Workload Generation by Modelling User Behavior in an ISP Subnet*

Christian Kurz, Helmut Hlavacs, Gabriele Kotsis

Institute for Computer Science and Business Informatics
University of Vienna
Lenaug.2/8, 1080 Vienna, Austria
Tel: +43 1 4277 38456, Fax: +43 1 4277 38451
{kurz, hlavacs, gabi}@ani.univie.ac.at

Keywords: network simulation, workload generator, user behavior, ISP

Abstract

In this paper the user-oriented BISANTE workload generator for network traffic simulation is presented. It enables the analyst to model both the behavior of network users interacting with applications or services and the network traffic at the packet level itself. This way, workload at different levels of abstraction can be constructed. Currently the workload generator is available for the simulation tools OPNET and NS.

The behavior of users is represented by hierarchical profiles stored in the profile database. In a case study modelling an internet service provider's (ISP) subnet we show how to create user profiles from observed low-level network data.

I. Introduction

Performance evaluation techniques, in particular techniques to predict the performance in terms of the degree to which QoS requirements can be met, are becoming a key issue in order to design and dimension innovative applications, services, and networks. In all such studies, it is necessary to first study the (expected) traffic characteristics and the impact of the network on the services and applications, as the user-perceived QoS may vary due to varying network conditions.

Simulation is a means of evaluating the effects of those interdependencies, requiring reliable, and prototypical models for the generation of artificial workloads. Traditional network simulators like NS [10] drive their simulation runs by using packet generators that create packets according to given probabilities.

In this paper, we present an approach, where a workload modeling framework is defined allowing network simulator users to specify not only stochastic processes for creating packets but also to model the *behavior of users* running applications which send traffic over a network.

II. Related Work

In the past, network traffic has mainly been modelled by capturing the characteristics of the stochastic process of the packet arrival [8,5]. Here, first and second-order characterstics as well as long-range phenomenons like self-similarity (e.g. in Ethernet, VBR encoded video, and web traffic) have been captured.

In contrast to those resource-oriented modeling approaches, user behavior models try to catch the sequence of user interactions at a higher level. Such models will in general be hierarchical, as the workload generated at a higher level will result in a stream of workload requests at a lower level. Examples for hierarchical user models can be found in [12,2,6,7]. Another popular method for catching inter-command dependencies is to apply Markov Chains for user behavior modeling [1,3,9]. In [4] users are represented by so-called "user equivalents", which are simple two-level heavy tailed On-Off processes. An advanced hierarchical application model has been implemented in the version 7 of the leading industrial network simulator OPNET [11]. Although having comparable features, this approach does not actually catch the behavior of network users,

---

* This work was performed in the BISANTE (Broadband Integrated Satellite Network Traffic Evaluations) Esprit project supported by the EC under research grant EP28425. See http://www.bisante.org

as it is not possible to define reactions to different levels of observed quality of service.

In this paper, we demonstrate, how such a hierarchical framework can be applied in simulation to link the effects of user behavior to the resulting network traffic and to finally being able to model and evaluate user satisfaction by means of simulation.

III. The Bisante Workload Generator

The BISANTE workload modeling methodology [7] is based on a hierarchical approach mapping from user behavior to network traffic.

This approach describes user behavior at different levels of detail. On the highest layer, users arrive at their workstations at a given rate. On the next lower layer, users start sessions, in which they may initiate one or more applications. At the application layer, users may issue commands, for example by clicking on hyperlinks or performing file downloads. At the lowest layer, the actual packets are created and inserted into the network. However, the number of layers is not fixed and may be specified by the user of the workload generator.

By defining simple interfaces between model layers, different modeling approaches can be integrated and the complexity of models can be hidden from the end user if desired. Modelled users may decide what to do next according to random or deterministic behavior, according to the time of day or the quality of service they observe for their application.

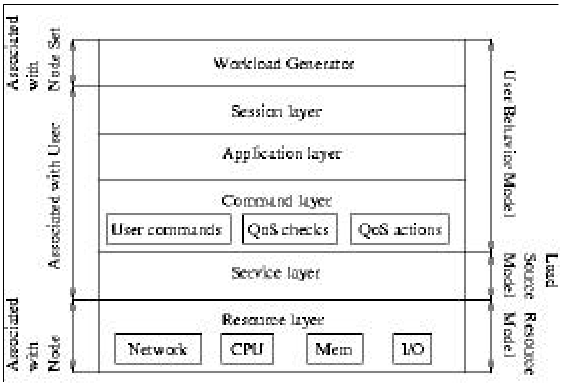

Figure 1: User behavior layers.

