  - pole at zero

$$f(\lambda) = \frac{\sigma^2}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ik\lambda} \sim \lambda^{1-2H} \qquad \lambda \to 0 \;;\; \lambda \in [-\pi, \pi]$$

SIEMENS

Self-Similarity
Estimating the Hurst Parameter

- Variance-Time Analysis
  - plot variance of aggregate versus aggregation time
  - simple, easy to understand
  - also gives second (variance) parameter
  - slightly unreliable
- R/S Analysis
  - classical approach for unknown mean and variance
  - plot rescaled adjusted range versus interval length
- Periodogram Analysis
  - shows increase of spectral density at zero
- Abry-Veitch Estimator
  - using wavelet theory
  - independent of stationarity
  - determines H and variance parameter from regression of Wavelet coefficients

Self-Similarity
Construction

- ON/OFF sources
- Poisson arrivals
- Pareto distribution of ON duration B

$$P\{B \le x\} = 1 - \left(\frac{x_0}{x}\right)^{\alpha}$$

- B has infinite variance and finite mean if $\alpha \in (1,2]$
- Resulting traffic is self-similar with

$$H = \frac{3-\alpha}{2}$$

Influences on Self-Similarity

- File size distribution
  - main cause
  - Heavy tail creates self-similarity
- Idle time distribution
  - also relevant, further increases H in certain cases
- Mixing traffic flows with different H
  - resulting H is somewhat interpolated
- Network Topology
  - no significant influence
- Protocol Stack
  - TCP (congestion and error control) modulates H
- Network performance decreases smoothly with increasing H
  - queue lengths are more sensitive

from Park/Kim/Crovella in Park/Willinger 2000

SIEMENS

Self-Similarity
Consequences of Long-Range Dependence

- Huge buffers needed for unelastic traffic
  - Burst level queueing is inefficient
  - Only link capacity can help (lower utilization)
- Elastic traffic not modelled correctly by stiff self-similar traffic
  - TCP adapts to available bit rates and prevents excessive packet loss
- Successive measurements not independent
  - Huge confidence intervals for measurements and simulation
  - To reduce confidence interval by a factor of 10, measurement must be $10^{1/(1-H)}$ longer
  - Consider this in traffic measurement and simulation!

SIEMENS

Multifractal Scaling

- Different scaling behavior observed on different time scales
- Scaling is not described exactly by monofractal law and one (Hurst) parameter
- Different causes on different time scales
  - Application Layer (file size distribution, connection arrivals)
  - TCP layer (flow&error control, timers, round trip time)
  - Media access

SIEMENS

Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

Information and Communications

SIEMENS

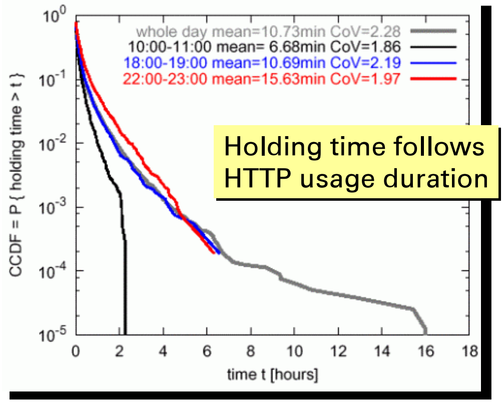
5. User Behaviour Measurements

- Dial-up traffic
- Influence of tariffing
- Applications usage
- Volume versus usage time
- Asymmetry

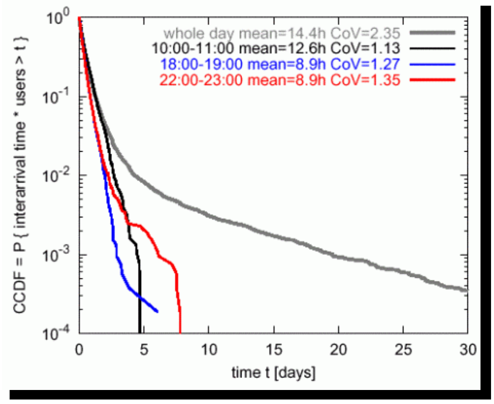
Information and Communications

Access Session Characteristics Distributions

Holding Time

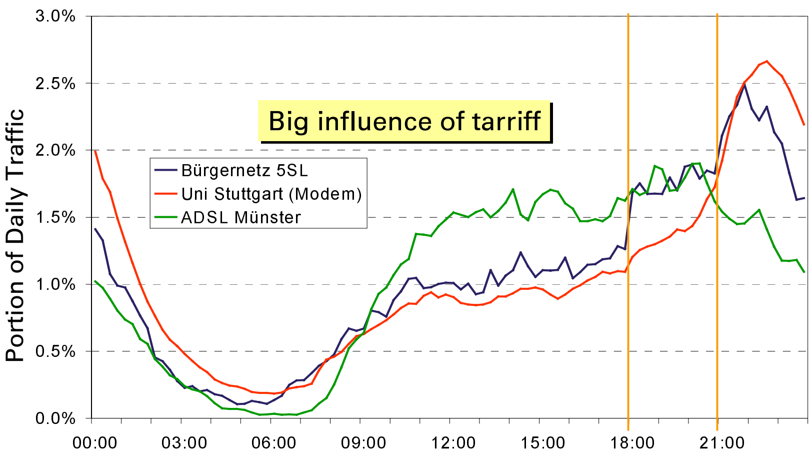


Cumulative Interarrival Time per User



Source: University of Stuttgart

Dialup Behaviour Daily Traffic Curves

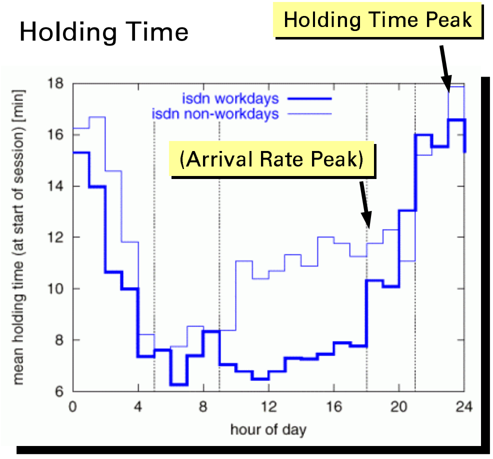


Data: 5SL, Uni Stuttgart, ADSL Münster: 15min values

SIEMENS

Access Session Characteristics
Daily Variation (2)

