D I M A C S

Center for Discrete Mathematics and Theoretical Computer Science

DIMACS SPECIAL YEAR ON NETWORKS: 1998 ANNUAL REPORT

DIMACS Center • Rutgers University • 96 Frelinghuysen Road • Piscataway, NJ 08854-8018 • USA
TEL: 732-445-5928 • FAX: 732-445-5932 • EMAIL: center@dimacs.rutgers.edu
WWW: http://dimacs.rutgers.edu/ • TELNET: telnet info.rutgers.edu 90

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs - Research, Bell Labs, NEC Research, and Telcordia Technologies (formerly Bellcore)

Overview:

The Special Year on Networks began in July 1996 and will continue through August 2000 with the help of a three year grant received from the National Security Agency. The program is stressing postdoctoral training and workshops that will reach a broad segment of the DM and TCS research community as well as engage practitioners.

Organizers:

- Stuart Haber, Bellcore / Surety Technologies
- David Johnson, AT&T Labs - Research
- Mihalis Yannakakis, Bell Labs

Scientific Advisory Board: Baruch Awerbuch, Johns Hopkins; Badri Badrinath, Rutgers; Endre Boros, Rutgers; Ding-Zhu Du, Minnesota; David Goodman, Rutgers; Jack Holtzman, Rutgers; Frank Hsu, Fordham; Srinivasan Keshav, Cornell; Arjen Lenstra, Citibank; Udi Manber, Arizona; Ray Miller, Maryland; Debasis Mitra, Bell Labs; Clyde Monma, Bellcore; Panos Pardalos, Florida; K.K. Ramakrishnan, AT&T Labs - Research; Greg Shannon, Spanning Tree Technologies; Scott Shenker, Xerox; Subhash Suri, Washington U. in St. Louis; Eva Tardos, Cornell; Peter Winkler, Bell Labs; and Yacov Yacobi, Microsoft

Sponsors and Collaborating Institutions:

The National Science Foundation
The New Jersey Commission on Science and Technology
The National Security Agency
AT&T Labs - Research
Bell Labs
Bellcore
Institute for Mathematics and its Applications, University of Minnesota
Rutgers University
Princeton University
NEC Research Institute

Theme of Program:

This special year is intended to bring together researchers and practitioners in a wide variety of fields connected to the common theme of "Networks" and ranging from questions of infrastructure to ones of applications. Special emphasis is placed on Internet-related topics, as this is a young and growing area where theoretical computer science and discrete mathematics have a chance to make a real impact.

The special year's topics can be grouped into four overlapping areas:

- **Network Security:** protocols for electronic commerce, public key infrastructure, network security and threat models.
- **Network Applications:** information retrieval, databases and the Web, multimedia transmission and coding, distributed and collaborative computing.
- **Network Control:** routing, congestion and admission control, network protocols, mobile computing.

- **Network Design:** topology and connectivity, switch design, facilities location.

The goal is to engage theoreticians with these problems and foster communication between them and real-world practitioners. Practitioners have much to tell theorists about what the key issues are, and theorists may themselves be able to provide crucial insights and solutions to practitioners.

Workshops, Miniworkshops and Special Events, 1998

- **Workshop on Ethical and Social Dimensions of Information Technology**
  - Dates: February 28-March 1, 1998
  - Location: Princeton University
  - Attendance: 51
  - Organizers: Helen Nissenbaum, Princeton University; Bernard Chazelle, Princeton University; Joan Feigenbaum, AT&T Labs - Research

  The Workshop was dedicated to a discussion of how computer and information systems are shaped by societal and ethical values, including broadly encompassing values such as fair distribution of goods and power, freedom, autonomy, sovereignty, and privacy as well as more specific human ends such as wealth, efficacy, and rights to free expression, association, privacy and property. Panel presenters, representing the fields of computer science, the social sciences, philosophy, and policy studies, discussed values embedded in specific systems, including but not limited to the net, encryption, security, autonomous agents, educational software, user-interfaces, and the structure of information systems. The panelists were guided by questions such as: How do values influence or determine the shape of computer and information systems? Whose ends, interests or values are best, and most frequently, represented in contemporary systems? By what means are values embedded in systems – public policy, markets, or the discretion of individual scientists and engineers? Are some of these sources more "legitimate" than others? What values ought to shape computer and information systems? Is there some shared sense of public, community or individual welfare that ought to drive the design of systems? Is it enough to "let the market decide"?

