Active

Project #: C-36-X57 Cost share #: C-36-365 Rev #: 0
Center # : 10/24-6-R8547-0A0 Center shr #: 10/22-1-F8547-0A0 OCA file #:
Work type : RES
Contract#: CDA-9422033 Mod #: Document : GRANT
Prime #: Contract entity: GTRC

Subprojects ? : N CFDA: 47.070
Main project #: PE #: N/A

Project unit: COMPUTING Unit code: 02.010.300
Project director(s):
SCHWAN K COMPUTING (404)894-2589

Sponsor/division names: NATL SCIENCE FOUNDATION / GENERAL
Sponsor/division codes: 107 / 000

Award period: 950615 to 960531 (performance) 960831 (reports)

Sponsor amount	New this change	Total to date
Contract value	104,000.00	104,000.00
Funded	104,000.00	104,000.00
Cost sharing amount		52,000.00

Does subcontracting plan apply ?: N

Title: INTERACTIVE DISTRIBUTED COMPUTING

PROJECT ADMINISTRATION DATA

OCA contact: Jacquelyn L. Bendall 894-4820

Sponsor technical contact

CAROLINE WARDLE
(703)306-1980

NATIONAL SCIENCE FOUNDATION
4201 WILSON BLVD.
ARLINGTON, VA 22230

Sponsor issuing office

ANNE B. TENNEY
(703)306-1212

NATIONAL SCIENCE FOUNDATION
4201 WILSON BLVD.
ARLINGTON, VA 22230

Security class (U,C,S,TS) : U
Defense priority rating : N/A
Equipment title vests with: Sponsor GIT X

ONR resident rep. is ACO (Y/N): N
NSF supplemental sheet

Administrative comments -
INITIATION OF PROJECT. THIS IS AN EQUIPMENT GRANT.

SK494

U

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

NOTICE OF PROJECT CLOSEOUT

Closeout Notice Date 09/03/96

Project No. C-36-X57 ____ Center No. 10/24-6-R8547-0A0_

Project Director SCHWAN K ____ School/Lab COMPUTING ____

Sponsor NATL SCIENCE FOUNDATION/GENERAL ____

Contract/Grant No. CDA-9422033 ____ Contract Entity GTRC

Prime Contract No. ____

Title INTERACTIVE DISTRIBUTED COMPUTING ____

Effective Completion Date 960531 (Performance) 960831 (Reports)

Closeout Actions Required:	Y/N	Date Submitted
Final Invoice or Copy of Final Invoice	N	____
Final Report of Inventions and/or Subcontracts	N	____
Government Property Inventory & Related Certificate	N	____
Classified Material Certificate	N	____
Release and Assignment	N	____
Other ____	N	____

Comments ____
LETTER OF CREDIT APPLIES. 98A SATISFIES PATENT REPORT. ____

Subproject Under Main Project No. ____

Continues Project No. ____

Distribution Required:

Project Director	Y
Administrative Network Representative	Y
GTRI Accounting/Grants and Contracts	Y
Procurement/Supply Services	Y
Research Property Managment	Y
Research Security Services	N
Reports Coordinator (OCA)	Y
GTRC	Y
Project File	Y
Other ____	N
____	N

OMB Number 3145-0058

1

NATIONAL SCIENCE FOUNDATION
4201 Wilson Blvd.
Arlington, VA 22230

PI/PD Name and Address

Professor Karsten Schwan
Georgia Institute of Technology
College of Computing
801 Atlantic Drive
Atlanta, GA 30332-0280

NATIONAL SCIENCE FOUNDATION FINAL PROJECT REPORT

PART I - PROJECT IDENTIFICATION INFORMATION

1. Program Official/Org.	Caroline Wardle, CDA
2. Program Name	CISE, Office of Cross-Disciplinary Activities
3. Award Dates (MM/YY)	From: 6/15/95 To: 5/31/96
4. Organization and Address	Georgia Tech Research Corp. Administration Building Atlanta, GA 30332
5. Award Number	CDA-9422033
6. Project Title	Interactive Distributed Computing

NSF Grant Conditions (Article 17, GC-1, and Article 8, FDP-11) require submission of a Final Project Report (NSF Form 98A) to the NSF Program Officer no later than 90 days after the expiration date of the award. Final Project Reports for expired awards must be received before new awards can be made (NSF Grants Policy Manual Section 340).