Holding Time



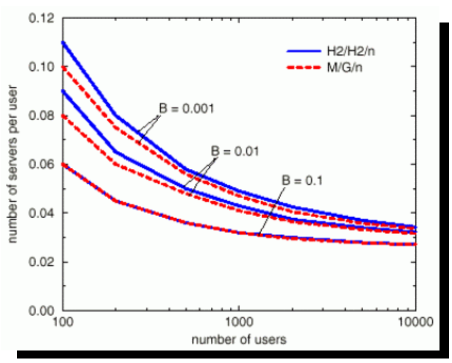
Few Long
Late Night Sessions

Source: University of Stuttgart

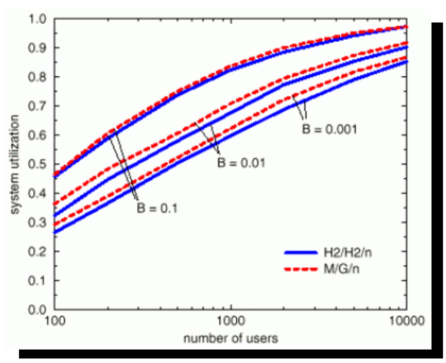
SIEMENS

Access Session Characteristics
Dimensioning Examples

Economy of Scale



System Load

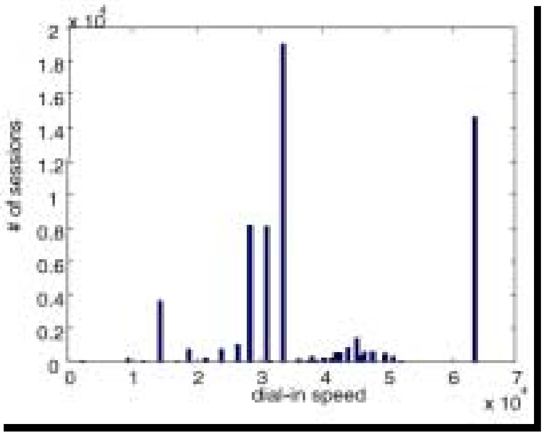


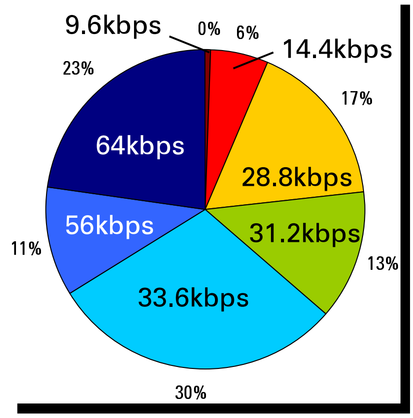
H2/H2/n-0 more realistic than M/G/n

Source: University of Stuttgart

SIEMENS

Achieved Dialup Access Speeds

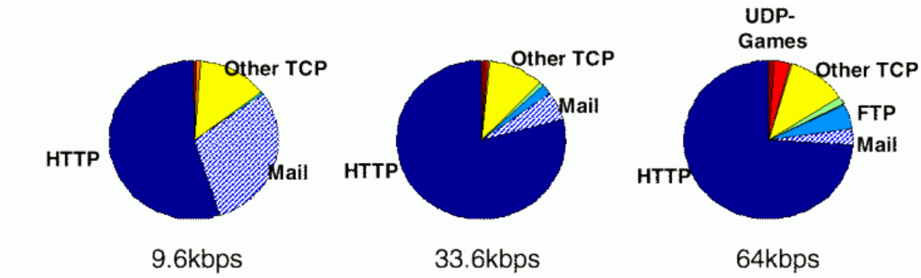




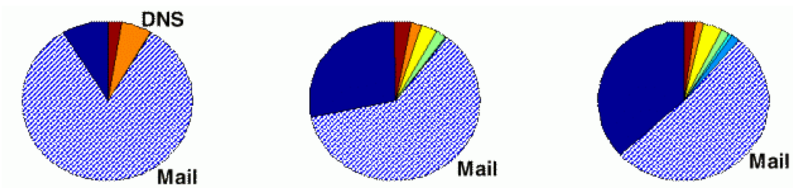
Source: University of Würzburg

SIEMENS

Application Mixes



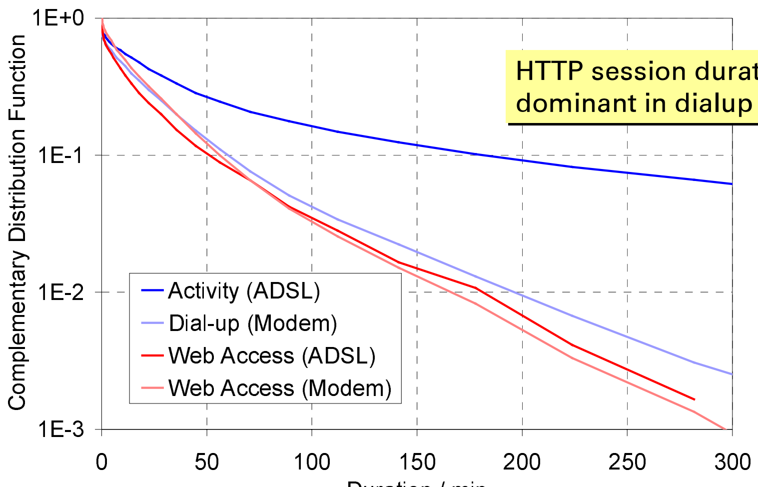
... in sessions < 100sec:



Source: University of Würzburg

SIEMENS

Dialup Behaviour
Session Duration (1)

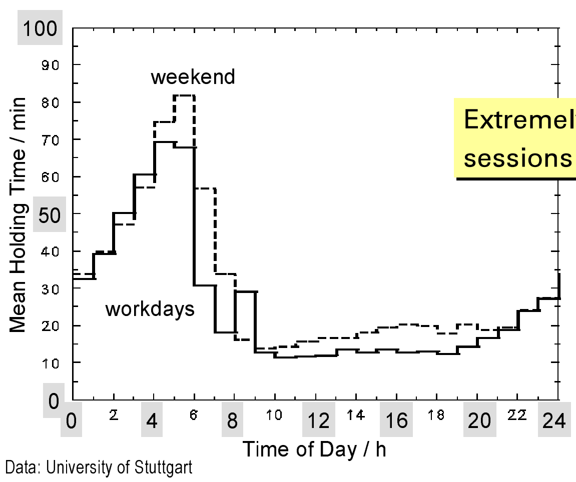


HTTP session duration is dominant in dialup session

Data: 5SL, ADSL Münster

SIEMENS

Dialup behaviour
Session Duration (2)

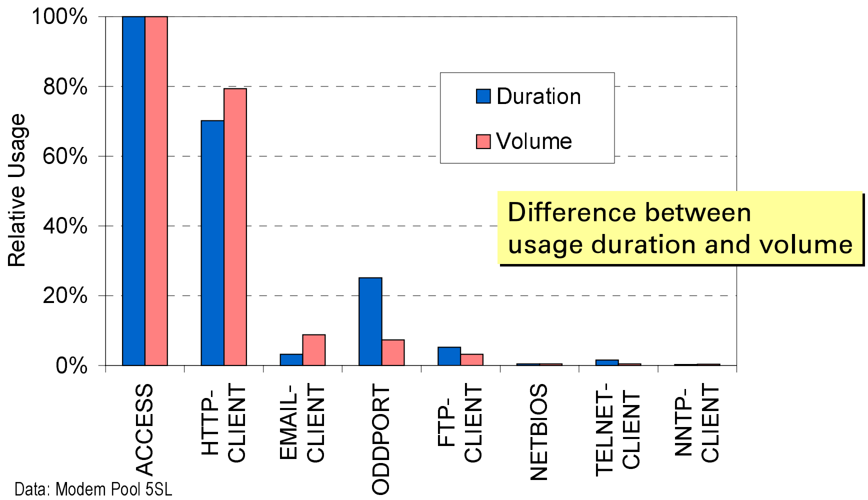


Extremely long dialup sessions during the night

Data: University of Stuttgart

SIEMENS

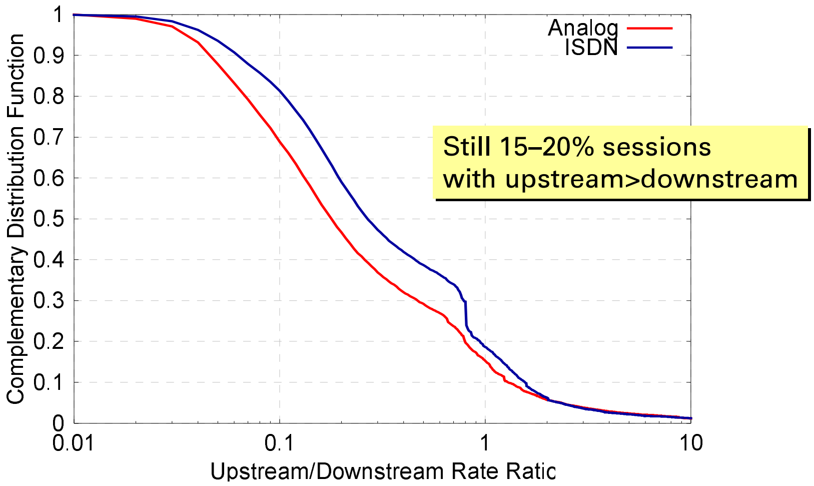
Applications Usage Volume vs. Duration



Data: Modem Pool 5SL

SIEMENS

Overall Access Traffic Asymmetry



Data: 5SL

SIEMENS

Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

SIEMENS

6. Backbone Measurements

- Traffic mixes (Protocols, Applications)
- Periodic Changes
- Symmetry Relations
- Internet Network Characteristics
- Flow aggregation characteristics

SIEMENS

Traffic Mixes

In addition to local traffic:

- Domain Name System (DNS)
- Network News (nntp)
- Routing Protocols
- Network Management

- CAIDA (www.caida.org) March 2000 (5 min @22:07)
  - 30% HTTP
  - 25% FTP data
  - 13% other TCP
  - 6% Napster
  - 6% Squid Web Cache
  - 5% SMTP
  - 5% Liquid Audio
  - and many more ...

SIEMENS

Periodic Changes and Symmetry

- Daily patterns
  - less activity during early morning (2:00–7:00)
  - prime time during the day (10:00–18:00) or evening (depends on ISP)

- Weekly patterns
  - less activity on Saturday / Sunday

- Asymmetry on international links varies during the day
  - determined by client/server relations
  - mostly export of documents from U.S.

SIEMENS

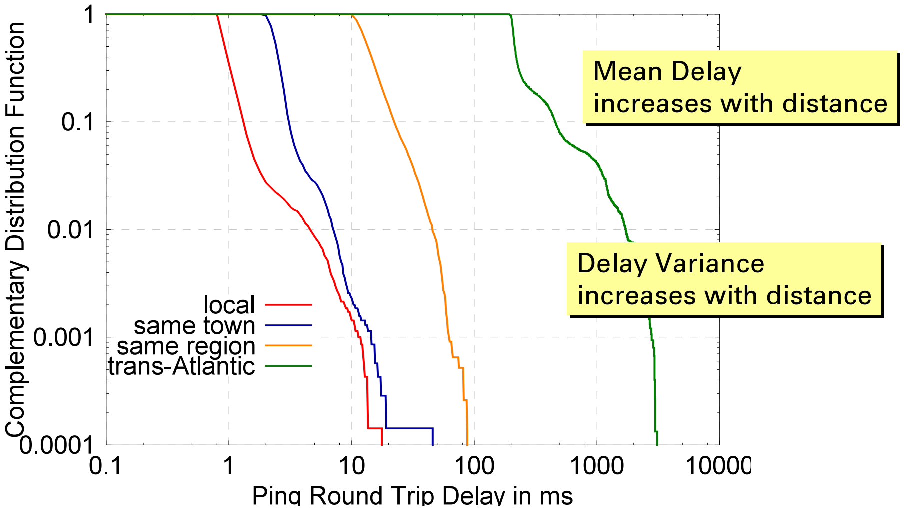
Network Characteristics

Paxson's Measurements show

- Internet connections are mostly
  - "busy" (relatively high packet loss rate of 4–20%) or
  - "quiescent" (no packet loss)
- Consecutive packets mostly take the same path
  - reverse path may be different

SIEMENS

Network Characteristics
Delay Distributions



SIEMENS

Flow Characteristics

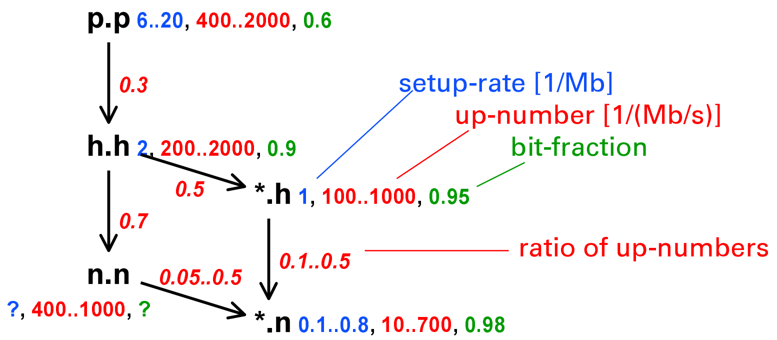
- Caution!
  - Flow volume and duration have heavy-tailed distributions
  - expected duration may be infinite
  - volume variance may be infinite
  - mean values are very hard to measure correctly

Normalized characteristics allow comparison:

- Setup rate
  - (new flows per second) / (total traffic bit rate) = 1/(mean flow volume)
- Up-number
  - number of active flows per transported traffic bit rate
- bit fraction
  - fraction of traffic that could be transported in a shortcut

SIEMENS

Flows
Aggregation Properties



- bit-fraction improves significantly with aggregation
- up-number typically a few hundred shortcuts per Mb/s with strong variation; small aggregation gain
- smaller up-numbers only for *.n shortcuts beginning close to destination network

data deduced by Paul Schlüter from Copeland et al. (1999), Feldmann et al. (1998), Nagami et al. (1999)

SIEMENS

Outline

Information and Communications

SIEMENS

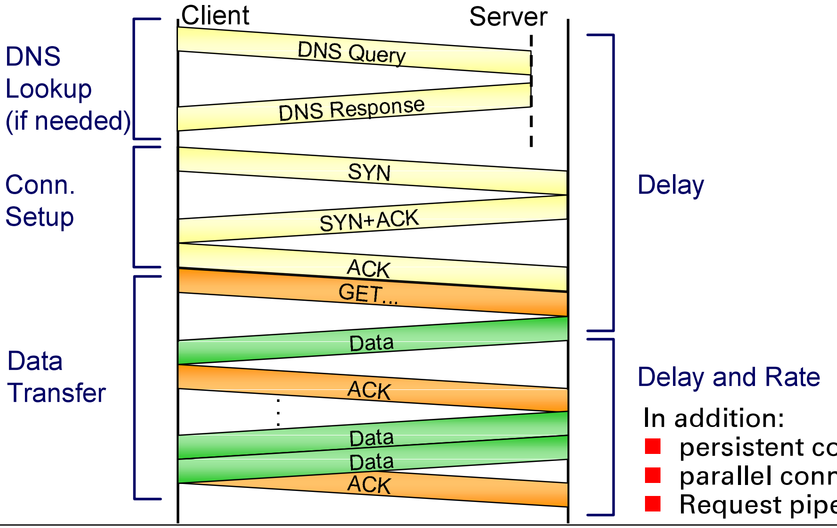
7. Web Performance

- Delay and Rate influence Performance
- Delay Components
- Delay Distributions
- Protocol and Architecture Options

Information and Communications

SIEMENS

Reminder
DNS-HTTP-TCP



In addition:

- persistent connections
- parallel connections
- Request pipelining

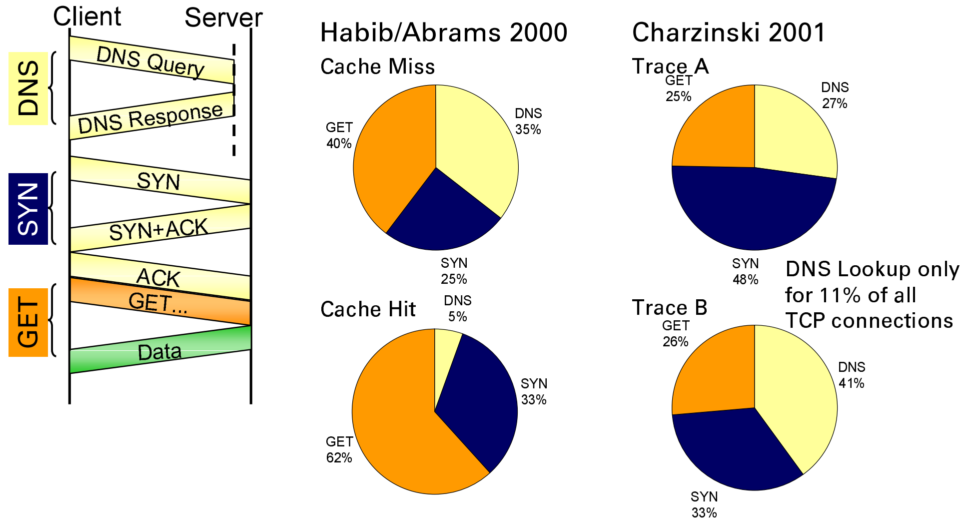
SIEMENS

Web Delay Components

- DNS latency
  - server reaction time
  - retries due to multiple (false) domain name extensions
- Network delay
  - propagation delay
  - processing delay (routing)
  - queueing delay (limited link bandwidth)
- Server reaction times
  - Connection establishment (socket creation, answer to SYN packet)
  - answer to GET request
  - database lookup, page construction time for dynamic pages
- Client reaction times
  - reaction to DNS answer
  - reaction to connection set-up (SYN+ACK)
- Content transmission time (determined by bottleneck bandwidth)
- Others
  - HTTP redirections
  - protocol mismatches (GET causing RST)

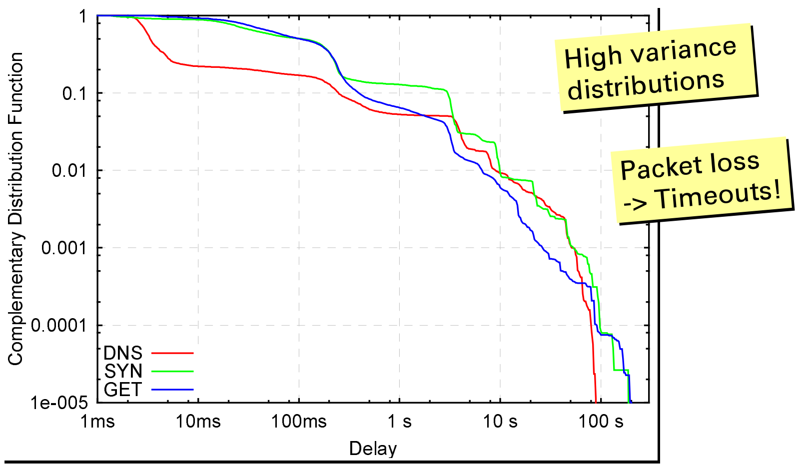
SIEMENS

Web Delay Components
Mean shares during DNS-SYN-GET (for first item)



SIEMENS

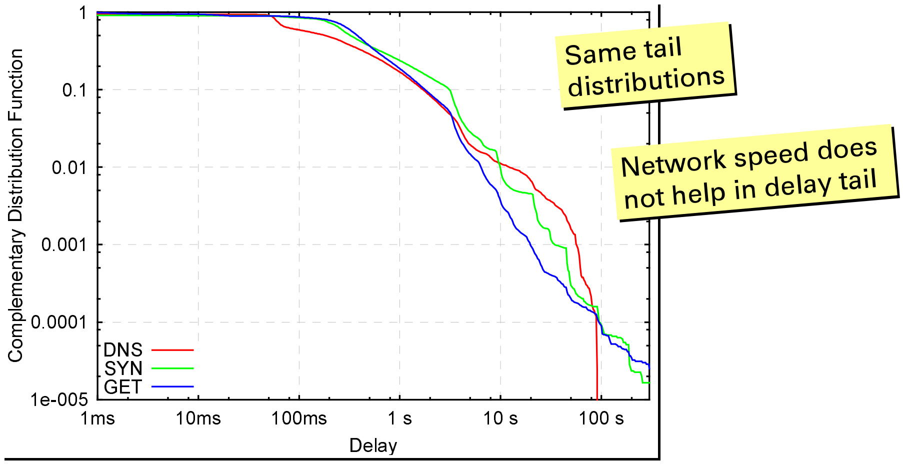
Measured Delay Distributions



Data: ADSL Münster 1998 (Trace A)

SIEMENS

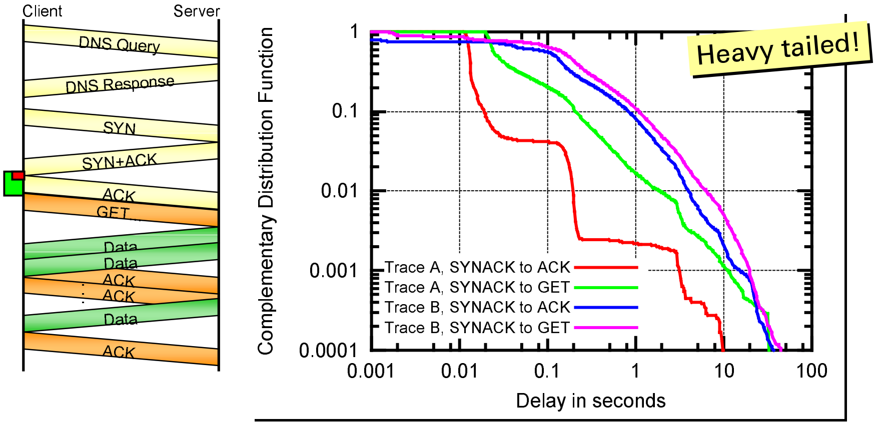
Measured Delay Distributions



Data: 5SL Modem/ISDN 1999 (Trace B)

SIEMENS

Client Response Times

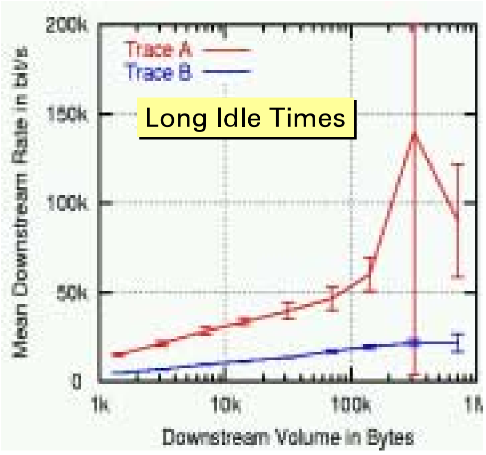


- 10% of clients needed longer than 200ms (1s) to send GET request
  - main reason: parallel connection handling

