# Resource Engineering for Internet Applications


Ajit K. Jena
Computer Center
Indian Institute of Technology
Bombay-400076, India
Email: ajit@cc.iitb.ac.in

Adrian Popescu
Dept. of Telecommunications
and Signal Processing
Blekinge Institute of Technology
371 79 Karlskrona, Sweden
Email: apo@bth.se

Arne A. Nilsson
Dept. of Telecommunications
and Signal Processing
Blekinge Institute of Technology
371 79 Karlskrona, Sweden
Email: ano@bth.se



***Abstract*— The focus of the paper is on resource engineering for supporting Service Level Agreements (SLAs) in IP networks. We review some of the recent developments in Internet service architecture and protocol developments as well as some of the most important challenges still to be considered. An important performance metrics that must be considered in developing the new Internet is regarding the delay and the delay variation/jitter. Connected with this, a case-study is reported on end-to-end delay performane in a simple model for best-effort Internet where the bandwidth is not a commodity. The case-study is based on a client-server simulation model. The client and the server nodes are running applications like HTTP and SMTP, and they are placed in two sites separated by an Internet cloud. Performance issues of Short Range Dependence (SRD) and Long Range Dependence (LRD) traffic under different resource control regimes are compared. The results highlight the importance of the queueing delay at the WAN ingress point, which is due to the significant bandwidth differences that may exist, in this case, between the LAN and WAN link layers. The results also highlight the role TCP window size and Frame Relay Permanent Virtual Channel (FR PVC) control mechanisms play in the provision of delay performance for Internet services.**


## I. Introduction

The recent advent of IP-based telephony and multimedia services has offered providers of teleservices a great opportunity to push for the development of new services that integrate different media and also exploits the power of intelligent terminals. Communications can now encompass not only voice, but also audio, video, shared applications, and even virtual reality. At the same time, Internet, with the access to information that is offered to everyone, is acquiring an everincreasing importance for business community as well as for the population at large.

Recent proliferation of heterogeneous networking technologies has also created new opportunities and challenges for network deployers. On one hand, there is the choice of extension of the new teleservices to cellular phones by developing more powerful user interfaces and also gateways to the public network. On the other hand, there is the challenge of the provision of Quality of Service (QoS) guarantees on an end-to-end basis, i.e., at the level of applications running in user space of general purpose operating systems. Furthermore, while the access to Internet has become a commodity, there is also a strong need for differentiating the service offerings of different Internet Service Providers (ISPs), under the frame of QoS as well as of value-added Internet services.

The focus of the paper is on resource engineering for supporting Service Level Agreements (SLAs) in IP networks. Two levels are considered, SLA at the link (packet) level and SLA at the application (end-user) level. A case-study is considered for these purposes, which is based on an object-oriented simulation model for client-server interactions able to capture real message exchanges as well as traffic payloads as seen in typical Internet access scenarios. The OPNET simulation environment is used and the parameters for simulation are derived from real Internet traffic flows. Internet applications like HTTP and SMTP are considered in the study, and statistical models have been developed that are based on real traffic measurements. The core of the network is formed by a Frame Relay (FR) WAN, with a topology that allows for diverse traffic mixtures.

The focus of the experiments is on the delay performance together with controls and traffic engineering to cater for a specific SLA. The goal is to do dimensioning of network resources with reference to SLA fulfilment as well as to find out the implied performance bounds when diverse operating conditions are varied in a certain range. The impact of link and transport layer controls is studied and the SLA is related to TCP and FR control parameters. The influence of Long Range Dependence (LRD) properties on the delay performance at packet and application level is reported as well. Starting with Short Range Dependence (SRD) traffic conditions, a bandwidth and buffer allocation scheme is worked out such as a specified SLA is obtained at both the packet and application level. Subsequently, under LRD traffic conditions, the utilization profiles and the resource allocation schemes are adjusted so as to achieve similar performance levels.

The results confirm the importance of the queueing delay at the WAN ingress points, which is due to the significant bandwidth differences between LAN and WAN link layers. The results also bring out the significant role TCP window size and Frame Relay PVC control mechanisms play in the provision of delay performance for Internet applications. Actually we are not aware of similar results reported in the literature.

The paper is organized as follows. The first sections are devoted to reviewing the fundamental concepts of traffic engineering (section 2) and the main challenges still to be considered (section 3). Section 4 is devoted to Service Level Agreement (SLA) definitions and descriptions. Performance metrics for SLA are considered. Section 5 is focused on the description of the case-study considered for experiments. In particular, settings for various aspects in the model (mix of applications, TCP parameters, PVC settings, and diverse router parameters) are presented in a detailed way. The next part (section 6) reports the performance obtained at the packet level, for both SRD and LRD traffic. Finally, section 7 is devoted to the performance obtained at the application level.

## II. Internet Traffic Engineering

Internet Traffic Engineering (TE) is part of the Internet network engineering that addresses the issue of performance optimization. The main functions of traffic engineering are traffic and capacity management (to maximize the network performance at a minimum cost) as well as network planning (to adapt to forecasted traffic growth) [3], [4]. Main objectives are reliable network operations and performance enhancing. A number of tools are used for traffic engineering, e.g., measurement, characterization, modeling and control of Internet traffic as well as a number of protocols, e.g., related to QoS resource management (connection admission control CAC, bandwidth reservation, scheduling and buffer management, etc.), or related to routing (table) management (policing, bandwidth broker BB, resource reservation protocol RSVP, etc.) [3]–[5]. Because of the complexity of Internet traffic models and protocols, traffic engineering has been proved to be a complex procedure itself [4], [16], [32].

Traffic engineering is a critical part of the network design and operation, and an important goal is to balance the traffic loads in the network by redirecting the traffic on alternative paths. Generally, the traffic engineering problem can be treated as an optimization problem, where the (pro-active and/or reactive) control of Internet traffic is exercised on multiple levels of temporal resolution. For instance, the capacity planning component considers very large temporal scales, ranging from days to possibly months and even years. On the other hand, certain traffic management functions, like packet level processing (e.g., queue management and scheduling), act on small time scales, which are ranging from microseconds to milliseconds.

Another aspect that complicates things is the fact that the Internet traffic engineering is acting on different domains, intra-domain (within a specific Autonomous System AS) as well as inter-domain (covering more ASs). Especially in the case of best-effort IP networks, with routers with limited functional capabilities, effective Internet traffic engineering has been shown to be difficult to implement. However, new developments like Multiprotocol Label Switching (MPLS) as well as faster and more intelligent routers have (partially) opened up this bottleneck, starting so the migration toward new network architectures that are more intelligent, more robust and offering better performance [1], [4], [13], [16], [19], [27].

## III. Internet Challenges

In spite of the above-mentioned successes, the Internet faces new challenges, which in turn may influence the solutions adopted for traffic engineering. Some of the most important challenges are:

### A. Provision of End-to-End Guarantees

The new models put forth for IP QoS means that the focus of data networking has been shifted from secure data delivery to data delivery with guaranteed delay performance (in the form of end-to-end delay and jitter). The new Internet is expected to provide services where several parameters (delay, jitter and packet loss) are minimized as much as possible. New control mechanisms are under development, to help network operators providing reliable service levels and also differentiating vs competitors, where possible differentiators are metrics like connection setup and delay (within the frame of Service Level Agreement SLA).