Below, or on a separate page attached to this form, provide a summary of the completed projects and technical information. Be sure to include your name and award number on each separate page. See below for more instructions

PART II - SUMMARY OF COMPLETED PROJECT (for public use)

The summary (about 200 words) must be self-contained and intelligible to a scientifically or technically literate reader. Without restating the project title, it should begin with a topic sentence stating the project's major thesis. The summary should include, if pertinent to the project being described, the following items:

- The primary objectives and scope of the project
- The techniques or approaches used only to the degree necessary for comprehension
- The findings and implications stated as concisely and informatively as possible

PART III - TECHNICAL INFORMATION (for program management use)

List references to publications resulting from this award and briefly describe primary data, samples, physical collections, inventions, software, etc., created or gathered in the course of the research and, if appropriate, how they are being made available to the research community. Provide the NSF Invention Disclosure number for any invention.

I certify to the best of my knowledge (1) the statements herein (excluding scientific hypotheses and scientific opinion) are true and complete, and (2) the text and graphics in this report as well as any accompanying publications or other documents, unless otherwise indicated, are the original work of the signatories or of individuals working under their supervision. I understand that willfully making a false statement or concealing a material fact in this report or any other communication submitted to NSF is a criminal offense (U.S. Code, Title 18, Section 1001).

	8-29-96
Principal Investigator/Project Director Signature	Date

IMPORTANT:
MAILING INSTRUCTIONS
Return this *entire* packet plus all attachments in the envelope attached to the back of this form. Please copy the information from Part 1, Block I to the *Attention block* on the envelope.

NSF Form 98A (Rev. 7/95)

PART IV - FINAL PROJECT REPORT - SUMMARY DATA ON PROJECT PERSONNEL

	SENIOR STAFF		Post Doctoral		Graduate Student		Under Graduate Student		OTHER	
	M	F	M	F	M	F	M	F	M	F
A. Total U.S. Citizen					15	4				
B. Total, Permanent Residents										
U.S. Citizens or Permanent Resident:										
American Indian or Alaskan Native Asian										
Asian										
Black, Not of Hispanic Origin					3					
Hispanic										
Pacific Islander										
White, Not of Hispanic Origin					12	4				
C. Total, Other Non-U.S. Citizens					9	2				
Specify Country 1. India										
2. Costa Rica										
3. Rumania										
D. Total, All Participants (A+B+C)					**24**	**6**				

PART IV — FINAL PROJECT REPORT — SUMMARY DATA ON PROJECT PERSONNEL

(To be submitted to cognizant Program Officer upon completion of project)

The data requested below are important for the development of a statistical profile on the personnel supported by Federal grants. The information on this part is solicited in response to Public Law 99-383 and 42 USC 1885C. All information provided will be treated as confidential and will be safeguarded in accordance with the provisions of the Privacy Act of 1974. You should submit a single copy of this part with each final project report. However, submission of the requested information is not mandatory and is not a precondition of future award(s). Check the "Decline to Provide Information" box below if you do not wish to provide the information.

Please enter the numbers of individuals supported under this grant.
Do not enter information for individuals working less than 40 hours in any calendar year.

	Senior Staff		Post-Doctorals		Graduate Students		Under-Graduates		Other Participants[1]	
	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.
A. Total, U.S. Citizens										
B. Total, Permanent Residents										
U.S. Citizens or Permanent Residents[2]: American Indian or Alaskan Native . .										
Asian										
Black, Not of Hispanic Origin										
Hispanic										
Pacific Islander										
White, Not of Hispanic Origin . . .										
C. Total, Other Non-U.S. Citizens										
Specify Country 1.										
2.										
3.										
D. Total, All participants (A + B + C)										
Disabled[3]										

☐ Decline to Provide Information: Check box if you do not wish to provide this information (you are still required to return this page along with Parts I-III).

[1]Category includes, for example, college and precollege teachers, conference and workshop participants.