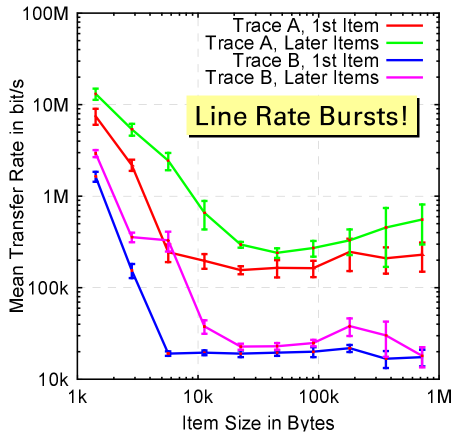
SIEMENS

TCP and HTTP Keep-alive Connections
Mean rates versus volume class

Per HTTP/TCP Connection (P2P Flow)



Per Item (Response to one GET Request)



SIEMENS

Web Performance
General Observations

- DNS lookups can take significant time
- Connection establishment
  - routes and servers show “cold” and “warm” states
- small files:
  - most delay between GET request and start of transfer
  - server load is critical
- large files:
  - most delay during transmission
  - network load is critical (timeouts, fast retransmits)
- All delays show heavy-tailed distributions (!)
- High throughput needs good OS scheduling and I/O performance
  - on both sides

SIEMENS

Protocol and Architecture Options

- Caching
  - validation time can be significant [Krishnamurthy/Wills]
  - does not help with dynamic content
- Persistent connections (HTTP/1.0 or HTTP/1.1)
  - can reduce network load
  - bad if server memory is a bottleneck [Barford/Crovella]
- Request pipelining
  - reduces influence of round-trip times to GET more items
  - problem with servers closing connections
    (unclear client/server interaction) [Krishnamurthy/Wills]
- Browser/Proxy options [Cohen/Kaplan]
  - pre-resolving
  - pre-connecting
  - pre-warming (dummy HTTP HEAD)

Information and Communications

SIEMENS

Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

Information and Communications

SIEMENS

8. Models

- Models on Different Activity Levels
- Multilevel Models
- Traffic Model Validation
- Link Dimensioning Models
- Multiplexing Gain
- TCP Models

SIEMENS

Model Types

- Different models for different levels
  - Layer 3 traffic models to drive lower layer simulations
  - user/application models to drive TCP simulations
  - session level models to drive loss simulations
- User and application models
- Single user / backbone traffic models
- Network models
  - e.g. for TCP behaviour
- Multilevel models
  - e.g. for HTTP users or HTTP traffic

SIEMENS

Input Traffic on Different Activity Levels
Access Session Traffic

Access Session

- Holding time
- Arrival rate (single / multiple subscribers)
- Daily traffic variations
- Tradeoff between holding time and arrival rate
  - important e.g. in ISDN or AO/DI access with extremely short conection set-up times
- Influence of tariffs

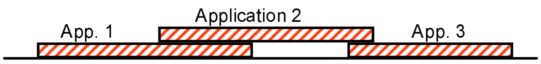
Models

- M/G/n-0 -> Erlang-B formula
- GI/G/n-0

SIEMENS

Input Traffic on Different Activity Levels
Application Session Traffic

Within an Access Session:



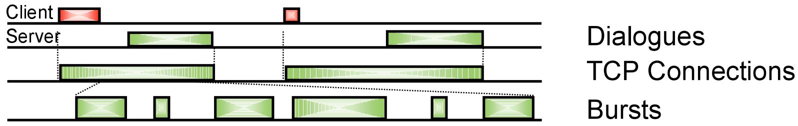
Application Sessions

- Holding time
- Application usage share
  - correlated with access session holding time!
- HTTP: holding time corresponds to dial-up session h.t.
- Application preferences change over the years
  - e-mail -> Web access -> audio streaming
  - video, games, 3D chat, ...
- Applications evolve over the years
  - lynx, Mosaic, Netscape, Internet Explorer
  - animated GIFs, Frames, Java, Javascript, ...

SIEMENS

Input Traffic on Different Activity Levels
Dialogues, TCP Connections, Bursts

Within an HTTP Application Session:



- User actions
- Element size distributions
- number of elements per user request
- connection duration distribution
- parallel connections
- Feedback from Network
- TCP: Interaction with other users’ connections
- Bursts within one connection also due to
  - download of multiple elements
  - database look-up in server
  - TCP flow control

SIEMENS

Input Traffic on Different Activity Levels
Dialogues, TCP Connections, Bursts (2)

- Connection interarrival times
  - Weibull distribution
  - timer driven
- TCP’s flow control adapts to available bit rate
  - Limit measured from packet loss (or excessive RTD)
  - When is a link “correctly dimensioned”?
  - Packet traffic traces cannot be re-used in another scenario
  - Packet loss is an unsuitable QoS criterion
  - “fun factor” dimensioning
- Models
  - M/G/R-PS, ON/OFF Fluid Flow
  - Markov models for TCP behaviour
  - FBM Fluid Flow

Input Traffic on Different Activity Levels
Packet Traffic

Within a Burst:

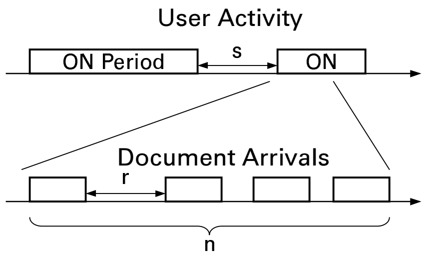
Packets (Client-Server and Server-Client)

- Packet Size Distribution
- Packet Interarrival Time
- Correlations
- Upstream/Downstream Correlations due to TCP
  - Media Access Control issues
  - Time Division Duplex issues

Multilevel Models
Shuang Deng, ICC'96

- ON-OFF model
- Single HTTP user
- Upstream or downstream traffic
- Burst Level

- Duration n of ON period
  - Weibull
- Duration s of OFF period
  - Pareto (infinite mean)
- Interarrival time r
  - approx. Weibull

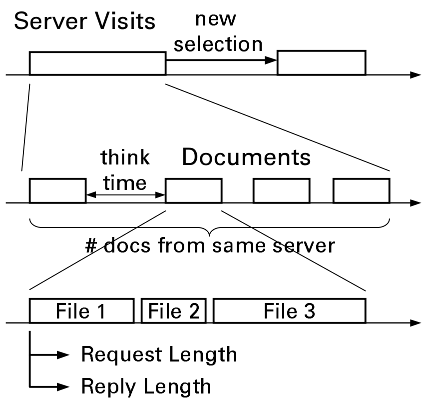


- Unclear: infinite mean OFF time

SIEMENS

Multilevel Models
Bruce Mah, Infocom'97

- Session and Burst Level
- Single HTTP user
- Upstream and downstream traffic
- Request Length
  - bimodal
- Reply Length
  - Pareto, $\alpha \approx 1.04 – 1.14$
- Number of files per doc.
- Think Time
- Number of docs per server
- Server Selection
  - Zipf's Law

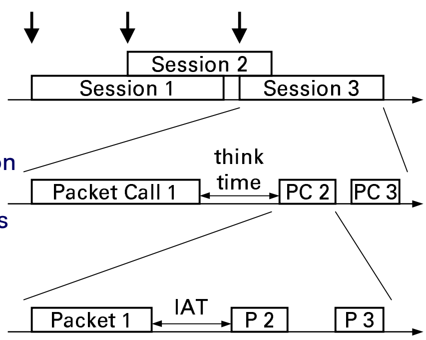


- Think Time includes idle time between sessions -> extremely long!

SIEMENS

Multilevel Models
"ETSI/3GPP UMTS" Traffic Model

- Session arrivals (Poisson Point Process)
- "Packet Calls"
  - number of packet calls per session (geometric)
  - reading time between packet calls (geometric)
- Packets
  - number per packet call (open)
  - packet IAT (geometric/open)
  - packet size (cut-off Pareto)

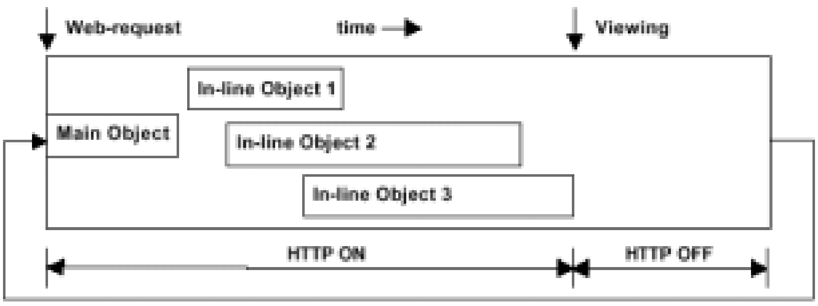


- Packet Call length instead of packet size should be Pareto

SIEMENS

Multilevel Models
Choi and Limb, ICNP'99

Choi/Limb 1999

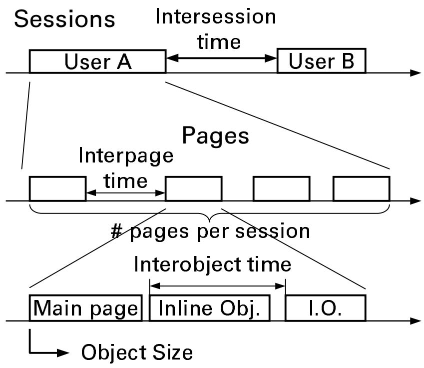


- parallel connections
- detailed structure of HTTP downloads

SIEMENS

Multilevel Models
Feldmann et al, Sigcomm 1999

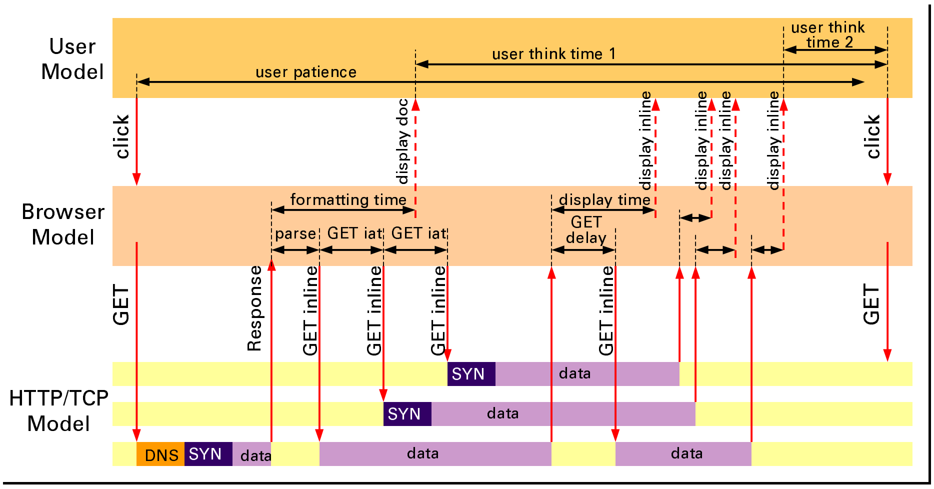
- Session and Burst Level
- used to drive TCP simulations
- Weibull or timer-driven connection interarrival times