Actually, the only one way to provide end-to-end delay guarantees is to create an end-to-end data flow and to reserve resources in the network. That means that connection-oriented (CO) subnetworks must be used. Typical examples of CO subnetworks are Asynchronous Transfer Mode (ATM), Multiprotocol Label Switching (MPLS), Frame Relay (FR) as well as the Integrated Services (IntServ) model. The Differentiating Services (DiffServ) model is also typical for this case, except that performance guarantees can only be provided when the network is lightly loaded [9], [23].

The ideal subnetwork for IP is however connectionless (CL). Two technologies have therefore been proposed for reducing the delay in this case. These are the header compression and using of fragmentation/reassembly [2], [12]. The area of applicability is however limited, e.g., header compression is used only for low-speed serial links.

Furthermore, due to the complexity of Internet traffic as well as of Internet protocols, spanning from Medium Access Control (MAC) protocols at the link layer up to specific control mechanisms at the application layer, several other aspects have come to play an important role in the provision of end-to-end delay guarantees. These are the traffic self-similarity, the multilevel network control and the so-called routing flaps.

### B. Scalable Solutions for QoS

A lot of efforts have gone into the development of new systems and technologies to provide networking with QoS. Several network models for QoS have been developed so far, but they have been also shown to have limitations of different kinds. For instance, the Integrated Services (IntServ) model [10] (which is used together with the Resource ReSerVation Protocol (RSVP) [11]), has the problematic issue of RSVP scalability, in spite of the advantage of per-flow control. Further, the alternative solution developed to overcome the scalability limitation, the Differentiated Services (DiffServ) model [9] (which is used together with MPLS and TE, and with aggregate-flow control), faces other limitations, e.g., in the form of imprecise QoS guarantees [23]. A better model for

QoS seems to be the so-called Hybrid model, where peripheral subnetworks are IntServ- and RSVP-aware, whereas the core networks are DiffServ-aware [8]. This model is however still under study [2]. Fundamental questions like ”How much complexity to incorporate in access networks and how much in core networks?”, ”How to do suitable resource reservation in the core networks?”, and also regarding policing and signaling, need still to be answered.

### *C. Traffic Self-Similarity*

Traffic measurement studies from a wide range of working packet networks (including Ethernet LANs, WANs, CCSN/SS7, ISDN and VBR video over ATM) have convincingly shown the presence of self-similar and Long-Range Dependence (LRD) properties in both local area and wide area traffic traces, which means that similar statistical patterns may occur over different time scales that can vary by many orders of magnitude (i.e., ranging from milliseconds to minutes and even hours) [7], [14], [21], [26], [31]. Today, there is mounting evidence that LRD is of fundamental importance for a number of traffic traffic engineering problems, such as traffic measurements, queueing behavior and buffer sizing, admission control and congestion control [32]. Unlike traditional packet traffic models, which give rise to exponential tail behavior in queue size distributions and typically result in optimistic performance predictions and inadequate resource allocations, LRD traffic models predict hyperbolic or Weibull (stretched exponential) queue size distributions and could therefore result in longer waiting times at the network processing elements (e.g., routers), affecting so the control and the management of the Internet [15], [25].

The global Internet has also seen tremendous growth in terms of nodes and user base as well as of types of applications. One of the most important consequences of this growth is related to an increased complexity of the traffic experienced in these networks. Further, transporting packets across the network between the application end points is subject to delays and errors. They may even be discarded by a congested router/switch. Because of the presence of LRD phenomenon across many types of networks, metrics of network performance such as throughput, packet loss, latency and buffer occupancy levels are affected. Accordingly, some of the most important consequences are [32]: packet delays and consequently application level delays have a heavy-tailed distribution; transport layer protocols like TCP estimate the round trip timer values from the peer acknowledgements and hence are influenced by it; congestion situations are unavoidable and they appear as short-lived impulses; with increase in load, congestions appear more frequently while maintaining the impulsive behavior; only increasing buffer sizes does not result in significant improvements in packet loss behavior.

### *D. Multilevel Network Control*

Another aspect that must be considered in IP QoS is the growing importance of proper integration of control mechanisms acting at diverse time scales and layers. For instance, it is well-known that the Internet may occasionally corrupt, drop, duplicate or reorder data packets. Depending on the layer where they are implemented, error control mechanisms can operate either on a point-to-point basis (e.g., reliable link layers specified for IEEE 802.2, 802.11, X.25, HDLC) or on an end-to-end basis (e.g., TCP error control, or application-based error control) or combinations. Error recovery at the link layer is especially useful in the case of subnetworks with (relatively) high error rates, e.g., wireless and other noisy links. The ultimate responsibility for error recovery (in the case of CO applications), is however at the end points. Thus, point-to-point and end-to-end control mechanisms may often co-exist, and this may lead to situations when these mechanisms react destructively to each other resulting so in performance deterioration. This is a direct consequence of improper protocol implementations, to cater with such unfavorable situations.

The interplay between point-to-point and end-to-end control mechanisms, and between protocols at different layers (e.g., HTTP, TCP) is generally a complex process, and difficult to study analytically. Other requirements, such as efficiency of resource allocations, further complicates the picture. Protocol design that considers interplay with other protocols is a rather new area, which is in its infancy of exploration.

### *E. Routing Flaps*

Some of the most serious impairments of the Internet, which could introduce delay variations too big to be compensated in the buffers at receivers, are traffic congestion (especially at the ingress routers) and routing flaps (at the core routers). The routing flaps may occur when changes in network routing occur, and they are like shock waves that propagate through the Internet’s backbone. For instance, when a major core router, or a link, goes down, the other routers have to reconfigure the routing for the traffic addressed to the damaged router. Further, the information about the new (routing) configuration is also propagated to the other routers, across multiple routing domains, to the farthest corners of the Internet. Should other changes in the Internet occur before the first change has completely propagated to the corners, a new set of ripples may appear that collide with the previous one, creating so a situation when routers are ”flapping” instead of routing, and finally generating a continual background roar of changes in the routing and the routing tables [6]. By this, different transient forwarding blackholes and/or loops may be created, affecting so the delay performance of the Internet. Routing protocols like Open Shortest Path First (OSPF) and Border Gateway Protocol (BGP) are especially sensitive to this kind of impairments. Intensive research activity has been started to minimize the negative effects of routing flaps. Some of the best ways to do that seem to be by using increased computing power in routers as well as by using of specific dampening strategies to improve the overall stability of the Internet routing tables and to off-load the CPUs of core routers [6], [28].