  There were panels on philosophical perspectives, technical perspectives, social science perspectives, and policy perspectives. There were discussions on democratic implications of Internet privacy by L. Jean Camp; distance education (pedagogical versus economic values) by Marvin J. Croy; metadata platforms by Joan Feigenbaum; oppression, resistance and information technology by Lucas D. Introna; and anonymity, surveillance, and crypotography by Deborah G. Johnson. There were also presentations on how the marriage of management and computing intensifies the struggle for personal privacy by Rob Kling and Jonathan P. Allen; policy in public key infrastructures by B. A. LaMacchia; information technology and freedom by Abbe Mowshowitz; the digital dilemma by W. R. Neuman; and ethics and information technology from a social science perspective by Susan Leigh Star. Finally, Lodis Rhodes talked about choice and control as essential values in a community computer network; M.J. van den Hoven talked about moral responsibility, public office and information technology; and Lorrie Faith Cranor talked about bias and responsibility in neutral social protocols.

- **Workshop on Multichannel Optical Networks: Theory and Practice**
  - Dates: March 16-19, 1998
  - Location: Rutgers University

– Attendance: 52
– Organizers: Peng-Jun Wan, Ding-Zhu Du, Panos M. Pardalos

Recently, widespread interest in information infrastructures has heightened interest in the high performance achievable in multichannel optical networks. WDM networks offer potential advantages, including higher aggregate bandwidth per fiber, new flexibility for automated network management and control, noise immunity, transparency to different data formats and protocols, low bit-error rates, and better network configurability and survivability–all leading to more cost-effective networks. In the last several years, truly astounding progress has been made in optical device technology. The research challenges in multichannel optical networks involve novel approaches to build cost-effective networks from these available device components. The main focus of this workshop was to bring together researchers working on the theoretical and practical aspects of network design problems such as network architecture, interconnection design, medium access control, channel assignment, wavelength routing, scalability analysis, reconfigurability analysis, cost-performance analysis, and network management.

Tutorial survey talks were presented by John Strand on network architectures and transport network design and Biswanath Mukherjee on optical network building blocks, mainly for wavelength-division multiplexing (WDM) and on various architectures, protocols, and algorithms for two classes of optical WDM networks, broadcast-and-select networks for local applications and wavelength-routed networks for wide-area networks. Among the other speakers, G. Persiano and Gordon Wilfong discussed the benefit of wavelength converters and their impact on the wavelength routings in trees and rings; Peng-Jun Wan, Suresh Subramaniam and Thomas Erlebach gave algorithms for maximizing the number of connections with given set of communication requests and the number of available wavelengths; Ornan Gerstel raised the issue of minimizing the SONET ADMs in SONET/WDM ring networks; Rami Melhem, K. Sivalingam and Klaus Jansen discussed the dynamic wavelength/lightpath assignments; and Bruno Beauquier, Eric Harder, Jean-Claude Bermond and Chunling Zhou presented the routing and wavelength assignment for various multicasting/broadcasting communications patterns. Also, Biswanath Mukherjee and Olivier Crochat described how to support fault tolerance and protection; Jadranka Skorin-Kapov, Liwu Liu, Rocky KC Chang and Radim Bartos discussed various aspects of the virtual topology embedding into optical broadcast-and-select optical networks; Chunming Qiao presented the polymorphic control for cost-effective design of optical networks; and Cheng-Chi Yu described the support of real-time services.

- **Workshop on Robust Communication Networks: Interconnection and Survivability**
  – Dates: November 18-20, 1998
  – Location: Rutgers University
  – Attendance: 81
  – Organizers: Nathaniel Dean, Rice University; Frank Hsu, Fordham University; R, Ravi, Carnegie Mellon University