- Good description of parameters

SIEMENS

Multilevel Models
Färber 2001 (unpublished)



SIEMENS

Multilevel Models
Parameters

- Many parameters but easy to understand
- How to determine parameters?
  - Often difficult to extract even from packet trace measurements
- Parameters depend on other constraints
  - network speed
  - computer (client / server) speed
  - delays
  - tariffs
- Contrast: Multi-fractal models
  - canonical set of parameters (per time scale octave)
  - measurement based
  - with little physical meaning

SIEMENS

Traffic Model Validation

- Equivalence of generated traffic:
  Check statistics
  - size / interarrival time distributions
  - correlation structure
- Validation within system/network model of interest:
  Check model against measured trace in simulation
  - same queue length / delay statistics or
  - same loss probability
- don't just
  - generate optically equivalent traffic
  - postulate user behaviour

SIEMENS

Link Dimensioning

- Bufferless models
  - burst scale approximation
  - assumption: enough buffer for packet scale multiplexing
  - no burst buffering capability
- Buffer models
  - take burst scale buffering into account
  - inefficient with long-range dependent traffic
- unelastic models
  - take traffic as inevitable
  - compute loss probability (and packet delay)
- elastic models
  - include TCP's flow control
  - compute total delay e.g. for transfer of a given file

Worst-Case (low loss) models
Rate Envelope Multiplexing (REM)

- Make sure that the considered link is not a bottleneck
  - make the loss rate on this link low enough
  - -> TCP will not be influenced by this link
  - use bufferless burst scale model to make sure small packet scale buffers will suffice
- Example: ON-OFF Rate Envelope Multiplexing
  - n ON/OFF sources with ON probability β and rate r_{ON}

$$p_{loss} = \sum_{i=\lceil c_L / r_{ON} \rceil}^{n} \binom{n}{i} \beta^i (1-\beta)^{n-i} \frac{i\, r_{ON} - C_L}{\rho C_L}$$

Norros' Fractional Brownian Motion Model

- Input traffic is FBM with drift
  $A_t = mt + \sqrt{ma} Z_t$
  - mean rate m, Hurst parameter H, variance coefficient a
- Captures long-range dependence
  - buffer can be dimensioned including LRD effects
  - several approximations reduce accuracy of results
  - rate dimensioning results do not differ significantly from approaches that take the variance into account but ignore LRD

$$C = m + (H^H (1-H)^{1-H} \sqrt{-2\ln\varepsilon})^{\frac{1}{H}} a^{\frac{1}{2H}} x^{\frac{H-1}{H}} m^{\frac{1}{2H}}$$

SIEMENS

M/G/r-PS
Processor Sharing Model

- Input traffic
  - Poisson arrivals of fluid ON/OFF traffic streams
  - arrival rate λ, maximum ON rate r_{ON}
  - generally distributed (e.g. heavy-tailed) ON volume
- TCP flow control modeled by Processor Sharing discipline
  - simple approximation for TCP behaviour on a single link
  - result: (time) mean of delay factor f_R

$$f_R = 1 + \frac{1}{r(1-\rho)} \frac{\frac{(r\rho)^{r_g}}{r_g!}}{(1-\rho)\sum_{i=0}^{r_g-1}\frac{(r\rho)^i}{i!} + \frac{(r\rho)^{r_g}}{r_g!}}$$

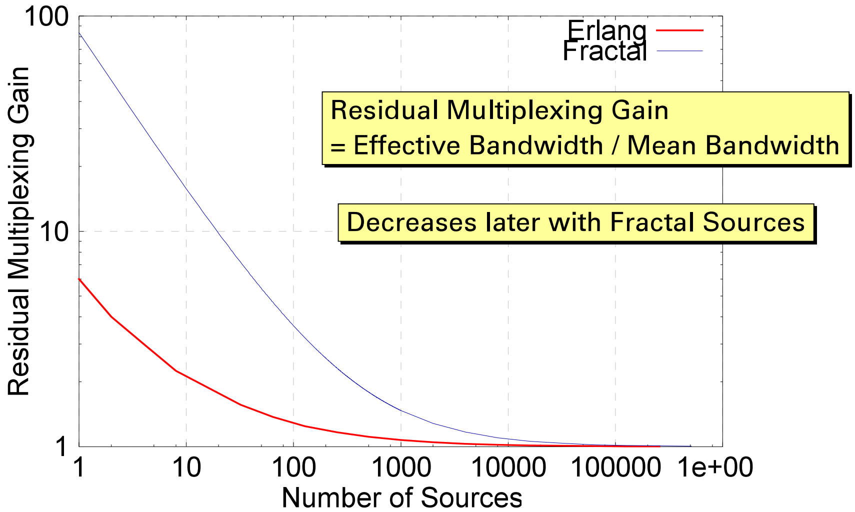
SIEMENS

Multiplexing Gain

- In order to achieve the same level of QoS, less resources are needed when traffic is aggregated.
  - Packet (cell) level: "Multiplexing Gain"
  - Connection (call) level: "Economy of Scale"
- gain can be achieved as long as there is variance in the traffic
  - Self-similar traffic also shows multiplexing gain
  - Self-similar traffic still shows multiplexing gain at high aggregation levels

SIEMENS

Multiplexing Gain
Example



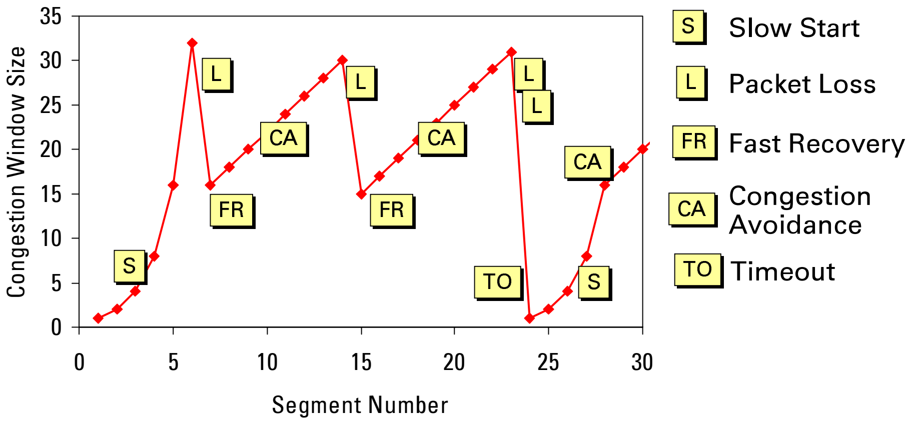
SIEMENS

TCP Models

- Analytical models include relevant system states
  - basic idea: get stochastic distribution of system state occupancy and derive other measures (e.g. throughput) from that
- CWND
  - congestion window
  - sender can send CWND segments until acknowledgement is needed
- Threshold
  - CWND value at which exponential increase ("slow start") is replaced by linear increase ("congestion avoidance")
- Often additional assumptions
  - greedy source
  - fresh connection
  - independent packet losses
- There are different TCP versions around!

TCP Models
General Behaviour

- TCP Reno trace



TCP Models
Main Result

- Steady state bandwidth [Padhye, Firoiu, Towsley, Kurose 1998]

$$B \sim \frac{1}{RTT} \cdot \sqrt{\frac{3}{2 \cdot b \cdot p}}$$

B achievable throughput
RTT round trip time
p packet loss probability (low)
b number of acknowledget packets per ACK received

- fast decrease for higher packet loss rates
- lower throughput for long round trip times

SIEMENS

Outline

1. Introduction
2. User and Application Behavior
3. Application Behavior Measurements
4. Self-Similarity
5. User Behavior Measurements
6. Backbone Measurements
7. Web Performance
8. Models
9. Implications for Simulation
10. Implications for Quality of Service

SIEMENS

9. Implications for Simulation

- Steady State and Confidence
- Infinite Expectations
- Input Parameters
- Deterministic Scenarios

Steady State and Confidence

Effects of long-range dependent traffic

- steady state reached slowly
  - stochastic generators (input processes!)
  - observed system state (e.g. queue length)
- High variability at steady state
  - high probability of "swamping" observation
- Standard deviation of batch means decreases slowly
  - To reduce batch means standard deviation by a factor of 10: simulate factor of $10^{1/(1-H)}$ longer
  - H=0.5: factor 100 longer
  - H=0.9: factor 10 000 000 000 longer!