## IV. Service Level Agreement

An SLA is a formal definition of the relationship that exists between a supplier of services and customers. SLA addresses issues like the nature of service provided, reliability, responsiveness, the process for reporting problems, time frame for response to reported problems and problem resolution, methodologies for auditing service levels, penalty and escape clauses, etc [30]. In the context of Internet, service issues have become an important tool for Internet Service Providers (ISPs) to retain their client portfolio and to attract new customers. An ISP may provide specific SLAs to customers who may use the ISP network in different ways, which can be generally described as a combination of three basic modes, namely to access the public Internet, to interconnect two or more sites, and to access proprietary, industry specific networks such as in the case of enterprise networks. However, the terms of SLA that govern each of the access profiles are likely to be different. For example, when the objective is to connect two sites of a given customer, the assurances about the performance level will focus on the path between the pair of access routers and is likely to be more clearly defined and controlled. On the other hand, when the goal is to access the Internet in general, the ISP may not be in a position to control the performance obtainable. However, in such cases, the ISP may be in a position to certify that the performance obtained will be more or less decided by the external Internet cloud and that the ISP's access network will not be the bottleneck.

In the context of computer networks, service specification for SLA can be done at different protocol layers. At the application layer, the focus is on application sessions. Primitives like throughput, object access delay, and transaction update turnaround time are relevant here. Further, the overall performance aspects obtained at the application layer is a combination of the stochastic nature of payload contents sizes, the characteristics of application layer protocol, and of the transport and lower layer protocols. Application layer protocol works as a feedback loop on the end-to-end path and thus influences the performance significantly.

Another key issue to be considered is that the performance at the application level is also dependent on the number of concurrent application sessions. A SLA specification geared toward the ***busy hour traffic*** may be more meaningful here.

At the transport layer, primitives like connection throughput and goodput are considered for SLA. On the contrary, at the link layer, the focus is on packets. Given the expected size of user population at the client site, major parameters considered here are the bandwidth allocation and the port buffer sizing. Also, there may be rate and burst control issues as well (e.g., as used in ATM or FR). Depending on particular environments, a combination of these layer specific SLAs may be chosen.

Generally, there are two aspects related to SLA of networked systems: availability and responsiveness [30]. These are applicable at both the application level and the network layer. Host or router uptime, link outage frequency, and error rates, they all fall in the category of system availability. Most of these parameters can be monitored and detected using diverse network management tools (e.g., SNMP). Further, some of the major points that can be grouped under the category of responsiveness are: one way end-to-end delay at the link layer; application level turnaround time; TCP connection setup latency; TCP retransmissions; packet loss rates; and available bandwidth. Some of these metrics can be collected from regular SNMP or RMON type statistics databases whereas others (e.g., TCP and application layer metrics) can only be audited via sophisticated monitors. Some SLAs may also take the form of guaranteed transmission profiles (such as Continuous Bit Rate CBR in ATM and Committed Information Rate CIR in FR PVCs and SVCs), to allow users to inject a specific amount of data into the network with an assured bandwidth and loss characteristics.

## V. Case-Study: SLA on a Best-Effort Network

The goal is to study a client-server model driven with Internet applications like HTTP and SMTP. This involves understanding and simulation of application sessions and the constituent transactions. Toward this end, an object-oriented view has been developed into a client-server simulation framework [17]. The model closely mimics the real-life events that occur in an Internet consisting of clients and servers. The client and the server nodes running the above-mentioned applications are placed in two sites (S1, respectively S2), which are separated by an Internet cloud (fig 1). The OPNET simulation tool [24] is used as a simulation environment.

The ingress and egress points to the Internet is via routers, which feed into a FR WAN link via Frame Relay Access Devices (FRADs). The clients and the servers are organized into two separate sites. The client nodes are 10BaseT Ethernet hosts and the server nodes are 100BaseT hosts.

The application traffic is modeled with a mixture of Uniform (or Lognormal) and Pareto distributions [18] and, accordingly, a fractional Brownian model ($fBm$) is used to model the network traffic for a large number of clients (according to the so-called "Joseph effect") [15].

Most of Internet client-server applications can be modeled as ON-OFF sources. The client makes a REQUEST for a service at a specific server (client side ON state) and the server responds with a RESPONSE (server side ON state), thus giving rise to a lock-step behavior. The ON periods on both sides are interspersed with OFF periods. The stochastic aspects of ON duration derive directly from those of the protocol message elements and user message contents [14]. The high variability in the content sizes results in clusters of packets that enter the network. The OFF periods are typically due to user inactivities and the variability in the OFF process may be due to different types of users accessing network services concurrently. The level of activity of human users generally varies from user to user.

In order to understand the end-to-end transactions, it is important to quantify the effects of network control mechanisms on the application-level properties. Transporting packets across the network between the application end points is subject to


Fig. 1. Network architecture used for study

delays, errors, and losses that occur at congested router/switch. Even though the presence of LRD has been known for quite some time now, the impact on end-to-end performance at the packet or application layer is not fully known [32]. The aim of the simulation experiment is focused on the end-to-end delay performance while maintaining packet losses within specified limits. The motivation is that as far as end-user performance is concerned, packet losses and delay are related. Packet losses result in retransmissions at the TCP layer and indirectly affect the object access latency at the application layer. By minimizing losses and targeting the delay performance the overall performance may be optimized and at the same time may result in simplification of the simulation experiment setup.

### A. Setting Up the Environment

Global Internet is an example of best-effort collaborative service spanning across many geographical, cultural, and ethnic limits. Individual organizations own and administer parts of this network as per their own policies. When the users access services on the Internet, the end-to-end path sometimes literally spans across continents. Along the end-to-end path, the packets may cross many different organizations, network technologies, and routing and switching policies. Because of these diverse conditions, in general, it is very difficult to make performance evaluations to a specific degree of accuracy. Simulation oriented investigations may however be possible under some specific operating conditions. Even so, it is extremely important to identify important environmental conditions and choose appropriate parameter values for them. Some of the most important parameters, and suitable choice for values of the same are as follows.

*1) Network architecture:* The case-study focuses on a hypothetical network architecture that is appropriate for a corporate Intranet or an interconnection of two sites of an organization separated by hundreds of kilometers of geographical distance. The network consists of three components: the local networks labeled "Site 1" and "Site 2", and the "IP Cloud" interconnecting them. The local networks are owned and administered by the corporate organization whereas the ISP handles the interconnectivity. The ISP's infrastructure to support this connectivity is considered to be a single T1 link.

The relation between the corporate organization and the ISP takes the form of a SLA. A threshold of 1 sec was selected as SLA for the WWW delay at application level. The goal of the study is to do dimensioning of network resources with reference to SLA fulfilment as well as to find out the implied performance bounds when diverse operating conditions are varied in a certain range.

*2) Mix of applications:* WWW constitutes today about 60-80% of the Internet traffic load and the second application in terms of traffic volumes seems to be E-mail [17]. Given this background, a mix of WWW and E-mail traffic has been considered as an application environment for the experiment. The proportion of the traffic belonging to the respective categories is controlled by controlling the number of client hosts on either sites and/or by adjusting the session arrival rates. The simulation investigates the effects during a ***busy hour***. During the ***busy hour***, session arrivals for both classes of applications are modeled as Exponential processes with a mean inter-arrival rate of 15 seconds.