Architectural interconnection and survivability play important roles in the design, construction, operation and application of robust communications networks. Due to the advent of VLSI and fiber optics technologies, it has become possible and feasible to design and construct large scale high performance and high speed wireline and wireless communication networks that are also robust. This opens many challenging issues and problems for both the theory community and practitioners. Topics and problems discussed at the RCN:IS workshop included network structure and architecture (topology/architecture, reliability/latency, routing/switching, and bandwith); WAN backbone and access network design and construction (capacity efficiency, cost, restoration methods, congestion control, trade-offs between different architectures, and interconnection between LANs and the Internet); container construction and fault tolerant properties (multipath, all pair shortest paths, multicast, wide-diameter and fault-diameter); and primal dual approximation methods (survivable network design). Other topics included highly connected networks (including k-connected graphs, algorithms for special classes of graphs, etc.) and vulnerability (minimum degree problems, network upgrading, network attack and inhibition); survivability in optical networks (double coverings, SONET, T/W/CDM) and wireless networks (switch assignments, interconnections between base stations); and network design support systems (design heuristics, add/delete algorithms) and novel networks.

Selected Research Results

The following **abbreviations** are used throughout: RU = Rutgers University, PU = Princeton University, ATTL = AT&T Labs, BL = Bell Labs, BC = Bellcore, NEC = NEC Research Institute, PM = permanent member of DIMACS, PD = DIMACS postdoctoral fellow, GS = Graduate Student, V = DIMACS visitor.

Synchronization and Packet Routing in Networks Boaz Patt-Shamir (V) worked jointly with Rafi Ostrovsky (PM, BC) and solved an important open problem: Can optimal synchronization be attained efficiently when clocks are drifting (except for a "source" clock which provides standard time), if communication is constrained in some reasonable way? It was well known that if clocks are drift-free, then there is an optimal efficient solution; it is also known that if clocks can drift and communication is arbitrary, then no algorithm is both efficient and optimal. In this work Ostrovsky and Patt-Shamir gave an algorithm whose complexity is polynomial in the number of "pending messages" (messages the processor knows to be sent, but the processor does not know when received). This result has immediate applications, since virtually all synchronization algorithms in use today have bounded (typically linear in the number of processors) pending messages in any given time.

Efficient Scheduling Algorithms under the "Adversarial Queueing Theory" Model Bill Aiello (PM, BC, now ATTL), Rafi Ostrovsky (PM, BC) and Boaz Patt-Shamir (V) looked at the question of efficient scheduling algorithms under the "adversarial queuing theory" model. They partially settled the primary open problem in this field by proving that the "longest in system" policy builds up queues of only linear size when the paths used are shortest paths, and they gave an algorithm guaranteeing polynomial size queues for any paths.

Strictly Nonblocking Switching Networks Frank Hwang (PM) published a major new volume on switching networks for telephone traffic. Benes published *Mathematical Theory of Connecting Networks and Telephone Traffic* in 1965. Hwang's book is the first attempt to update the classical theory on nonblocking and rearrangeable networks with the thirty-some years of intensive research since 1965. But it also brings in new subject matter like multimedia and broadcast networks. In particular, the book includes a new sufficient condition on strictly nonblocking networks which was applied to the EGS photonic switching patented by AT&T. The paper describing this result was delivered in preliminary form by Richards and Hwang at the DIMACS workshop on Nonblocking Switching Networks.

Placement Algorithms for Hierarchical Cooperative Caching A basic problem in distributed memory environments is to provide efficient access to remote objects (e.g., words of memory, files). This problem arises in several applications including distributed file systems and the World Wide Web. For example, consider a hierarchical network of machines in which each machine periodically issues a request for an object drawn from a fixed set of unit-size objects. Assume that: (i) the frequency with which each machine accesses each object is known, (ii) each machine has a cache of known capacity, (iii) any cache can be accessed by any machine, and (iv) any request is satisfied by the closest machine with a copy of the desired object. In such an environment, it is desirable to fill the available cache space with copies of objects in such a way that the average access cost is minimized. Madhukar Korupolu, Greg Plaxton and Rajmohan Rajaraman (PD) obtained both exact and approximate polynomial-time algorithms for this hierarchical placement problem. Their exact algorithm is based on a reduction to min-cost flow, and does not appear to be practical for large problem sizes. Thus they are motivated to search for a faster approximation algorithm. Their main result is a simple constant factor approximation algorithm for the hierarchical placement problem that admits an efficient distributed implementation.