IV. User Profiles

User behavior is represented by user profiles, describing a hierarchy of independent stochastic processes exchanging standardised messages among each other. Processes are modelled by finite state machines (FSM), the set of states determining the actions a user is able to perform. Depending on the FSM purpose, different FSM implementations may be chosen from. The user profiles are stored as human readable textfiles, which are organised in terms of a subtree of a Windows or Unix file system. Therefore they can easily be modified for setting up varying simulation parameters using any text editor. The profiles can be shared between all existing workload generator implementations, currently existing for the network simulators OPNET (Windows and Unix) and NS (Linux).

V. Case Study: ISP Subnet

One of the key features of the hierarchical modelling framework is that it provides the modeller with information on user behavior at various levels of detail. In this case study we will show how the data derived from an analysis of user behavior can be used to quantitatively evaluate network performance of a point-of-presence (POP) in Austria, Klagenfurt, belonging to one of the largest internet service providers (ISP) of Austria.

The POP is connected by a CISCO 7206 router to Vienna and therefore to the Internet. The connection consists of a 1984 KB line and a 64 KB backup line. The interconnections of the Klagenfurt POP are 100 Mbps Ethernet lines. Users can dial in via two US Robotics Hypersystems. These are connected to a Cisco W-SC2908-XL Switch which is connected to the router to Vienna.

The two terminal servers are each holding a set of modem banks for ISDN (64 kBit) and analog modem connections (max. 56 kBit). Currently, the terminal servers contains

	ISDN	**Analog**
Terminal Server 1	112	98
Terminal Server 2	64	56

modems. The question to answer in this case study is whether the 2 Mbit/s WAN connection to Vienna is sufficient to transport the data created by the ISP customers.

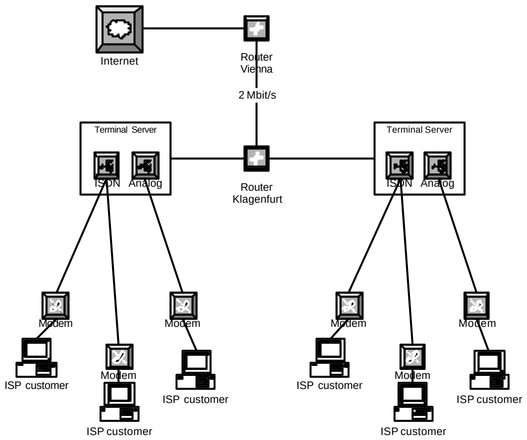

Figure 2: ISP service access point Klagenfurt.

VI. Creating User Profiles

Two logfiles containing one week's data from an ISP are processed to obtain layered user profiles. The first logfile contains data about when and where users connect to the ISP using analog modems or ISDN connections. The stored parameters for each session are a unique user-ID, the start- and stoptime of each session, the IP-address, the connection speed, the modulation type and a parameter indicating whether the connection was established via modem or ISDN. The second logfile describes packets sent across the ISP's subnet. Each record contains information about start- and stoptime of the flow, source IP-address and port, destination IP-address and port, number of bytes transferred, number of packets transferred and the type of IP-protocol used (TCP or UDP).

First an analysis of transferred data volume per source- and destination port was done. These ports were mapped to the services which use them. Afterwards ports with most transfer volume and selected ports with know services were selected. With these ports as one parameter and the users as the other parameter a matrix was built. For each user it was calculated how many bytes were transferred using the selected ports. This table was used to build clusters of users applying different clustering criteria, e.g. transfer volume or similar prozentual transfer volume on the same ports. For clustering, a hierarchical cluster analysis using the Ward-Method and squared Euclidean distance was carried out. To proof that the calculated clusters are significantly different from each other, Kruskal-Wallis tests were performed.

Furthermore, the clustered data was analysed for creating the actual user profiles. For the highest profile level all interarrival times and session length for each user of a cluster was calculated. From this data, two files containing the probabilities for interarrival times and session length were created. The probability to choose a user of a certain cluster was set according to the number of users divided by the sum of all users.

In the next lower profile level users start to use certain ports. Therefore the probability to use a certain port is computed according to the data volume transferred over the port. For each port the interarrival times and the size of packets sent were calculated and store in two respective files. Similar to the data sent by users the data sent back by servers was structured and computed.

VII. Model Validation

The ISP data describes the traffic created by users that dynamically connect their home computers via modem to modem banks belonging to an ISP. The data under consideration was taken from a log file describing the data sent over an ISP access point at Klagenfurt, Austria, over the period of one week.Users sending this traffic have been grouped together into six user classes, where each class will dynamically open connections to several servers in the internet and will then transfer data over their connections.