[2]Use the category that best describes the ethnic/racial status to all U.S. Citizens and Non-citizens with Permanent Residency. *(If more than one category applies, use the one category that most closely reflects the person's recognition in the community.)*

[3]A person having a physical or mental impairment that substantially limits one or more major life activities; who has a record of such impairment; or who is regarded as having such impairment. *(Disabled individuals also should be counted under the appropriate ethnic/racial group unless they are classified as "Other Non-U.S. Citizens.")*

AMERICAN INDIAN OR ALASKAN NATIVE: A person having origins in any of the original peoples of North America and who maintains cultural identification through tribal affiliation or community recognition.

ASIAN: A person having origins in any of the original peoples of East Asia, Southeast Asia or the Indian subcontinent. This area includes, for example, China, India, Indonesia, Japan, Korea and Vietnam.

BLACK, NOT OF HISPANIC ORIGIN: A person having origins in any of the black racial groups of Africa.

HISPANIC: A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.

PACIFIC ISLANDER: A person having origins in any of the original peoples of Hawaii, the U.S. Pacific territories of Guam, American Samoa, and the Northern Marinas; the U.S. Trust Territory of Palau; the islands of Micronesia and Melanesia; or the Philippines.

WHITE, NOT OF HISPANIC ORIGIN: A person having origins in any of the original peoples of Europe, North Africa, or the Middle East.

NSF Form 98A (Rev. 7/95)

Interactive Distributed Computing

Karsten Schwan, Mustaque Ahamad, Mostafa Ammar, Richard Fujimoto, and Scott Hudson

College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332-0280

This NSF CISE equipment grant has impacted Georgia Tech far beyond its actual funding level for two reasons: (1) due to the simultaneous awards of a substantial NSF Research Infrastructure grant and of an engineering equipment grant (jointly with engineering faculty) and (2) due to new research awards from other agencies, including DARPA. The principal equipment purchase made with the CISE equipment grant is a cluster of 16 SUN Ultrasparc machines networked via Myrinet, ATM, and 100 MBit Ethernet. This workstation cluster's utility is enhanced with additional equipment purchased via the other equipment grants, including a compute server (a 12 processor SGI PowerChallenge shared-memory multiprocessor), 18 Indy workstations as additional computational and visualization/graphics engines, and a Sun SMC 1000E video server. Moreover, the additional funding permitted us to network all of these machines via both point to point Ethernet and ATM, and with SCI interconnects linking the Ultrasparc machines to each other. The resulting rich computational and network fabric is enabling us to perform substantive research in the variety of topics addressed in our proposal, which include high performance applications, collaboration systems, efficient middleware for HPC and for multimedia applications and for collaboration systems, and high performance and multicast communication protocols.

1 Project Summary

The College of Computing at the Georgia Institute of Technology has purchased computing equipment for support of research in computer science and engineering. The equipment is being used in several primary research projects, and has also been accessible for use by other related research. Primary projects using this equipment address both applications and technologies supporting them:

- Interactive scientific programs – the on-line interaction with scientific simulations running on multiple, networked parallel machines is enabling scientists to gain more rapid insights into the physical problems being explored.

- Interactive simulations – end users aree able to run and view system simulations in conjunction with actual system operation, to diagnose problems or to understand operational system characteristics.
- Distributed collaboration – collaboration support systems offer scientific users seemingly 'close' interactions via shared 3D data spaces.
- Object-based middleware addresses the efficient and potentially on-line sharing of state on single parallel machines and across multiple networked workstations or supercomputers.
- Object-base middleware and applications are enhanced by experimentation with the underlying communication protocols required by the diverse types of state sharing being performed in both.

Specific progress on each of these topics is described below.

2 High Performance Applications

2.1 Interactive Scientific Computations

This NASA-funded project is developing a high performance, parallel atmospheric modeling code, in conjunction with researchers from the School of Earth and Atmospheric Sciences at Georgia Tech, and funded by a joint NASA grant. This code exhibits the following novel attributes:

- It is interactively executable, where researchers can 'steer' the ongoing computation by manipulation of selected parameters. Currently, such online steering may be performed with respect to the dynamics computations (i.e., constituent transport) being performed, specifically addressing alternative settings for vertical constituent motion.
- It has high computational and communication requirements, where the online monitoring and visualization of selected output data (and corresponding observational data) must be performed such that little or no performance penalties arise. This requires significant infrastructure support in terms of the transport and processing of monitoring, steering, and visualization information, especially when multiple machines are involved in these processes.
- Its steering interface utilizes data visualizations, rather than the typical code module or program visualizations employed by other efforts addressing program steering (or by

debuggers). As a result, end users may inspect and steer programs in terms with which they are familiar.