Dynamic Replication on the Internet Michael Rabinovich (ATTL) and Rajmohan Rajaraman (PD) analyzed a dynamic replication protocol designed by Michael and Irene Rabinovich. This protocol consists of an algorithm for deciding on the number and location of object replicas and an algorithm for distributing requests among currently available object replicas. Their analysis showed that if the frequency with which each machine accesses each object is fixed, then the system reaches a state in which the total number of replications, deletions, and migrations is bounded with high probability.

Contention Resolution in Multiple Access Networks Micah Adler (V), Phil Mackenzie, and Rajmohan Rajarman (PD) studied the problem of scheduling the transmission of packets from multiple senders over a number of Ethernet-like multiple-access networks. In particular, they were interested in exploiting the additional power allowed by the ability of a transmission to choose from multiple channels. They demonstrated a protocol that, when p senders transmit over $m \geq \log p$ channels, is stable and has an expected delay for any message of $O(1)$, if the arrival rate at each sender is at most $\lambda m/p$ for a constant λ sufficiently small. Their protocol has several advantages over existing protocols: it is considerably simpler and has much better constants than any previous $O(1)$ expected delay protocol. In addition, this is the first $O(1)$ expected delay protocol for asynchronous environments. All previous $O(1)$ delay protocols required a tightly synchronized clock and all previous protocols for asynchronous systems had an expected delay of $\Omega(\log p)$. This protocol works for pessimistic models of asynchrony. They also showed a lower bound on the delay of any protocol if the arrival rate is sufficiently high. In particular, if the arrival rate at each sender is at least $\lambda m/p$, where λ is a sufficiently high constant less than 1, then the average expected delay of any protocol is $\Omega(p)$. This bound is tight because a natural generalization of the round-robin protocol to multiple channels incurs $O(p)$ delay for any $\lambda < 1$.

Scheduling to Minimize Average Stretch Traditional metrics for analyzing on-line scheduling algorithms such as average completion time, average flow time, and maximum flow time, do not take into account the differing job sizes in the objective function. That is, these metrics tend to treat the delay incurred by a job in a uniform manner regardless of the size of the job. Thus, the penalty for a large delay incurred by a small job is identical to the penalty for the same delay incurred by a large job. For several applications, including scheduling in web servers, this is undesirable. In order to address this problem, Bender, Chakrabarti, and Muthukrishnan introduced a new metric called stretch. The stretch of a job in a schedule is the ratio of its flow time (the time the job spends in the system) to its job size. Rajmohan Rajaraman (PD) and Anthony Shaheen (DIMACS REU student) studied the on-line complexity of minimizing average stretch. They showed that a well-known scheduling algorithm, the shortest remaining processing time algorithm (SRPT), is 4-competitive with respect to average stretch. S. Muthukrishnan (PM, ATTL) and Rajaraman are studying the complexity of average stretch in multi-processor and non-clairvoyant (where the size of a job is not known a priori) environments.

Allocation of Resources in a Multimedia Network Tina Heikkinen (GS, RU) worked on the modeling of a telecommunication network applying linear programming. Given a telecommunication system described in terms of a matrix G describing the link fading (gain) coefficients of the users of the network and a second matrix F describing the interference fading coefficients, the general problem is to define a resource allocation rule, i.e., an allocation of transmit powers $x = (x_1, .., x_n)$ of the n users, such that $(Gx)_i/(Fx)_i > a$, for all i. Heikkenen extended this framework to the case of multimedia. Two aspects of multimedia were modelled; first, by allowing for non-square matrices G and F each user was allowed to have access to a multiplicity of datatypes, modelled by varying the user-by-user data rates; second, temporal dynamics were introduced, to capture the dynamics of the resource allocation problem. The latter is quite new; the modelling of dynamics was done by defining a dynamic game against nature.