The second important aspect considered here is the choice of the parameters for content distributions at the application layer. This has been adjusted so as to create two separate environments, one with LRD traffic characteristics (Hurst estimate of roughly 0.9) and the other with SRD traffic (Hurst estimate of roughly 0.55). The WWW traffic is considered as the main traffic and E-mail traffic serves only as a background traffic. The LRD properties of the mixture traffic are determined by the traffic with highest LRD component [15]. Accordingly, the SMTP traffic parameters were selected so as to result in a traffic mix where the behavior of the WWW traffic is dominant.

*3) TCP parameters:* The TCP layer is an important element in the end-to-end path of the application service loop. Each connection uses TCP window mechanism to dynamically determine the available capacity in the network and accordingly to decide how much data it can safely have in transit. The TCP window size is an important parameter in flow and congestion control of the Internet traffic.

However, when implementing an application service, there is no actually guiding principle as to how to select a particular window size. The implementation of the same application service under different operating systems tends to use different window sizes (when initiating TCP connections), e.g., Sun Solaris uses a window size of 8760 whereas RedHat Linux uses a value of 32120 [22]. It is unclear how this kind of shaping mechanism interacts with the application layer properties and what is the effect on performance. Accordingly, the simulation experiments are performed for window sizes of 8760, 16384, and 32120 bytes.

*4) PVC settings:* The physical connectivity between the two sites of the network under investigation is provided by a standard (1.5 Mbit/sec) T1 link with a FR PVC creating a logical channel over it. The characteristics of the logical channel are described as a set of three parameters, i.e., the Committed Information Rate ($CIR$), the Committed Burst Size ($B_c$), and the Excess Burst Size ($B_e$) [29].

To the best of our knowledge, there are no methodologies actually available as to how to select parameters to ensure optimal resource allocation. In order to find a meaningful value for $CIR$, a calibratory approach has been adopted. To begin with, a *Zero CIR* allocation strategy is chosen for the PVC. Under such conditions, the FR rate control does not take any effect. All packets that enter the PVC are marked ***discard eligible***. Even though all packets are admitted into the network, their delivery is subject to resource conditions on the physical path. A set of about 100 client hosts are used to populate the two sites in network to allow for a sufficient number of TCP connections on the link, such as the network traffic approaches the *fBm* model. Further, to start with, the application parameters for the sources are chosen so as to create more or less Markovian conditions. The link resource conditions are basically the port buffer size at the routing or switching stages. The buffers were optimized to keep packet loss within prescribed limits. With this setup, the simulation experiment was performed and the packet trace collected. From this, the byte count was collected over non-overlapping intervals of 10 seconds and the average bandwidth computed was 500 Kbytes, which corresponds to 50 Kbytes per second. Using this is as a thumb rule, a *CIR* of 64 Kbytes/sec seems to be adequate, i.e., 512 kbit/sec. There are a few intervals where the *CIR* limit is indeed exceeded but they can be easily accommodated by the excess physical bandwidth available on the link.

Bandwidth allocation is essentially studied as a scaling problem. Having decided on a *CIR* allocation for the traffic with SRD properties, the number of hosts is then reduced gradually for the LRD case while maintaining similar end-user QoS levels, thus determining the bandwidth allocation applicable for the specific traffic.

The choice for the burst control parameters $B_c$ and $B_e$ has also no formal principle. The ratio $\frac{CIR}{B_c}$ is denoted as $T_c$, the measurement interval. Survey of relevant literature has revealed that prescribed values for $T_c$ vary in the range 0.1 to 1.0 but the reasoning behind the specific limits are unknown [29]. For the parameter $B_e$, no suitable ranges were found. Again, choosing an exploratory strategy, the values for the parameters $T_c$ were varied over a range and $B_e$ was varied in relation to $T_c$. Simulation runs were carried out for several such profiles in order to estimate the performance impacts of such variations.

*5) Router/Switch parameters:* The choice of parameters for router or switching elements along the end-to-end path is an important factor in performance estimations. The first parameter of interest is the time taken per packet by the routers and switches. This is a value independent of the packet size. The routers use their FRAD interface to put the packet onto the WAN and similarly the switches also have to just do a table lookup to determine the output port. As each packet arrives at the input port of a switch, it raises an interrupt. The switching element determines the output port for the specific packet and schedules the packet to join the queue at that port exactly after one switching time delay. However, the packets may have to wait at the output port because of multiplexing. The time delay per packet at the switching or routing stages has been set to 20 *µsec*. This is a typical value for modern, fast packet processors. The other important parameter is the size of the port buffer. Here too, a heuristic approach was applied. For different combinations of TCP window sizes and PVC burst control parameters, the loss performance was monitored via simulation. The port buffers were dimensioned not in bytes or packets but based on the speed of the link to which the port is attached. For example, for a T1 link, a port buffer dimension of 100 milliseconds indicates a port buffer size of 15360 bits. Through this experimental procedure the link buffers are gradually increased so as to keep the losses limited to well below 0.5%. For the experimental environment used in the case-study, a buffer of 250 milliseconds seems to be adequate for our purposes.

## VI. Packet Level Performance

The packets generated by the application sources are aggregated inside the edge router. FR control mechanisms feed these packets into the network subject to the packets meeting the PVC control criteria. For the given case study, there is spare capacity available inside the network because the PVC *CIR* allocation is about one third of the link capacity. Under such conditions, once the packets are injected into the network (as per the configured PVC settings), the performance is not much affected by link conditions. The packets also encounter switching delay that is influenced by the switch hardware. The bottleneck link does not cater to any cross traffic flows and thus behaves as a single delay line consisting of a cascade of queueing stages. At a typical switch, the queueing delay experienced by successive packets can be expressed using the Lindley's equation [20].

The simulation results reported in this section are only about the queueing delays. The transmission and propagation components of the delay are disregarded in order to focus purely on the traffic generation process.

The queueing process at the ingress router is predominantly due to bandwidth mismatch between Ethernet LAN and the FR WAN, number of concurrent application sessions, and burst characteristics of the generated traffic. Further, we consider the burst characteristics as being captured by the Hurst estimates of the byte count process at the link layer. The impact of Hurst estimates on the queueing process is shown in fig. 2. The empirical CCDF of the delay across the PVC for both SRD and LRD traffic scenario are plotted using a logarithmic scale on the y-axis. The almost linear behavior in the SRD case indicates an exponential behavior whereas the LRD case appears to show a behavior typical of Weibull distributed random process, indicating so a *fBm* traffic model in the link. The delay characteristics shown in the plot are for the case of high utilization conditions with *Zero CIR* PVC settings and when the TCP connections use a window size of 8760.

Further, for the above case, the CCDF of the end-to-end delay is compared with the delay experienced by the packets at the PVC entry point (fig 3). As the figure indicates, the queueing delay at the WAN ingress point dominates the end-to-end performance, with the conclusion that the effect of queueing and switching delays at the switches inside the WAN is marginal.