The project has been focussing on interactivity of the application with a moderate number of distributed end users, as well as with porting the application from a shared memory platform to mixed shared and distributed memory platforms:

- The transport (dynamics) component of this application has been parallelized on both shared memory and distributed memory machines. It currently runs on cluster machines (SGI or SUN Sparc machines), on the IBM SP-2, and on the SGI PowerChallenge. It has been run on larger-scale remote machines, as well, including machines at NCSA. It has been used in an I-Way demo by our group in SC'95.

- A more complex chemistry code is currently being added, entailing interfacing the existing C code with a Fortran-based chemistry package. The chemistry package performs complex ozone chemistries, which are required for more realistic global models of the atmosphere's behavior.

- The enhanced model is being ported to utilize both the Utrasparc cluster and its enhanced communication ablilities as well as the various shared and distributed memory supercomputers available in the College of Computing and elsewhere at Georgia Tech, which include the IBM SP-2 and an R10000-based SGI Powerchallenge.

- Extensive recent efforts on this project are addressing the on-line visualization and manipulation of the large-scale data sets being produced and utilized in scientific models like these. Specifically, interactive model interfaces now being constructed in standard environments (ie., OpenInventor and OpenGL and NOT using the previous Explorer interfaces utilized by this project) and are being extended to offer convenient constructs for 3D data visualization as well as for collaboration among multiple scientists working on remote machines. This data-based collaboration via 3D user interfaces will be described in more detail later in this report.

Publications and a PhD thesis arising from this work are listed below:

1. T. P. Kindler, "The Development of Supercomputing Tools In a Global Chemistry Transport Model (CTM) and Its Application to Selected Problems in Global Atmospheric Chemistry", Dec. 1995 (School of Earth and Atmospheric Sciences - William L. Chameides, thesis chair).

2. T. Kindler, K. Schwan, D. Silva, M. Trauner, and F. Alyea, "Parallelization of Spectral Models for Atmospheric Transport Processes", Concurrency: Practice and Experience, to appear 1996.

3. M. Trauner, F. Alyea, W. Ribarsky, G. Eisenhauer, V. Martin, "I-Way Demo: An Interactive Atmospheric Modeling Code", *Supercomputing '95*, San Diego, Dec. 1995.

4. Y. Jean, T. Kindler, W. Ribarsky, W. Gu, G. Eisenhauer, K. Schwan, and F. Alyea, "Case Study: An Integrated Approach for Steering, Visualization, and Analysis of Atmospheric Simulations", *Visualization '95*, Oct. 1995.

5. K. Schwan, F. Alyea, W. Ribarsky, and M. Trauner, "Integrating Program Steering, Visualization, and Analysis in Parallel Spectral Models of Atmospheric Transport", *Science Newsletter: Information Systems*, NASA, July 1995.

6. B. Schroeder, G. Eisenhauer, J. Heiner, V. Martin, K. Schwan, and J. Vetter, "From Interactive Applications to Distributed Laboratories", journal submission.

2.2 Interactive and Distributed Simulation

This project is developing high performance parallel and distributed simulation tools to support parallel and potentially interactive simulations of large-scale telecommunication networks. Over the past nine months, our parallel simulator (Georgia Tech Time Warp, or GTW) that was originally designed to execute on shared memory multiprocessor platforms has been adapted to execute on the equipment puchased with NSF funds. Experiments have been performed executing simulations of wireless networks on the PowerChallenge, Sun Ultrasparc workstations, and SGI Indy workstations.

Using the NSF equipment, experimental research completed over the first year of the grant focused on (1) partitioning ATM network simulations for parallel execution, (2) developing new dynamic load balancing algorithms to enable Time Warp simulations to execute "in background" on the NOW platform (a system that performs dynamic load management is now operational), (3) developing techniques to efficiently perform incremental state saving in Time Warp, transparent to the simulation application, and (4) developing a variety of techniques to enable efficient execution of Time Warp programs.