Infinite Expectations

- ... can never be simulated ☺
- M/G/1 has infinite expected waiting time if the "G" has infinite variance
  - Mean Waiting Time $E[W] = E[S]\frac{\rho(1+c_S^2)}{2(1-\rho)}$
  - Residual Lifetime $E[R] = \frac{c_x^2+1}{2}E[X]$
- model carefully
  - consider packets instead of bursts or
  - for TCP, use M/G/r-PS or M/D/1-PS instead of M/G/1
- limit distributions
  - check validity of assumption (e.g. do simulation results change with limit?)
  - introduce corresponding mechanisms into networks (e.g. limit on e-mail sizes)

SIEMENS

Input Parameters

- Don not use the Normal distribution
  - Finite probability for X<0
- Use input parameters that have a meaning
  - and make sure the corresponding random variables have finite mean
- TCP traces are generally invalid
  - if simulation includes TCP model -> use file sizes
  - if simulation does not include TCP -> only binary result possible
    - YES: the simulated network does not disturb TCP
    - NO: the simulated network disturbs TCP and results will be fundamentally different
- The “mean packet size” is generally uninteresting
  - Packet sizes have multimodal distributions

SIEMENS

Deterministic Scenarios

- Be careful not to simulate trivial scenarios ad infinitum
- Ensemble statistics vs. single source statistics
- Applications:
  - Voice over IP on packet level
  - other constant rate sources
- Solutions
  - in simple models: identify period and change phase cyclically
  - use phase changing generators
  - use frequency shifted generators

SIEMENS

Outline

Information and Communications

SIEMENS

10. Some Remarks on QoS

- Traffic Types
- QoS for Elastic Traffic?
- Problems of Admission Control for Web Traffic

Information and Communications

SIEMENS

Traffic Types

- Elastic Traffic
  - transported over TCP
  - given amount of data are to be transmitted "as fast as possible"
  - error control with automatic repetition (ARQ)
  - utilization of maximum available bandwidth (without excessive loss)
  - Classical (stream oriented) traffic descriptors (Leaky Bucket) are invalid

→ Network bandwidth determines "fun factor"

- Stream Traffic
  - given bit rate
  - using e.g. RTP/UDP as transport protocol
  - delay, loss and delay variation to be limited
  - blocking to be limited

→ Network bandwidth determines usability

SIEMENS

QoS for Elastic Traffic?

- Web access is an important service in IP networks
- reason for investing in IP access (besides e-mail)
- fundament for other modern services
  - e-commerce, e-business, etc

- Our performance picture of the Internet is determined by the Web!

→ Web performance is an important issue

- QoS should not only be considered for streaming applications

SIEMENS

QoS for Elastic Traffic

- Problem: compliance with bandwidth sharing model
- Use e.g. congestion marking/pricing

- Guaranteeing QoS to elastic traffic needs admission control
- ... and effective pricing!

- Flat rate and no admission control leads to highway situation
  - congestion
  - priority helps you get faster to the next point of congestion [P.Key]

SIEMENS

Problems of Admission Control with Web Browsing

Admission Control on a per-TCP connection basis:

- wrong unit of transfer
  - multiple requests per connection with pause in between
- penalizes many short connections instead of one long one
  - long conn. can restart transmission at arbitrary byte number in HTTP/1.1
  - average size is more variable than arrival rate
- application view
  - bad QoS even for admitted “flows” if one out of n connections for loading a page is blocked
  - basis for admission control should be multi-connection transaction rather than single TCP connections
  - problems with predictability and describability of traffic

References

General Literature

D. Comer, "Internetworking with TCP/IP", 4th. Ed., Prentice Hall, Englewood Cliffs, NJ, USA, 2000

S. Floyd, V. Paxson, "An Overview of Internet Engineering, Measurement, and Modeling", Tutorial International Teletraffic Congress ITC-15, Washington D.C., USA, June 1997

J.W. Roberts, "Traffic Theory and the Internet." IEEE Communications Magazine 39'(1) 2001, pp. 94–99

Internet growth and traffic figures

As statistics for Internet growth and Internet traffic have a very short life, we mostly give URLs here. URLs themselves often also have a relatively short life, but they still live somewhat longer than the data they point to.

www.nw.com or http://www.isc.org/dsview.cgi?domainsurvey/index.html

http://nis.nsf.net/statistics/nsfnet/

http://www.ripe.net/statistics/index.html

http://www.dfn.de/win/allinfo/statistik/

V. Paxson, "Growth Trends in Wide-Area TCP Connections", IEEE Network 8(4)1994, pp.8-17

K.C. Claffy, H.-W. Braun, G.C. Polyzos, "Tracking Long-Term Growth of the NSFNET." Communications of the ACM 37(8) Aug. 1994, pp. 34-45

Routing

C. Huitema, "Routing in the Internet", Prentice-Hall, 1995

V. Paxson, "End-to-End Routing Behavior in the Internet." Proc. ACM SIGCOMM'96, Stanford, CA, USA, Aug. 1996, pp. 25–38. Available at ftp://ftp.ee.lbl.gov/papers/routing.SIGCOMM.ps.Z

C. Labovitz, G.R. Malan, F. Jahanian, "Internet Routing Instability." Proc. ACM SIGCOMM´97, Cannes, France, Sep. 1997, pp. 115–126

Backbone Traffic

J. Apisdorf, K. Claffy, K. Thompson, R. Wilder, "OC3MON: Flexible, Affordable, High-Performance Statistics Collection." Online Proceedings of Inet97, Kuala Lumpur, Malaysia, June 1997. Available at http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/F1/F1_2.HTM

R.L. Carter, M.E. Crovella, "Measuring Bottleneck Link Speed in Packet-Switched Networks." Boston University, CS Dept, Boston, MA 02215 Technical Report BU-CS-96-006, Mar. 1996. Available at ftp://cs-ftp.bu.edu/techreports/96-006-measuring-bottleneck-link.ps.Z

K.C. Claffy, H.-W. Braun, G.C. Polyzos, "A Parametrizable Methodology for Internet Traffic Flow Profiling." IEEE JSAC 13(8) Oct 1995, pp. 1481-1494

K. Claffy, G. Miller, K. Thompson, "The nature of the beast: recent traffic measurements from an Internet backbone." Proceedings of the INET'98 Conference, April 1998. Available at http://www.caida.org/Papers/Inet98/index.html

J.A. Copeland, R. Abler, K.L. Bernhardt, "IP Flow Identification for IP Traffic Carried over Switched Networks." Computer Networks and ISDN Systems 31(5), 1999, pp. 493-504

Feldmann A, Rexford J, Cáceres R. 1998. Efficient Policies for Carrying Web Traffic Over Flow-Switched Networks. IEEE/ACM Trans. Netw. 6, 673-685

S. Lin, N. McKeown, "A Simulation Study of IP Switching." Comp. Comm. Rev. 27(4) 1997, pp. 15-24

S. McCreary, kc claffy, "Trends in Wide Area IP Traffic Patterns: A View from Ames Internet eXchange." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

K. Nagami, H. Esaki, Y. Katsube, O. Nakamura. Flow Aggregated, Traffic Driven Label Mapping in Label-Switching Networks. IEEE J. Select. Areas Comm. 17, 1999, pp. 1170-1177

V. Paxson, "End-to-End Internet Packet Dynamics." ACM Computer Communication Review 27(4) Oct. 1997, pp. 139-152. Available at ftp://ftp.ee.lbl.gov/papers/vp-pkt-dyn-sigcomm97.ps.Z

K. Thompson, G.J. Miller, R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics." IEEE Network 11(6) Nov./Dec. 1997, pp. 10-23. Available at http://www.vbns.net/presentations/papers/MCItraffic.ps

Application and user traffic and models (mostly HTTP)

M.F. Arlitt, C. L. Williamson, "A Synthetic Workload Model for Internet Mosaic Traffic." Proc. SCSC'95, Ottawa, Canada, July 1995, pp. 852-857

M.F. Arlitt, C.L. Williamson, "Web Server Workload Characterization: The Search for Invariants." ACM Performance Evaluation Review 24 (1) May 1996, pp. 126-137. Available at ftp://ftp.cs.usask.ca/pub/discus/paper.96-3.ps.Z

D. Bauer, J. Charzinski, V. Held, "Virtuelle Universität durch ADSL-Technologie: Studenten im Geschwindigkeitsrausch durch breitbandigen Internet-Zugang." ITG-Fachtagung "Internet-Frischer Wind in der Telekommunikation", Stuttgart, Okt. 1999 (in German)

M.S. Borella, "Source Models of Network Game Traffic", Proc. Networld+Interop'99, Las Vegas, USA, May 1999. Available at http://www.xnet.com/~cathmike/MSB/Pubs/game-traffic.ps.Z

L.D. Catledge, J.E. Pitkow, "Characterizing browsing strategies in the World-Wide Web." Computer Networks and ISDN Systems 27(6) April 1995, pp. 1067-1073

J. Charzinski, "Internet Client Traffic Measurement and Characterisation Results", Proc. ISSLS, Stockholm, Sweden, June 2000.

J. Charzinski, J. Färber, M. Frank, J. Tölle, "The AMUSE Residential Multimedia Trials: Phase 1 Monitoring Results Summary." Proc. NOC'98, Manchester, UK, June 1998, pp. 86-93

H.-K. Choi, " A Behavior Model of a Web Traffic", 7th International Conference on Network Protocols 99' (ICNP 99'), Toronto, Canada, Nov. 1999

A.E. Conway, S.B. Moon, P. Skelly, "Synchronized Two-Way Voice Simulation Tool for Internet Phone Performance Analysis and Evaluation." Proc. 9th Conf. Computer Performance Evaluation, Modelling, Techniques and Tools, St. Malo, France, June 1997. Available at ftp://gaia.cs.umass.edu/pub/Moon96:Synch.ps.gz

C.R. Cunha, A. Bestavros, M.E. Crovella, "Characteristics of WWW Client-based Traces." Boston University, CS Dept, Boston, MA 02215 Technical Report TR-95-010, April 1995. Available at ftp://cs-ftp.bu.edu/techreports/95-010-www-client-traces.ps.Z

S. Deng, "Empirical Model of WWW Document Arrivals at Access Link." Proc. ICC´96, Dallas, TX, USA, June 1996, pp. 1797-1802

D.J. Ewing, R.S. Hall, M.F. Schwartz, "A Measurement Study of Internet File Transfer Traffic." Tech. Rep. CU-CS-571-92 Univ. of Colorado, Boulder, USA, Jan. 1992. Available at ftp://ftp.cs.colorado.edu/pub/cs/techreports/schwartz/FTP.Meas.ps.Z

J. Färber, M. Frank, J. Charzinski, "The WWW-Service in the AMUSE Field Trials: Usage Evaluation and Traffic Modelling." Proceedings of the Expert ATM Traffic Symposium, Mykonos, Greece, Sep. 1997. Available at http://www.ind.uni-stuttgart.de/IND/MA/Fa/papers/wwwusage.ps.gz