The message streams that originate at the application layer are partitioned into TCP segments. The segments are channeled along the end-to-end path subject to flow and congestion control mechanisms of TCP. In other words, the TCP layer has a kind of non-linear (modulating) impact on the data streams

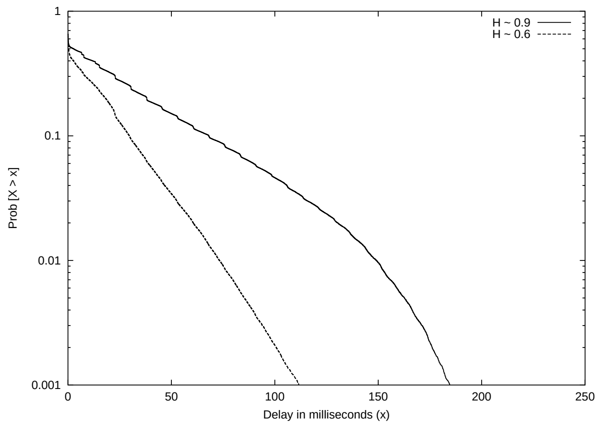


Fig. 2. Impact of LRD on the end-to-end packet delay

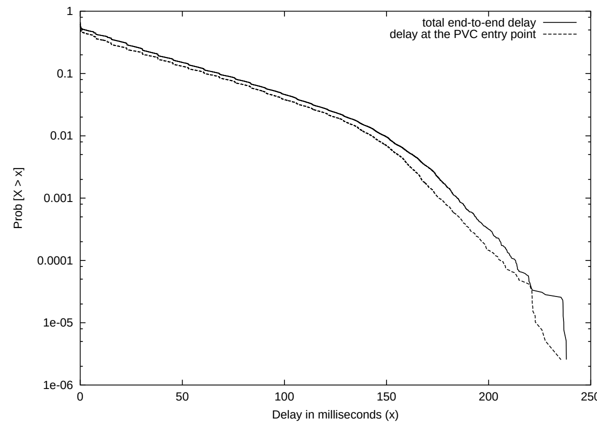


Fig. 3. End-to-end delay vs queueing delay at the PVC entry point

passing through it. Each source tries to determine the available capacity in the network so that it can maximize the number of in-transit segments. In fact, TCP repeatedly increases the load it imposes on the network in an effort to locate the point where congestion occurs, and then it backs off from there. Thus, it creates losses to find the available bandwidth for the connection. In cases where TCP operates on top of links with own admission and flow control mechanisms (e.g., Frame Relay, ATM), this may lead to situations where the two protocols react destructively resulting in performance deterioration. This is because TCP has no way of distinguishing between packet drops due to link layer control mechanism and those due to genuine congestion situations inside the network. In the present case-study, the resources have been configured so as to eliminate congestion conditions happening in the network. On the other hand, the FR controls may create limited loss situations causing a throttling effect on TCP flows. To get a view of the packet delay performance that has not been impaired by the FR controls, but only by the TCP window size, the CCDF of the packet delays under various load configurations are computed using *Zero CIR* settings. Fig 4 shows the behavior under SRD and LRD conditions, for similar load conditions, low (20%) and high (80%). It is observed that the delay behavior tends to be more heavy-tailed in the case of LRD traffic as the window size is increased whereas in the case of SRD the behavior is same irrespective of the window size. It is also noted that, in the case of LRD traffic and for high load conditions, the buffer at the PVC entry point must be increased to three times more than that of the corresponding SRD traffic conditions, so as to keep the packet loss ratio under 1%.

The effects of FR rate and burst control mechanisms on SRD and LRD traffic are presented in fig 5. In each category, five different control structures are studied in approximately similar utilization profiles. Again, the LRD traffic seems to be sensitive to controls. Even though the packet loss has been limited to below 1%, it acts indirectly in regulating delays because each loss puts the TCP source into the congestion control phase (*multiplicative decrease*) wherein the sources try to readjust to the prevailing conditions. For FR controls, the ratio $\frac{B_e}{B_c}$ seems to be an important parameter. Additionally, under high LRD contents, the absolute values of $B_c$ and $B_e$ are also significant in their impact on the packet delays across the two ends of the PVC.

## VII. End-User Performance

The perception of the user when accessing Internet services is mainly concerned with reliability and responsiveness [30]. The main concerns in reliability aspects of Internet is with irregularities in domain name resolution. On the contrary, the main criteria in responsiveness are page access turnaround time in WWW, and throughput rates in the case of SMTP and FTP. In the previous section, the dependence of packet level performance on various factors such as LRD properties of the traffic, utilization profiles, TCP and FR layer controls, and sizing of buffers was presented. The same factors also influence the end-user performance.

The dependence of access turnaround of WWW on the utilization levels is presented in the scattered plots in fig 6. The network resources and the load profiles were dimensioned such that at least 90% of the WWW page requests complete within a deadline of 1 second. In other words, the deadline of 1 sec was selected as something equivalent to the application layer SLA. Given the simulation conditions, and under LRD conditions, this is achieved by reducing the number of clients to approximately 50% of the SRD levels. The response delay does not seem to depend too much on the utilization levels. As explained in the previous section, the main component of delay comes from the delay experienced by the packets at the WAN ingress point. This delay is mainly determined by the Hurst parameter and is not sensitive to utilization levels.

An important point to consider is the fact that big payloads will entail longer response times. A big postscript document may take a few minutes to download. Thus, it is meaningful to correlate the number of bytes downloaded per page to the respective response times. Accordingly, in fig 6 the relationship is plotted on a log-log scale. A typical usage pattern in WWW may involve a user periodically accessing a particular page