This research is already beginning to have a substantial impact outside the Georgia Tech community. Within the last year MITRE has developed a commercial air traffic simulation on top of GTW that allows them to perform simulations that previously took 1.5 *hours* using a SIMSCRIPT-based simulator in only one or two *minutes* using GTW on a multiprocessor Sparc, while producing comparable simulation results. MITRE is planning to provide this tool to the FAA for use in air traffic management. Our research has also heavily influenced the *High Level Architecture* effort in DoD in defining a common distributed simulation infrastructure for all simulations in the DoD.

Publications that describe this work in greater detail include:

1. F. Hao, K. Wilson, R. M. Fujimoto, and E. Zegura, "Logical Process Size in Parallel ATM Simulations," to appear in *1996 Winter Simulation Conference*, December 1996.

2. C. Carothers and R. M. Fujimoto, "Background Execution of Time Warp Programs," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

3. D. West and K. Panesar, "Automatic Incremental State Saving," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

4. S. R. Das and R. M. Fujimoto, "Adaptive Memory Management and Optimism Control in Time Warp," submitted for publication.

5. S. R. Das and R. M. Fujimoto, "An Empirical Evaluation of Performance-Memory Trade-offs in Time Warp," submitted for publication.

6. R. M. Fujimoto and M. Hybinette, "Computing Global Virtual Time in Shared Memory Multiprocessors," submitted for publication.

7. R. Rönngren, R. Ayani, R. M. Fujimoto, and S. Das, "Efficient Implementation of Event Sets in Time Warp," submitted for publication.

8. R. M. Fujimoto and R. M. Weatherly, "Time Management in the DoD High Level Architecture," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

9. R. M. Fujimoto and R. M. Weatherly, "HLA Time Management and DIS," *14th Workshop on Standards for the Interoperability of Distributed Simulation*, March 1996.

10. R. M. Fujimoto, "Parallel and Distributed Simulation," *1995 Winter Simulation Conference*, December 1995, pp. 118-125.

3 Technology Developments

Several research projects are developing underlying, enabling technologies to support high performance applications on NOW machines and on parallel supercomputers. Three are primarily concerned with "middle-ware" software and technologies that are directly utilized by the two applications exposed above. The first project is studying the dynamic monitoring, adaptation, and interactive steering of high performance computations for on-line control of "virtual laboratory instruments" and for "what-if?" experimentation with complex simulation models by distributed laboratory users. The second project is exploring the efficient execution of simulation programs, especially discrete-event simulations of telecommunication network models, on multi-granular compute servers (servers containing both tightly-coupled multiprocessors and loosely-coupled workstations). The third project is exploring the collaboration technologies required in distributed laboratories.

The fourth project is concerned with the distributed systems and telecommunication networking technologies underlying distributed laboratories. Components of this work are already being used to support the middleware efforts listed above. The emphasis in distributed systems research is on support for shared state in multi-granular parallel and distributed computing environments. Networking research is concerned with the communications in such environments, including multicast protocols and protocol configuration.

3.1 Interactive Steering

This project is exploring the ways in which supercomputer applications may be made interactive. It is well known that such interaction is desirable and perhaps, essential in order to increase the effectiveness of researchers working on highly complex problems. The project's topics are driven by the atmospheric modeling application described above:

- targetting on-line monitoring and steering, including the development of infrastructures for both.
- offering the ability to span many heterogeneous machines and deal with dynamic behaviors in terms of the number of application components being monitored and the amounts of data being extracted via monitoring and injected via steering.
- providing the ability to interact with multiple user interfaces, so that end users can employ data visualizations meaningful in their research domains, while developers can view and

manipulate target applications in terms of program modules/variables, view results in terms of performance, etc.

Our successes to date include the development of an innovative infrastructure for monitoring and steering parallel, shared memory programs. This infrastructure will be re-released to the broader research community July 1, 1996. We are working with NCSA researchers (Dan Reed) to evaluate its capabilities outside Georgia Tech. In addition, one of our students is spending the summer at Los Alamos National labs to apply the infrastructure to other application programs.