A. Feldmann, A.C. Gilbert, P. Huang, W. Willinger, "Dynamics of IP traffic: A study of the role of variability and the impact of control", Proc. ACM Sigcomm, Cambridge, MA, USA, August 1999. Available at http://www.cs.uni-sb.de/~anja/feldmann/papers/sigcomm99_trace_sim.ps.gz

J. Judge, H.W.P. Beadle, J. Chicharo, "Modeling World-Wide Web Request Traffic." Proceedings of the SPIE Conference 3020, San Jose, CA, USA, Feb. 1997, pp. 92-103

D.P. Heyman, T.V. Lakshman, A.L. Neidhardt, A.L., "A New Method for Analysing Feedback-Based Protocols with Applications to Engineering Web Traffic over the Internet." ACM Performance Evaluation Review 25(1) June 1997, pp. 24-38

P. Karlsson, Å. Arvidsson, "The Characteristics of WWW Traffic and the Relevance to ATM." Tech. Rep., Dept. Telecomm. and Math., Univ. of Karlskrona/Ronneby, Sweden; COST 257TD(97)21, May 1997. Available at ftp://www-info3.informatik.uni-wuerzburg.de/pub/cost/cost257/may97/257td97021.ps

B. Mah, "An Empirical Model of HTTP Network Traffic." Proc. IEEE Infocom'97, Kobe, Japan, Apr. 1997. Available at http://www.ca.sandia.gov/~bmah/Papers/Http-Infocom.ps

I. Marsh, "IP Telephony Tracing and Simulation." www.sics.se/~ianm/Telephony/iptel.html

H.F. Nielsen, J. Gettys, A. Baird-Smith, E. Prud´hommeaux, H.W. Lie, C. Lilley, " Network Performance Effects of HTTP/1.1, CSS1, and PNG." Proceedings of ACM SIGCOMM´97, Cannes, France, Sep. 1997, pp. 155-166. Available at http://www.w3.org/Protocols/HTTP/Performance/Pipeline.html

V.N. Padmanabhan, J.C. Mogul, "Improving HTTP Latency." Computer Networks and ISDN Systems 28, Dec. 1995, pp. 24-35

J. Rosenberg, H. Schulzrinne, "The IETF Internet Telephony Architecture and Protocols." IEEE Network 13, May/June 1999, pp. 18-23

H. Schulzrinne, J. Rosenberg, "A Comparison of SIP and H.323 for Internet Telephony." Proc. NOSSDAV, Cambridge, UK, July 1998. Available at http://www.cs.columbia.edu/~hgs/papers/Schu9807_Comparison.ps.gz

J. Sedayao, "World Wide Web Network Traffic Patterns." IEEE COMPCON´95 Technologies for the Information Superhighway, San Francisco, CA, USA, Mar. 1995, pp. 8-12

A. Woodruff, P.M. Aoki, E. Brewer, P. Gauthier, L.A. Rowe, " An investigation of documents from the World Wide Web." Computer Networks and ISDN Systems 28, May 1996, pp. 963-980. Available at http://epoch.cs.berkeley.edu:8000/~woodruff/inktomi/

Self similarity and models for fractal traffic

P. Abry, D. Veitch, P. Flandrin, "Long-Range Dependence: Revisiting Aggregation with Wavelets." Journal of Time Series Analysis 19(3) May 1998, pp. 253-266. Available at http://www.serc.rmit.edu.au/~darryl/A2.ps

R. Adler, R.E. Feldman, M.S. Taqqu, "A Practical Guide to Heavy Tails." Birkhäuser, Boston, Basel, Berlin, 1998

V. Bolotin, J. Coombs-Reyes, D. Heyman, Y. Levy, D. Liu, "IP Traffic Characterization for Planning and Control." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 425-436

M. Caglar, K.R. Krishnan, I. Saniee, "Estimation of Traffic Parameters in High-Speed Data Networks." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 867-876

M.E. Crovella, A. Bestavros, "Explaining World Wide Web Traffic Self-Similarity." Tech. Rep. TR-95-015, Boston University, CS Dept, Aug. 1995. Available at ftp://cs-ftp.bu.edu/techreports/95-015-explaining-web-self-similarity.ps.Z

M.E. Crovella, A. Bestavros, "Self Similarity in WWW traffic: Evidence and possible causes." Proc. ACM SIGMETRICS'96. Available at http://cs-www.bu.edu/faculty/best/res/papers/sigmetrics96.ps

A. Feldmann, A.C. Gilbert, W. Willinger, T.G. Kurtz, "The Changing Nature of Network Traffic: Scaling Phenomena." ACM SIGCOMM Computer Communication Review 28(2) Apr. 1998, pp. 5-29

A. Feldmann, W. Whitt, "Fitting mixtures of exponentials to long-tail distributions to analyze network performance models." Performance Evaluation 31, 1998, pp. 245-279

A. Feldmann, A.C. Gilbert, W.Willinger, "Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic." Proc. SIGCOMM'98, Vancouver, BC, Canada, Sep. 1998, pp. 42-55

L.J. Forys, A. Erramilli, J.L. Wang, "New Traffic Analysis and Engineering Methods for Emerging Technologies." Proc. GLOBECOM'95, New York, NY, USA, Nov. 1995, pp. 848-854

R.G. Garroppo, S. Giordano, M. Isopi, M. Pagano, "On the Implications of the OFF Periods Distribution in Two-State Traffic Models." IEEE Communications Letters 3(7) Jul. 1999, pp. 220-222

M. Greiner, M. Jobmann, C. Klüppelberg, "Telecommunication Traffic, Queueing Models, and Subexponential Distributions." Queueing Systems, 1998

M. Grossglauser, J.-C. Bolot, "On the Relevance of Long-Range Dependence in Network Traffic." Proceedings of ACM Sigcomm'96. Available at http://www.inria.fr/rodeo/personnel/mgross/WWW/Papers/sigcomm96.ps.gz

D.P. Heyman, D. Liu, "Assessing the Effects of Short-Range and Long-Range Dependence on Overflow Probabilities. "Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

J.K. Jerkins, A.L. Neidhardt, J.L. Wang, A. Erramilli, "Operations Measurements for Engineering Support of High-Speed Networks with Self-Similar Traffic." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 895-906

W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson, "On the Self-Similar Nature of Ethernet Traffic." Proc. ACM SIGCOMM´93, San Francisco, CA, USA, 1993, pp. 183-193

I. Norros, A. Simonian, D. Veitch, J. Virtamo, "A Benes formula for the fractional Brownian storage." COST 242 Technical Document (95)004, Version 2, May 1995. Available at http://www.serc.rmit.edu.au/~darryl/COST.ps

K. Park, W. Willinger, Eds. "Self-Similar Network Traffic and Performance Evaluation." Wiley, New York, 2000

M. Roughan, D. Veitch, M. Rumsewicz, "Computing Queue-Length Distributions for Power-Law Queues." Proceedings of Infocom'98. Available at http://www.serc.rmit.edu.au/~darryl/Infocom98_camera.ps

M. Roughan, D. Veitch, "A Study of the Daily Variation in the Self-Similarity of Real Data Traffic." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 67-76

M. Roughan, D. Veitch, P. Abry, "Real-Time Estimation of the Parameters of Long-Range Dependence." IEEE/ACM Trans. Networking 8(4)2000 pp. 467–478

Z. Sahinoglu, S. Tekinay, "On Multimedia Networks: Self-Similar Traffic and Network Performance." IEEE Comm. Mag. Jan. 1999, pp. 48-52

H.-P. Schwefel, L. Lipsky, "Performance Results for Analytic Models of Traffic in Telecommunication Systems, Based on Multiple On-Off Sources with Self-Similar Behavior." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 55-65

D. Starobinski, M. Sidi, "Modeling and Analysis of Heavy-Tailed Distributions via Classical Teletraffic Methods." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 45-54

E. Van den Berg, "Parameter Estimation for Long Range Dependent Data: Single Fine vs. Multiple Coarse Measurements." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

D. Veitch, "Novel Models of Broadband Traffic." Proc. GLOBECOM'93, Houston, TX, USA, Dec.1993

D. Veitch, P. Abry, "A Wavelet Based Joint Estimator of the Parameters of Long-Range Dependence." IEEE Trans. Information Theory 45(3) Apr. 1999.

W. Willinger, M.S. Taqqu, A. Erramilli, "A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks." in F.P. Kelly, S. Zachary, I. Ziedins, "Stochastic Networks: Theory and Applications" Royal Statist. Lecture Note Series, Clarendon Press, Oxford, UK, 1996

W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level." IEEE/ACM Transactions on Networking 5(1) Feb. 1997, pp. 71-86

W. Willinger, V. Paxson, "Where Mathematics meets the Internet." Notices of the American Mathematical Society 45(8) Sep. 1998, pp. 961-970. Available at ftp://ftp.ee.lbl.gov/papers/internet-math-AMS98.ps.gz

Measurement and modelling of Dial-up traffic

V.A. Bolotin, "New Subscriber Traffic Variability Patterns for Network Traffic Engineering." Proc. 15th International Teletraffic Congress (ITC 15), V. Ramaswami, P.E. Wirth, Eds., Washington DC, USA, June 1997, pp. 867-878

V.A. Bolotin, Y. Levy, D. Liu, "Characterizing Data Connection and Messages by Mixtures of Distributions on Logarithmic Scale." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 887-894

J. Charzinski, "Good News about Heavy Tails", IEEE Conf. High Perf. Switching and Routing (ATM2000), Heidelberg, Germany, June 2000.