Data validation is carried out by comparing the traffic generated by the six user classes against the original traffic that was logged during the week. As a validation measure, the aggregated link utilization connecting the access point to the internet is taken.

It can be seen that the simulated models create a traffic comparable to the observed traffic.

The observed deviations stem from the fact that the user profiles are based on packet data recorded in one week, whereas each point of the observed utilization data (x,y) denotes one particular day, where x means the observed user arrival rate on this day, and y the observed average link utilization.

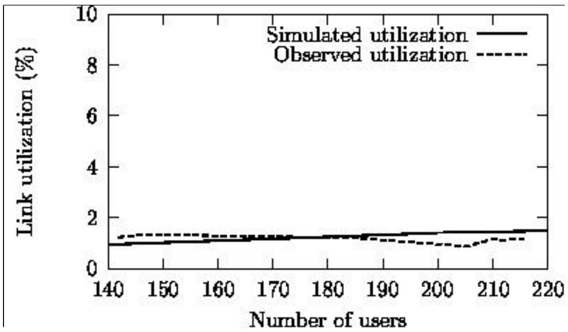

Figure 3: Real vs. simulated utilization.

VIII. Simulation Results

In the following simulation scenario, users connect at a certain rate (users/hour) to one of the terminal servers and create network traffic. The result of this simulation is the link utilization, showing whether the link has still enough capacity for additional connections in the future.

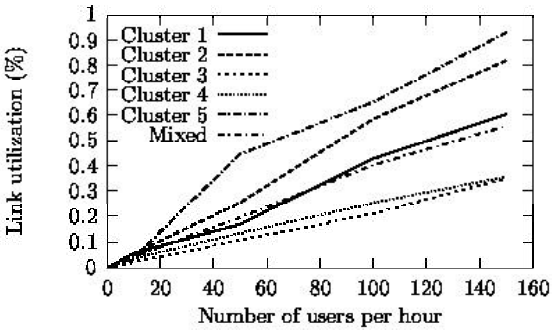

Figure 4: Link utilization depending on the number of users that connect to the access point each hour.

The results show that even if a large number of users connect to the access point, the link still provides enough bandwidth to service all users. Thus, the network traffic created by individual users is very small.

IX. Conclusions

In this paper it is shown how to use the NS implementation of the BISANTE workload generator to simulate an ISP subnet.

The BISANTE workload generator enables to model low-level packet genheration as wells as explicit modelling the behavior of network users reacting to different network conditions.

X. References

[1] M. Calzarossa, R. Marie, K. Trivedi, *System performance with user behavior graphs.* Performance Evaluation 11 (1990) , pp. 155—164.

[2] M. Calzarossa et al., *A hierarchical approach to workload characterization for parallel systems.* In “High Performance Computing and Networking” (B. Hertzberger and G. Serazzi, eds.). Springer LNCS, Berlin 1995, pp. 102—109.

[3] C. Chen. *Structuring and visualizing the www by generalized similarity analysis.* In “Proceedings of Hypertext '97”, 1997.

[4] M. Crovella, *Generating representative web workloads for network and server performance evaluation.* Performance Evaluation Review 26-1 (1998), pp. 151—160.

[5] H. Hlavacs, G. Kotsis, C. Steinkellner, *Traffic source modeling.* Technical Report TR-99101, Institute for Applied Computer Science and Information Systems, 1999.

[6] H. Hlavacs, G. Kotsis, *Modeling User Behavior: A Layered Approach.* In “MASCOTS'99”, IEEE Computer Society, Los Alamitos, California 1999, pp. 218—225.

[7] H. Hlavacs, E. Hotop, G. Kotsis, *Workload generation by modeling user behavior.* In "Proceedings OPNETWORKS 2000", http://www.opnet.com.

[8] D. Jagerman, B. Melamed, W. Willinger, *Stochastic modeling of traffic processes.* In Frontiers in Queuing: Models, Methods and Problems (J. Dshalalow, Edr.), CRC Press, 1996.

[9] D. Menasce, V. Almeida, R. Fonseca, M. Mendes, *Resource management policies for e-commerce servers.* In The 2nd Workshop on Internet Server Performance (WISP 99), 1999.

[10] *The network simulator Ns*, http://www-mash.cs.berkeley.edu/ns/ns.html

[11] *OPNET*, http://www.opnet.com

[12] S. Raghavan, P. Joseph, *Workload models for multi-window distributed environments.* In “Quantitative Evaluation of Computing and Communication Systems”, Springer-Verlag, Berlin 1995.