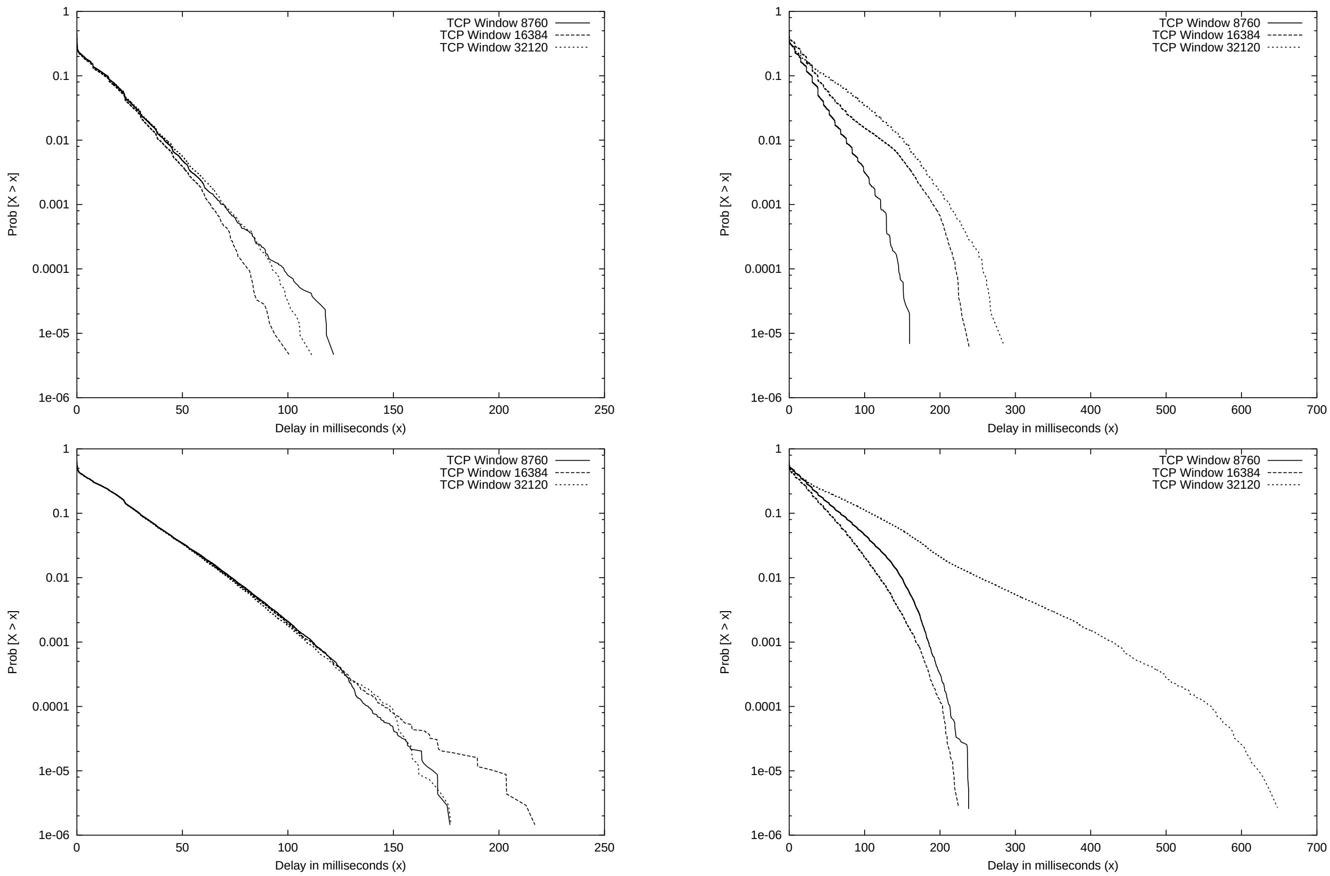


Fig. 4. Effects of TCP window size on SRD (left column) and LRD (right column) traffic, at low utilization levels (top) and high utilization levels (bottom)

(e.g., reading a newspaper or checking stock prices). In such a case, the user would associate a particular response profile with the specific page and thus will be rather sensitive to big variations in that. In such a case, repeatability of the response time is an important performance issue. Given a narrow range of page size in bytes, the spread of response times is observed. The spread widens as the utilization level increases, and this indicates the wide variability in service levels.

The effect of changing the TCP window size on end-user performance is shown in fig 7. This figure depicts the scenario under high utilization and with *Zero CIR* FR control settings. The difference between window sizes of 8760 and 32120 appears to be prominent, and the change in variability in the performance levels is apparent.

The effect of FR rate and burst controls on the end user performance is shown in fig 8. The left column is for $T_c = 0.8$ bits and the right column has $T_c$ set to 0.4 bits. In both cases, the top row is for the case when $\frac{B_e}{B_c} = 2$ and that for the bottom row is when $\frac{B_e}{B_c} = 1$. The effect of the ratio $\frac{B_e}{B_c}$ in controlling the variability in response times within the same page class is evident. It is observed the increase in the population of outliers under stricter burst control regimes of FR. On the contrary, the effect of the parameter $T_c$ can be seen by comparing the figures from left to right within the same row. This parameter seems to have only a marginal impact.

## VIII. Summary

Aspects related to resource dimensioning for best-effort IP networks and the impact of link and transport layer controls have been reported for the case of a simple (simulation) model composed by client and server nodes placed in two sites, which are separated by an Internet cloud. The influence of LRD properties on end-to-end delay performance has been studied as well. Starting with SRD traffic conditions, a bandwidth and buffer allocation scheme has been worked out such as a specified service level is obtained at packet and application layers. Utilization conditions and resource allocation schemes are adjusted for LRD traffic conditions with performance levels being maintained at those obtained under SRD scenario. The packet delay along the end-to-end path across a PVC interconnecting client and server groups has been studied when the port buffers at the routers and switches are dimensioned so as to optimize packet loss rates. The results show that, in this specific case, the queueing delay at the WAN ingress points is important. This delay occurs due to significant bandwidth differences between LAN and WAN link layers. The results also bring out the significant role TCP window size and Frame Relay PVC control mechanisms may play in delay performance of Internet services.