We are currently developing infrastructure that addresses the workstation clusters and multi-granular compute and visualization engines now used by most high performance researchers. Specifically, while on-line monitoring and steering already operates across multiple workstation and supercomputer systems, we are investigating issues of scalability by development of a useful infrastructure that can be configured dynamically to adjust to larger numbers of participants, to adapt to changes in traffic behavior, and to continue operation in the presence of bursty behavior. We are specifically addressing the potentially bursty behavior of supercomputer applications (both in terms of their data output and their monitoring and steering), the requirements of controlled or low latency in reacting to monitored program events, and the need to construct infrastructure that can provide high throughput across multiple network platforms and devices.

Publications in the past year include:

1. G. Eisenhauer, W. Gu, D. Silva, K. Schwan, J. Vetter, E. Kraemer, "Opportunities and Tools for Highly Interactive Parallel Computing", *Debugging and Performance Tuning for Parallel Computing Systems*, editors: A. H. Hayes, M. L. Simmons, J. S. Brown, and D. A. Reed, IEEE Press, May 1996.

2. G. Eisenhauer and K. Schwan, "Parallelization of a Molecular Dynamics Code", *Journal of Parallel and Distributed Computing*, May 1996.

3. J. Vetter and K. Schwan, "Progress: A Toolkit for Interactive Program Steering", *24th International Conference on Parallel Programming*, IEEE, Aug. 1995.

4. H. Bergmann, J. Vetter, K. Schwan, and D. Ku, "Development of a Parallel Spectral Element Method Code Using SPMD Constructs", *Parallel CFD '95*, June 1995.

5. J. Vetter and K. Schwan, "Models for Computational Steering", *Third International Conference on Configurable Distributed Systems*, IEEE, May 1996.

6. W. Gu, G. Eisenhauer, K. Schwan, and J. Vetter, "Falcon: On-line Monitoring and Steering of Large-Scale Parallel Programs", journal submission, Dec. 1994, revised Aug. 1995, second revision June 1996.

3.2 Distributed Simulation Technologies

This project is developing high performance parallel and distributed simulation tools to support the virtual telecommunication network application. Over the past nine months, our parallel simulator (Georgia Tech Time Warp, or GTW) that was originally designed to execute on shared memory multiprocessor platforms has been adapted to execute on the NSF equipment funds. Experiments have been performed executing simulations of wireless networks on the PowerChallenge, Sun Ultrasparc workstations, and SGI Indy workstations.

Using the NSF equipment, experimental research completed over the first year of the grant focused on (1) partitioning ATM network simulations for parallel execution, (2) developing new dynamic load balancing algorithms to enable Time Warp simulations to execute "in background" on the NOW platform (a system that performs dynamic load management is now operational), (3) developing techniques to efficiently perform incremental state saving in Time Warp, transparent to the simulation application, and (4) developing a variety of techniques to enable efficient execution of Time Warp programs.

This research is already beginning to have a substantial impact outside the Georgia Tech community. Within the last year MITRE has developed a commercial air traffic simulation on top of GTW that allows them to perform simulations that previously took 1.5 *hours* using a SIMSCRIPT-based simulator in only one or two *minutes* using GTW on a multiprocessor Sparc, while producing comparable simulation results. MITRE is planning to provide this tool to the FAA for use in air traffic management. Our research has also heavily influenced the *High Level Architecture* effort in DoD in defining a common distributed simulation infrastructure for all simulations in the DoD.

Publications that describe this work in greater detail include:

1. F. Hao, K. Wilson, R. M. Fujimoto, and E. Zegura, "Logical Process Size in Parallel ATM Simulations," to appear in *1996 Winter Simulation Conference*, December 1996.

2. C. Carothers and R. M. Fujimoto, "Background Execution of Time Warp Programs," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

3. D. West and K. Panesar, "Automatic Incremental State Saving," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

4. S. R. Das and R. M. Fujimoto, "Adaptive Memory Management and Optimism Control in Time Warp," submitted for publication.

5. S. R. Das and R. M. Fujimoto, "An Empirical Evaluation of Performance-Memory Trade-offs in Time Warp," submitted for publication.

6. R. M. Fujimoto and M. Hybinette, "Computing Global Virtual Time in Shared Memory Multiprocessors," submitted for publication.

7. R. Rönngren, R. Ayani, R. M. Fujimoto, and S. Das, "Efficient Implementation of Event Sets in Time Warp," submitted for publication.