The Capacity of a Multimedia Network Continuing with the study of resource allocation in multimedia networks described in the previous example, Tina Heikkenen (GS, RU) took the novel approach of applying the theory of non-cooperative games to define a globally optimal capacity denoted by $a(*)$, i.e., formally, the von Neumann growth model was inverted to max a with respect to x, s.t. for all i, $(Gx)_i/(Fx)_i > a$. Various characterizations for $a(*)$ were obtained. In continuing work on this topic, allowing for incomplete knowledge about the coefficients of the matrices G and F, Heikkenen used game theory/linear programming, with the recognition that due to the interference matrix, the game is an externality game and the resource allocation problem is an allocation problem under externalities. Knowledge of economic resource allocation under externalites was applied. Two solutions for resource allocation under externalities were described. First, it is possible to show that there exist general classes of matrix games with same integral optimal capacity. These games are called exchangeable. This result follows from the Leontief-property of the matrix G. Second, a Boolean formulation for the Leontief-von Neumann telecommunication system was presented, allowing for nonexchangeable matrix games; this formulation established a connection between zero-sum noncooperative games and cooperative games described by Boolean (characteristic) functions.

Enforcing E-Commerce Policies Using Structured Messages Naftaly Minsky (RU) and Victoria Ungureanu (GS, RU) have been working on the formulation of policies for electronic commerce and their enforcement. They proposed a general mechanism that can be used to formulate and enforce a wide range of policies for electronic commerce in a unified manner. This mechanism makes a strict separation between the formal statement of a policy and its enforcement. The policy is stated explicitly and can be examined by anybody who wishes to engage in it, and it is enforced by a distributed set of trusted controllers that regulate the activities of each agent under this policy.

Fast, Fair and Frugal Bandwidth Allocation in ATM Networks In connection with ATM networks, convergence issues arise for various algorithms that have been proposed in the context of the rate allocation problem. In this problem the aim is to allocate the bandwidth requests of various sessions such that the allocations are fair and capacity constraints are met. ATM networks are used to carry a variety of types of traffic. For some types of traffic, in particular Available Bit

Rate (ABR) traffic, the bandwidth of a network is typically insufficient to satisfy the requests of all the sessions, and so some fair allocation scheme must be devised. The ATM Forum, the standards setting body for ATM networks, has specified that the fairness criterion for ABR traffic should be Max-Min Fairness, which intuitively means that raising the bandwidth of any session comes at the expense of some other session of no greater bandwidth. Protocols to allocate bandwidth to sessions in a max-min fair manner are an important part of a network design. For a protocol to be realistic, it must conform to the Resource Management (RM) Cell mechanism specified by the ATM Forum. Such RM Cells get sent as a constant fraction of all cells sent by the source. However, they have only a few fields. RM Cells are the only means of communication allowed between switches so any reasonable protocol is totally distributed and asynchronous, since the RM Cell mechanism does not easily lend itself to synchronization. Finally, RM Cells must be handled very quickly at the switches. Call a protocol frugal if the switch spends O(1) computation on each RM Cell it receives, and it has O(1) local space for each session through it. Recently, several frugal RM Cell protocols have been proposed for ABR traffic, but none have been shown to converge to the max-min fair state. Protocols which are known to converge in a linear number of maximum roundtrip times require RM Cell processing which is linear in the number of sessions at a switch. Y. Bartal (BL), Martin Farach (PM, RU), Shibu Yooseph (PD), and Lisa Zhang (PM, BL) gave a frugal RM Cell protocol for ABR that matches the convergence time of the fastest known non-frugal protocol. A second type of ABR traffic is the Minimum Cell Rate (MCR) type, where every session can specify a minimum amount of bandwidth. The max-min fair allocation should then respect these MCR requests. They extended their results to give the first frugal RM Cell protocol for MCR and achieve a quadratic convergence rate.

Broadcast Disk Paging Broadcast disks are an emerging paradigm for massive information dissemination. In a broadcast disk, data is divided into n equal-sized pages and pages are broadcast in a round-robin manner by a server. Broadcast disks are effective because many clients can simultaneously retrieve any transmitted information. Paging is used by the clients to improve performance, much as in virtual memory systems. However, paging on broadcast disks differs from virtual memory paging in at least two fundamental aspects: A page fault in the broadcast disk model has a variable cost that depends on the requested page as well as the current state of the broadcast; prefetching is both natural and a provably essential mechanism for achieving significantly better competitive ratios in broadcast disk paging. Vincenzo Liberatore (GS, RU) designed a deterministic algorithm that uses prefetching to achieve an $O(n \log k)$ competitive ratio for the broadcast disk paging problem, where k denotes the size of the client's cache. He also showed a matching lower bound of $\Omega(n \log k)$ that applies even when the adversary is not allowed to use prefetching. In contrast, when prefetching is not allowed, no deterministic online algorithm can achieve a competitive ratio better than $\Omega(nk)$. Moreover, he obtained a lower bound of $\Omega(nlogk)$ on the competitive ratio achievable by any non-prefetching randomized algorithm against an oblivious adversary. These lower bounds are trivially matched from above by known results about deterministic and randomized marking algorithms for paging. An interpretation of these results is that in broadcast disk paging, prefetching is a perfect substitute for randomization.