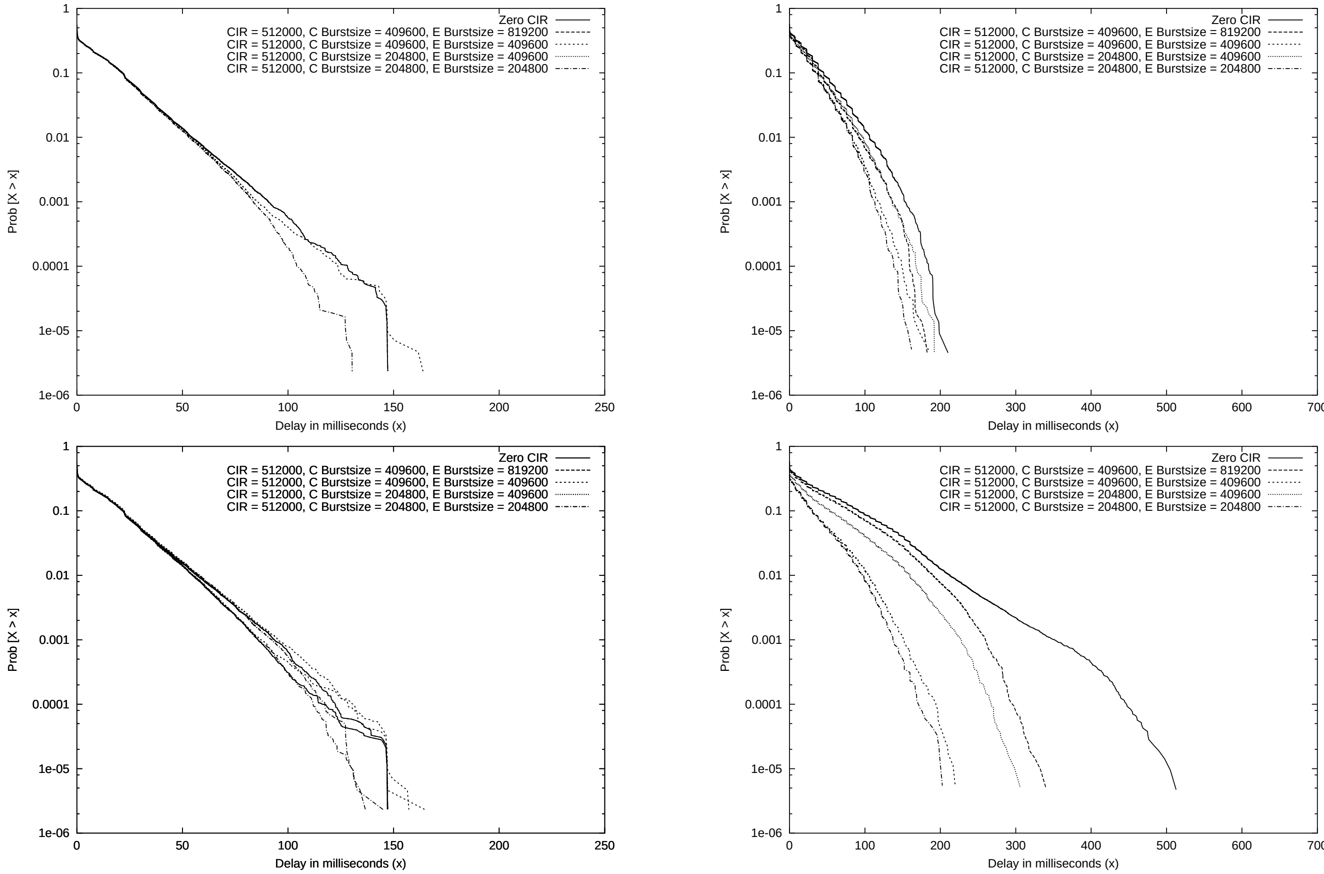


Fig. 5. Effects of FR PVC controls on SRD (left column) and LRD (right column) traffic at medium utilization level and different TCP window sizes: 8760 (top) and 32120 (bottom)

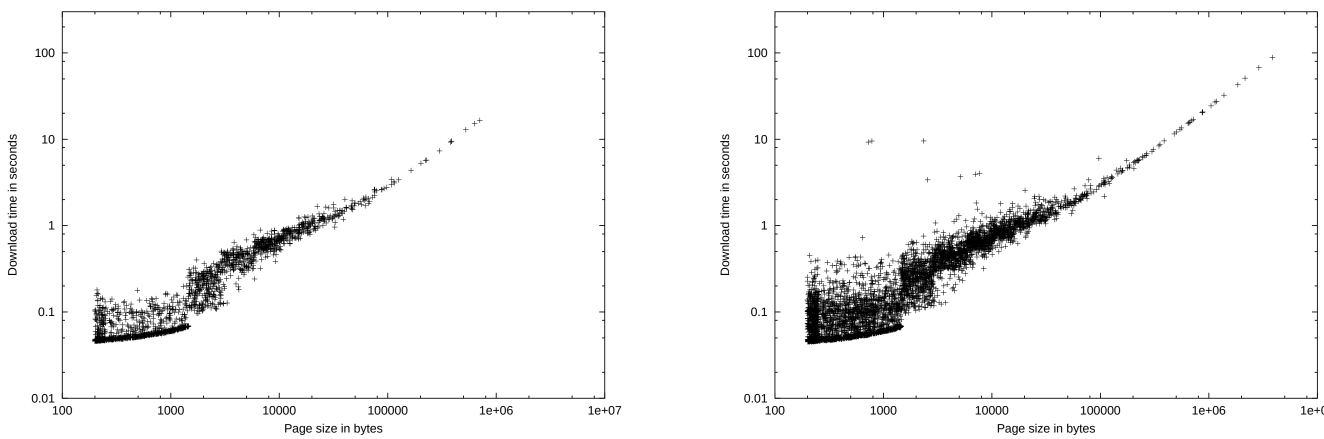


Fig. 6. End-user performance for WWW service at low utilization levels (left) and high utilization levels (right)

# REFERENCES

[1] Abarbanel, B. and Venkatachalam, S., *BGP-4 Support for Traffic Engineering*, work in progress, 2000.

[2] Armitage, G., *Quality of Service in IP Networks*, MacMillan Technical Publishing, USA, 2000.

[3] Ash, G., *Traffic Engineering and QoS Methods for IP-, ATM-, and TDM-Based Multiservice Networks*, Internet Draft, October 2001.

[4] Awduche, D.O., *MPLS and Traffic Engineering in IP Networks*, IEEE Communications Magazine, December 1999.

[5] Awduche D., Malcolm J., Agogbua J., O'Dell M. and McManus J., *Requirements for Traffic Engineering over MPLS*, RFC 2702, September 1999.

[6] Barber, T., Doran, S., Kerrenberg, D., Panigl, C. and Schmitz, J., *Route Flap Dampening*, work in progress, 1997.

[7] Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W., *Long-Range Dependence in Variable-Bit Rate Video Traffic*, IEEE Transactions on Communications, Vol. 43, No. 2/3/4, February/March/April 1995.

[8] Bernet, Y., Yavatkar, R., Ford, P., Baker, F., Zhang, L., Nichols, K. and Speer, M., *A Framework for Use of RSVP with Diff-serv Networks*,

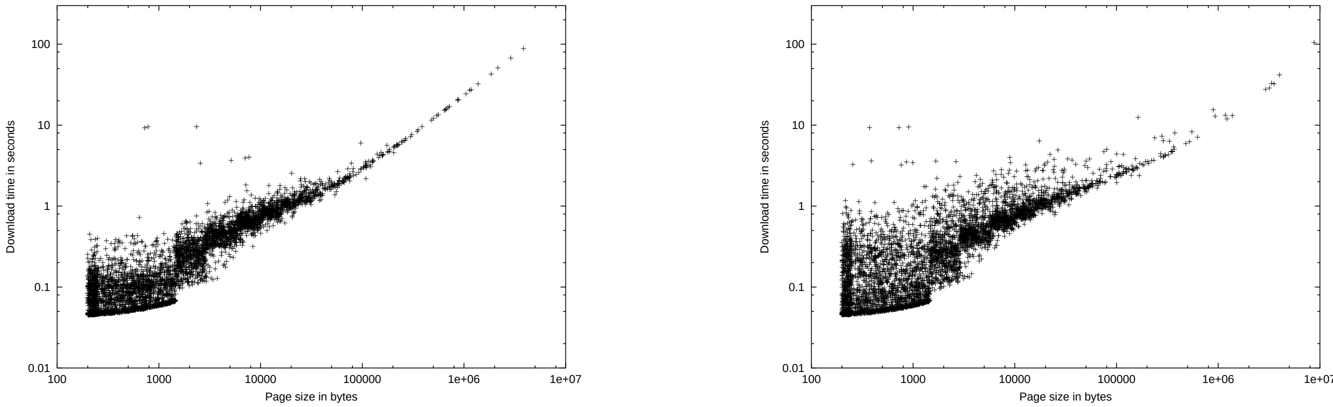


Fig. 7. WWW performance for TCP window size of 8760 (left) and 32120 (right)