8. R. M. Fujimoto and R. M. Weatherly, "Time Management in the DoD High Level Architecture," *1996 Workshop on Parallel and Distributed Simulation*, May 1996.

9. R. M. Fujimoto and R. M. Weatherly, "HLA Time Management and DIS," *14th Workshop on Standards for the Interoperability of Distributed Simulation*, March 1996.

10. R. M. Fujimoto, "Parallel and Distributed Simulation," *1995 Winter Simulation Conference*, December 1995, pp. 118-125.

3.3 Collaborative Systems Support

This project has concentrated on three areas. First, work has been done on generalized techniques for shared visualizations on the Open Inventor platform in conjunction with the Atmospheric Science application. This software is designed to provide a base capability to support shared views of visualization between distributed participants. Work on this aspect of the project has been done by a student jointly supervised by Karsten Schwan and Scott Hudson.

Second, work has been done on new algorithmic techniques to support multiple coupled views in a distributed setting. This work has been aimed at extending existing "constraint-based" techniques for user interface implementation, which have been very successful in the single user domain, to a distributed setting. Although constraint techniques provide very powerful, high-level abstractions for implementation of several aspects of user interfaces, until now the efficient update algorithms needed to support them have been limited to single user sequential systems. Work done this year has developed a new distributed algorithm for efficient (incremental and lazy) update of constraints distributed across a network. This result extends a prior optimal algorithm in the sequential domain and preserves its performance characteristics (and optimality when used sequentially). In addition, the new algorithm has been proven to provide an optimal level of concurrency (in keeping with preserving causal semantics of updates).

The final area of work in collaborative systems has been in systems to support awareness in distributed work groups. There are a large number of factors that make working in a distributed setting more awkward than working in a co-located group. Many of these reasons are informal and relate to awareness of one's colleagues. Awareness serves as a backdrop and catalyst upon which more explicit communications is built. This work has looked at several specific techniques for promoting awareness in distributed work groups. It has been particularly concerned with tradeoffs that occur between transmission of awareness information and privacy (for the sender), and between awareness and disruption (for the receiver).

Work next year will continue to concentrate on the same three areas: specific tools to be applied to collaboration in the atmospheric sciences application, use of constraint systems to support distributed user interfaces, and new techniques to support awareness in distributed work groups.

Additional details of this work are described in the following publications:

1. S. E. Hudson and I. Smith, "Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems," submitted to the ACM Conference on Computer Supported Cooperative Work, 1996.

2. K. Bharat, S. E. Hudson, "Supporting Distributed, Concurrent, One-Way Constraints in User Interface Applications," Proceedings of the ACM Symposium on User Interface Software, pp. 121-132, Nov. 1995.

3. S. E. Hudson, I. Smith, "Ultra-Lightweight Constraints," submitted to the ACM Symposium on User Interface Software, 1996.

4. S. E. Hudson, I. Smith, "Electronic Mail Previews Using Non-Speech Audio," CHI '96 Conference Companion, pp. 237-238, April 1996.

5. S. E. Hudson, I. Smith , "A Practical System for Compiling One-Way Constraints into C++ Objects," submitted to ACM Transactions on Computer Human Interaction.

3.4 State Sharing Infrastructure

This project concerns the support of both a shared address space and shared object abstractions, called the COBS project. Concerning shared address spaces,we will offer distributed shared memory abstractions to permit both scientific and collaborative applications to run across networked and parallel machines, coupled via diverse network links. Moreover, rather than having programmers perceive differences between shared page objects part of their address spaces and other shared objects, the COBS project will permit them to construct such objects within a uniform framework, across heterogeneous system components and with industry-standard methods of decomposing and coupling application components (ie., using the CORBA object standard). The high performance realizations of such objects will utilize a diversity of performance techniques, including the specialization of object representations and even of communication protocols used for object interaction.

A major success in our work this past year has been the award of a large DARPA project to fund the industrially relevant implementation of high performance object middleware. The resulting COBS project is being performed jointly with IBM TJ Watson Research Center and is likey to interact with ISV's, as well. Below, we provide brief description of the main results in these areas and list the publications that have resulted from this work.

COBS: Shared Memory and Configurable Objects.