J. Färber, S. Bodamer, J. Charzinski, "Statistical Evaluation and modelling of Internet dial-up traffic." Proc. SPIE Performance and Control of Network Systems III, Boston, MA, USA, Sep. 1999

A.A. Fredericks, "Impact of Holding Time Distributions on Parcel Blocking in Multi-Class Networks with Application to Internet Traffic on PSTN's." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 877-886

J.J. Gordon, K. Murty, A. Rayes, "Overview of Internet Traffic Issues on the PSTN." Proc. 15th International Teletraffic Congress (ITC 15), V. Ramaswami, P.E. Wirth, Eds., Washington DC, USA, June 1997, pp. 643-652

J. Kilpi, I. Norros: "Call Level Traffic Analysis of a Large ISP." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

S. Morgan, "The Internet and the Local Telephone Network: Conflicts and Opportunities." IEEE Comm. Mag. Jan. 1998, pp. 42-48

P. Orenstein, "A preliminary analysis of Work-at-Home Traffic Characteristics." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 907-918

TCP and TCP models

M. Allman, "On the Generation and Use of TCP Acknowledgements." ACM SIGCOMM Computer Communication Review 28(5) Oct. 1998, pp. 4-21

Å. Arvidsson, P. Karlsson, "On Traffic Models for TCP/IP." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 457-466

H. Balakrishnan, V.N. Padmanabhan, "How Network Asymmetry Affects TCP." IEEE Communications Magazine 39(4) 2001 pp. 60–67

C. Barakat, E. Altman, "Performance of Short TCP Transfers." Proc. Networking 2000, Paris, France, May 2000, pp. 567–579

T. Bonald, "Comparison of TCP Reno and TCP Vegas via Fluid Approximation", INRIA Rapport de Recherche No. 3563, Nov. 1998

C.P. Charalambos, V.S. Frost, J.B. Evans, "Performance of TCP Extensions on Noisy High BDP Networks." IEEE Commun. Letters 3(10)1999 pp. 294–296

S. Floyd, "A Report on Recent Developments in TCP Congestion Control." IEEE Communications Magazine 39(4) 2001 pp. 84–90

M. Mathis, J. Semke, J. Mahdavi, T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm." ACM Computer Communication Review 27(3), July 1997

J.C. Mogul, G. Minshall, "Rethinking the TCP Nagle Algorithm." ACM Computer Communication Review 31(1)2001 pp. 6–20

J. Padhye, V. Firoiu, D. Towsley, J. Kurose: ``Modeling TCP Throughput: A Simple Model and its Empirical Validation" ACM Computer Communication Review (Proc. ACM Sigcomm'98), Vol. 28 No. 4, Oct. 1998, http://www.acm.org/sigcomm/sigcomm98/tp/paper25.pdf

H.-P. Schwefel, "Behavior of TCP-like elastic traffic at a buffered bottleneck router." Proc. IEEE Infocom, Anchorage, AK, USA, April 2001. Available at http://wwwjessen.informatik.tu-muenchen.de/~schwefel/TCP_paper.ps.gz

I. Yeom, A.L.N. Reddy, "Modeling TCP Behavior in a Differentiated Services Network." IEEE/ACM Trans. Networking 9(1)2001 pp. 31–46

Measured Performance

P. Barford, M. Crovella, "Measuring Web Performance in the Wide Area." ACM Perf. Eval. Review 27(2)1999 pp. 37–48

P. Barford, M. Crovella, "A Performance Evaluation of Hyper Text Transfer Protocols." Proc. ACM Sigmetrics, Atlanta, GA, USA, May 1999

P. Barford, M. Crovella, "Critical Path Analysis of TCP Transactions." Proc. ACM Sigcomm Stockholm, Sweden, Sep. 2000

E. Cohen, H. Kaplan, "Prefetching the Means for Document Transfer: A New Approach for Reducing Web Latency." Proc. IEEE Infocom Tel Aviv, Israel, March 2000

M.A. Habib, M. Abrams, "Analysis of Sources of Latency in Downloading Web Pages." Proc. WebNet, San Antonio, USA, Nov. 2000

C. Huitema, S. Weerahandi, "Internet Measurements: the Rising Tide and the DNS Snag." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

B. Krishnamurthy, C.E. Wills, "Analyzing Factors that Influence End-to-End Web Performance." Computer Networks 33 (1–6) 2000 pp. 17–32

Dimensioning

S. Bodamer, J. Charzinski: "Evaluation of Effective Bandwidth Schemes for Self-Similar Traffic." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

J. Charzinski, "Fun Factor Dimensioning for Elastic Traffic", COST 257 Tech. Doc. (00)28, Oslo, May 2000. Avail. at http://www-info3.uni-wuerzburg.de/cost/

A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford: "Deriving Traffic Demands for Operational IP Networks: Methodology and Experience." Proc. ACM Sigcomm, Stockholm, Sweden, August 2000

D.P. Heyman, T.V. Lakshman, A.L. Neidhardt, "A New Method for Analysing Feedback-Based Protocols with Applications to Engineering Web Traffic over the Internet" Proc. ACM SIG-METRICS'97, Seattle, WA, USA; ACM Perf. Eval. Review Vol. 25, No. 1, pp. 24–38.

J.G. Klincewicz, J.A. Schmitt, R.T. Wong, "The Design of IP Enterprise Networks with QoS", Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

K. Lindberger, "Balancing Quality of Service, Pricing and Utilization in Multiservice Networks with Stream and Elastic Traffic" Proc. ITC 16, Edinburgh, UK, June 1999, pp. 1127–1136.

M. Nabe, M. Murata, H. Miyahara, "Analysis and modeling of World Wide Web traffic for capacity dimensioning of Internet access lines." Performance Evaluation 34, 1999, pp. 249–271

I. Norros, "On the Use of Fractional Brownian Motion in the Theory of Connectionless Net-works", IEEE JSAC Vol. 13, No. 6, 1995, pp. 953–962.

Server, Caching and Proxies

M.F. Arlitt, " A Performance Study of Internet Web Servers." Master Thesis, Dep. Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0W0, 1996. Available at ftp://ftp.cs.usask.ca/pub/discus/thesis_arlitt_co.ps.Z

N. Bhatti, A. Bouch, A. Kuchinsky, "Integrating User-Perceived Quality into Web Server Design." Comp. Netw. 33 (1–6) 2000 pp. 1–16

J.C. Bolot, P. Hoschka, "Performance Engineering of the World Wide Web: Application to dimensioning and cache design." Computer Networks and ISDN Systems 28, May 1996, pp. 1397-1405

H.-W. Braun, K.C. Claffy, "Web traffic characterization: an assessment of the impact of chaching documents from NCSA´s web server." Computer Networks and ISDN Systems 28, Dec. 1995, pp. 37-51

J. Dilley, M. Arlitt, "Improving Proxy Cache Performance: Analysis of Three Replacement Policies." IEEE Internet Computing Nov. 1999 pp. 44–50

Y. Fujita, M. Murata, H, Miyahara, "Performance Modeling and Evaluation of Web Systems with Proxy Caching." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 1179-1188

A. Loutonen, K. Altis, "World-Wide Web Proxies." Computer Networks and ISDN Systems 27 Nov. 1994, pp. 147-154

D. Neal, "The Harvest Object Cache in New Zealand." Computer Networks and ISDN Systems 28, May 1996, pp. 1415-1430

P.K. Reeser, R.D. van der Mei, R. Hariharan, "An Analytic Model of a Web Server." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 1199-1208

QoS

M. Asawa, "Measuring and Analyzing Service Levels: A Scalable Passive Approach." Proc. Sixth International Workshop on Quality of Service (IWQoS'98), May 1998, pp. 3-12

T. Bonald, J.W. Roberts, "Performance of Bandwidth Sharing Mechanisms for Service Differentiation in the Internet." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

J. Charzinski, "Problems of Elastic Traffic Admission Control in an HTTP Scenario." Proc. IWQoS, Karlsruhe, Germany, June 2001.

R.A. Guérin, V. Pla, "Aggregation and Conformance in Differentiated Service Networks – A Case Study." Proc. ITC Specialist Seminar on IP Traffic, Monterey, CA, USA, Sep. 2000

A. Kumar, M. Hegde, S.V.R. Anand, B.N. Bindu, D. Thirumurthy, A.A. Kherani, "Nonintrusive TCP Connection Admission Control for Bandwidth Management of an Internet Access Link." IEEE Communications Magazine 38(5) 2000 pp. 160–167

L. Massoulié, J. Roberts, "Arguments in favour of admission control for TCP flows." Proc. ITC 16, Edinburgh, UK, June 1999, pp. 33-44

M. May, J.C. Bolot, A. Jean-Marie, C. Diot, "Simple Performance Models of Differentiated Services Schemes for the Internet." Proc. IEEE Infocom New York NY, USA, March 1999

R. Mortier, I. Pratt, C Clark, S. Crosby, "Implicit Admission Control." IEEE JSAC 18(12) 2000 pp. 2629–2639

K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Service Field (DS Field) in the IPv4 and IPv6 Headers.", RFC 2474, Dec. 1998

J.W. Roberts, S. Oueslati-Boulahia, "Quality of Service by Flow Aware Networking." Phil. Trans. Royal Soc. Series A Vol. 358 No. 1773, August 2000

S. Shenker, "Fundamental Design Issues for the Future Internet." IEEE JSAC 13(7), Sep 1995, pp. 1176-1188

L. Zhang, S. Deering, D. Estrin, S. Shenker, D. Zappala, "RSVP: A New Resource ReSerVation Protocol." IEEE Network, Sep. 1993, pp. 8-18

Simulation

M.E. Crovella, L. Lipsky, "Long-Lasting Transient Conditions in Simulations with Heavy-Tailed Workloads." Proc. Winter Simulation Conf. 1997

S. Floyd, V. Paxson, "Why We Don't Know How To Simulate The Internet." Proc. 1997 Winter Simulation Conference, Atlanta, GA, USA, Dec. 1997. Available at http://www.aciri.org/floyd/papers/wsc97.ps

H.-P. Schwefel, L. Lipsky, M. Jobmann, "On the Necessity of Transient Performance Analysis in Telecommunication Networks." Feb. 2001. Available at http://wwwjessen.informatik.tu-muenchen.de/~schwefel/itc17.ps.gz