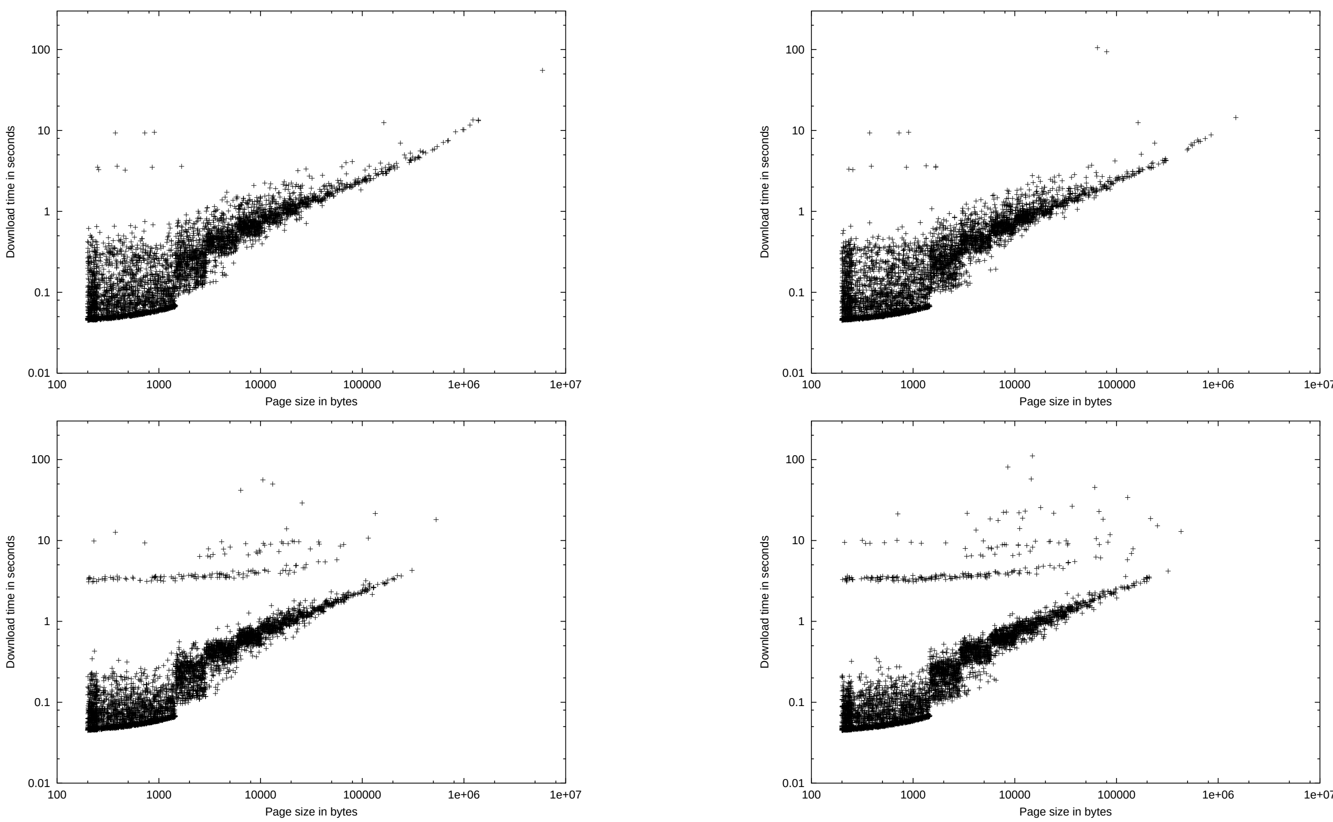


Fig. 8. WWW performance for different FR control regimes: soft control (left column) and tight control (right column), and for larger traffic burstiness accepted (top) and only smaller burstiness accepted (bottom)


Internet Draft, November 1998.

[9] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and Weiss, W., *An Architecture for Differentiated Services*, RFC 2475, December 1998.

[10] Braden, R., Clark, D. and Shenker, S., *Integrated Services in the Internet Architecture: an Overview*, RFC 1633, June 1994.

[11] Braden, R., Zhang, L., Berson, S., Herzog, S. and Jamin, S., *Resource ReSerVation Protocol (RSVP) - Version I Functional Specification*, RFC 2205, September 1997.

[12] Casner, S. and Jacobson, V., *Compressing IP/UDP/RTP Headers for Low-Speed Serial Links*, RFC 2508, February 1999.

[13] Crawley, E., Nair, R., Rajagopalan, B. and Sandick, H., *A Framework for QoS-Based Routing in the Internet*, RFC 2386, August 1998.

[14] Crovella, M.E., Bestavros, A., *Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes*, Proc. of ACM SIGMETRICS, 1996.

[15] Erramilli A., Narayan O. and Willinger W., *Experimental Queueing Analysis with Long-Range Dependent Packet Traffic*, IEEE/ACM Transactions on Networking, Vol. 4, No. 2, April 1996.

[16] Ghanwani, A., Jamoussi, B., Fedyk, D., Ashwood-Smith, P., Li, L. and Feldman, N., *Traffic Engineering Standards in IP Networks Using MPLS*, IEEE Communications Magazine, December 1999.

[17] Jena A.K., *Modeling and Evaluation of Internet Applications*, Licentiate Thesis, ISSN 1101-3931, Dept. of Communications Systems, Lund Institute of Technology, Sweden, 2000.

[18] Jena, A.K., Popescu, A. and Nilsson, A.A., *Modeling and Evaluation of Internet Applications*, research report, ISSN 1103-1581, Blekinge Institute of Technology, Karlskrona, Sweden, 2002.

[19] Katz, D., Yeung, D. and Kompella, K., *Traffic Engineering Extensions to OSPF*, work in progress, February 2001.

[20] Kleinrock L., *Queueing Systems, Volume 2: Computer Applications*, Wiley-Interscience, New York, 1976.

[21] Leland, W., Taqqu, M.S., Willinger, W. and Wilson, D., *On the Self-Similar Nature of Ethernet Traffic (Extended Version)*, IEEE/ACM Transactions on Networking, Vol. 2, No. 1, February 1994.

[22] *Linux Man Pages*, RedHat Linux, http://www.redhat.com.

[23] Mathy, L., Edwards, C. and Hutchison, D., *The Internet: A Global Telecommunications Solution?*, IEEE Network, July/August 2000.

[24] Mil3 Inc., *OPNET Modeler version 6.0*, http://www.mil3.com, Mil3 Inc. Washington DC, USA, 1999.

[25] Norros, I., *A Storage Model with Self-Similar Input*, Queueing Systems Theory and Applications, Vol. 16, Iss. 3-4, 1994.

[26] Paxson, V. and Floyd, S., *Wide Area Traffic: The Failure of Poisson Modeling*, IEEE/ACM Transactions on Networking, Vol. 3, No. 3, June 1995.

[27] Rosen, E., Viswanathan, A. and Callon, R., *Multiprotocol Label Switching Architecture*, RFC 3031, January 2001.

[28] Sangli, S.R., Rekhter, Y., Fernando, R., Scudder, J.G. and Chen, E., *Graceful Restart Mechanism for BGP*, Internet Draft, April 2002.

[29] Smith P., *Frame Relay Principles and Applications*, Addison-Wesley, 1994.

[30] Verma D., *Supporting Service Level Agreements on IP Networks*, Macmillan Technical Publishing, 1999.

[31] Willinger, W., Taqqu, M.S. and Erramilli, A., *A Bibliographical Guide to Self-Similar Traffic and Performance Modeling for Modern High-Speed Networks*, Stochastic Networks: Theory and Applications, Oxford University Press, 1996.

[32] Willinger, W., Paxson, V. and Taqqu, M.S., *Self-Similarity and Heavy-tails: Structural Modeling of Network Traffic*, A Practical Guide to Heavy Tails, Birkhäuser, Boston, 1998.