Associated Projects on Network Modeling and Parallel Simulations

DIMACS has two multi-disciplinary, multi-institutional research projects in network modeling and parallel simulations. The participants in these projects, both DIMACS members and outsiders, have repeatedly emphasized that the projects would not have happened without the organizational infrastucture provided by DIMACS. DIMACS brings scientific expertise and experience with large, complex projects that involve a variety of institutions and disciplines, and such projects provide opportunities for students, postdocs, faculty members, and industrial researchers to share in research experiences and learn more about each others' "cultures" and ways of doing science. Importantly, DIMACS provides a neutral site for collaboration: Industrial researchers find it easier to work on a joint high risk precompetitive project under the umbrella of an academic center which enhances the opportunities for interactions.

The two funded research efforts on simulations of integrated communications systems and scalable self-organizing simulations can be directly traced to collaborations initiated during our Special Year on Massively Parallel Computing, and are now closely related to our continuing Special Year on Networks. The first is funded by NSF and the second, followup project by DARPA. We describe these two projects together in this section. In 1998, WINLAB received an NSF Instrumentation Grant for a 14-processor SUN ultrasparc that is used in these simulation projects.

Project on Simulations of Integrated Communications Systems

This project is funded by the National Science Foundation for a three year period over 1995 - 1998, for a total of $1.5M.

Participating Scientists: Sandeep Bhatt (PM, BC and RU), Richard Fujimoto (Georgia Tech), Albert Greenberg (PM, ATTL), Jim Kurose (University of Massachusetts), David Nicol (William and Mary / Dartmouth), Andrew Ogielski (PM, RU), Diane Souvaine (PM, RU), Don Towsley (University of Massachusetts), and Paul Wright (PM, ATTL).

Collaborations: This project brings together experts in network modeling, design and analysis, distributed algorithms, and high performance discrete-event simulation to focus on the development of high-performance simulation tools for broadband networks and wireless communication services. This project directly brings together the expertise of scientists from two of the nation's leading telecommunications companies, in partnership with leading academic sites; and includes graduate and postdoctoral training that benefits students through exposure to industrial scale problems and internships.

Scalable Self-Organizing Simulations (S3)

This project is funded by DARPA for a three year period over 1996-1999, for a total of $3.1M. It represents a highly successful transition of the NSF-supported networks simulations project to DARPA. The theoretical underpinnings developed under NSF support are leading, in this project, to significant practical code creation and applications.

Participating Scientists: Sandeep Bhatt (PM, BC and RU), Marina Chen and Mark Crovella (Boston University), James Cowie (Cooperating Systems Corporation), Richard Fujimoto, Kalyan Perumalla and Ellen Zegura (Georgia Institute of Technology), Don Leskiw (Ultra Corporation), David Nicol (Dartmouth), Andrew Ogielski (PM, RU).

Collaborations: The daunting complexity of modern communications technologies on the one hand, and of the large software systems required on the other, demands that a project of this sort be carried out by an interdisciplinary team of experts in parallel simulations, networking, software engineering, and defense applications. We have put together a team of scientists and engineers from six institutions to work on the project, including participants from two small companies. Academic collaborators involve in the project postdoctoral fellows and many graduate research assistants.

Objective of the Projects

The goal of these projects is to achieve a generational transition in speed, scalability and manageability of simulations of very large scale multiprotocol communication networks. The challenge of enabling discrete event simulations of hundreds of thousands to millions of communicating entities requires novel, scalable network modeling techniques, and new, lightweight architectures for ultra-fast parallel simulation kernels. Current practice of network modeling and simulation cannot even approach problems of such size and complexity.

Approach

The success of these projects requires a significant collaboration across three domains: parallel discrete event simulations, software engineering, and network modeling. Therefore, the high level view of the project structure is as follows:

First, core software research involving the design and implementation of innovative software capable of reaching the "million entities target", including: i) Scalable, simulator-independent and user-friendly network modeling framework; ii) Very high speed, process-oriented parallel simulation kernels capable of supporting millions of logical processes; and iii) Highly usable, extensible runtime software environment. These project components will deliver the tools capable of efficient modeling and fast interactive simulations of very large communication networks.

Second, network modeling research on efficient model representations for network elements and protocols that are natural for networking researchers and engineers and strongly promote scalability; and on automation of network topology and traffic aggregation methods that can reduce model complexity without loss of accuracy. Very large models used in networking research also require generation of network topologies that approximate the global Internet, and automation of address and bandwidth/delay assignment in such models.

Third, practical validation of the new techniques by demonstration of scalability of simulations for a wide range of network types and protocol classes. Within these projects, these include: i) the Internet and ATM/PNNI internetworks - protocols, topology, traffic sources; and ii) Mobile wireless networks - protocols, radio propagation. External users have been experimenting with a variety of other models. The testing process involving internal and external users feeds back corrections and new ideas to simulator and tool developers. This project component will deliver a database of important models together with canonical methods for efficient parallelization.

Recent Project Accomplishments

S3 parallel simulators have demonstrated event execution rates and entity counts on realistic models well in excess of those achievable using the best tools available elsewhere.

New very high performance conservatively synchronized parallel simulator, Nops, was designed and implemented. Simulation speed exceeding million events/second, and scalability to several hundred thousand network elements were demonstrated.

Performance of the GTW optimistically synchronized parallel simulator improved over 100%, and scalability to over a million network elements was demonstrated.

An extensive suite of portable graphical control and visualization tools for TeD (Jane) was developed. TeD/GTW modeling capability was extended to arbitrarily nested process suspension.

This capability is not available in any other optimistic simulator.

The next generation modeling framework SSF was designed together with the reference Java implementation atop the JTED simulator. Scalability in parallel and distributed simulations was demonstrated. High performance implementation of SSF atop the Nops simulator is nearing completion. Progress was made in implementation of MegEnta, the ultra-high performance successor to GTW that will implement the SSF interface.

Novel techniques for simulated packet traffic aggregation were implemented. 10-fold increase of simulation speed and high accuracy were demonstrated in a model of ATM internetwork.

The model of long-range correlated (fractal) Web traffic generator Surge was implemented in SSF/JTED. Surge's original configuration is used worldwide, and has been adopted by the World Wide Web Consortium (W3C) as a core component of their HTTP-NG testbed.

Testing of complex models of mobile wireless networks was concluded. Parallel speedup and scalability with varying detail level were demonstrated.

Testing of ATM internetwork models with full PNNI protocols was concluded. Excellent parallel speedup and scalability with network size were demonstrated.

Porting of protocol models from other sequential simulators (ns) to TeD was demonstrated, and simulation scalability was shown.

The first prototype system for dynamic hierarchical visualization of very large network graphs was completed.

As a result of our work, this is the first time anybody can actually model and simulate the global Internet with a hundred thousand hosts and routers, running IP packets with real routing protocols at a rate of one million packet hops per second (more on newer machines).

Technology Transfer

The complete parallel simulation package Jane/TeD/GTW is being actively used on networking projects in over 20 corporations and universities, including Lucent Bell Labs, Bellcore, Newbridge Networks, Traveller Multimedia Networks, Mitre Corp., JPL, and also in projects in other areas, such as virtual factory simulations at Gintic Inst. of Manufacturing.

The GT-ITM network topology generation package is in use in several network research projects, including the ns v2 simulator (project VINT at UCB, USC, LBL), U. Oregon, and others.

Communications application requirements were coordinated with Air Force Research Laboratory (AFRL) focusing on the Air Force C2 initiatives for Global Awareness, Global Grid, and Dynamic Planning and Execution. Transitioning of the S3 technologies to Air Force simulation needs was defined.

Models of networks developed in project S3 for parallel simulation capability testing are being used in networking research. The WiPPET Wireless testbed is extensively used in research on wireless IPv6 protocols at WINLAB, Rutgers. Several wireless networking companies and ARMY CECOM have expressed interest in this package, and further distribution will be made.