This project's current results include:

1. Novel formulations of concurrent and distributed objects that address low-latency communications on high performance parallel machines, such as the IBM SP2 and shared memory platforms. The intent of this work is to support the communication and synchronization requirements of single, large-scale, parallel applications on the designated target platforms.

2. Object formulations so that programs running on high performance platforms can equally easily run on workstation clusters, such as machines connected by Myricom or ATM network boards. In effect, such distributed objects will effectively extend an application's address space across large-scale, distributed memory machines. The Indigo system developed in our earlier research will be one basis of this work, in conjunction with the configurable object model described in papers on the 'Kernel Toolkit'.

3. Objects range from 'memory' objects permitting programs to share unstructured, raw data stored in memory pages to typed and structured objects explicitly defined by application programs, so that page objects may be used when appropriate, but performance may be improved further by exploiting class information or even information about how object state is physically distributed (fragmented) across different machines, called 'Distributed Shared Abstractions'.

4. Object implementations take advantage of a variety of performance techniques, including replication and the dynamic adaptation of selected aspects of their implementations, so that performance-sensitive implementation attributes may be dynamically changed in response to runtime variations in application behavior or execution environments.

Shared memory objects.

The ' Indigo' software framework developed as part of our research permits end users to construct user-level shared abstractions on loosely coupled machines. This platform does not compete with existing DSM systems like Treadmarks,but it complements them by permitting developers to create both additional consistency protocols, like the causal consistency protocols developed by our group, and to create object-based shared abstractions that have consistency semantics suitable for specific application classes. The initial implementation of Indigo utilized PVM to span multiple computational platforms. A current revision of its implementation will use the COBS-provided infrastructure for distributed object systems to permit end users to construct memory objects with diverse consistency semantics along with other objects (fragmented, dynamically configured, etc.) utilized by their applications.

From shared memory to shared objects. The Indigo and COBS efforts jointly address recent hardware developments like the application of HPCC interconnects to networked computers and the construction of loosely coupled 'parallel machines' from sets of workstations linked with ATM network switches. In addition, parallel machines of moderate sizes are becoming ubiquitous, ranging from SMP nodes in distributed memory architectures to workstations using multple processors for increased throughput. The resulting availability of such heterogeneous architectures is causing users to demand that both shared and distributed memory machines can be programmed and used in a similar fashion.

Demands for increased programmability of distributed memory machines have given rise to research efforts resulting in Distributed Shared Memory (DSM) libraries layered on top of machines' networking and virtual memory systems. They have also resulted in the development of distributed object-oriented systems offering user programs network-wide access to shared services, like Fresco, as well as object-oriented concurrent programming layers on top of shared and distributed memory machines.

Our contribution to the emerging field of heterogeneous parallel programming is the integration of object- and memory-based parallel and distributed programming. Using the small set of calls offered by Indigo, developers can efficiently implement objects of any size or type, ranging from the 'page' objects required by DSM implementations, to distributed shared abstractions (DSA) in which object state and functionality may be distributed across multiple machines' memory units, to the dynamically configurable objects provided by COBS, to heavier weight implementations of remote objects like those of Fresco. This breadth of support also distinguishes our work from communication libraries like active messages that do not offer the specific functionality required by DSM implementations or like Tempest, which is primarily intended to experiment with alternative hardware-relevant caching strategies at the software level. Similary, in comparison with the Nexus library developed for concurrent C++ and with the lower RPC-like layers of distributed object systems like Spring, Chorus,, and distributed objects. Indigo attempts to combine both the functionality required by remote objects and by distributed shared memory. Our hope is that libraries like Indigo may provide one basis for the development of standard, lower level interfaces, perhaps even supported by operating systems or hardware, based on which next generation parallel programming models may be implemented.

Ongoing research includes evaluations with alternative communication protocols, the inclusion of additional consistency semantics, especially those addressing distributed multimedia and collaborative applications, and the use of online protocol configuration to improve the performance of Indigo applications that exhibit highly variable communication behavior.

1. R. West, K. Schwan, and M. Ahamad, "Exploiting Application-level Consistency Information to Improve the Performance of Distributed Multimedia Applications", in preparation.

2. P. Kohli, M. Ahamad and K. Schwan, "Indigo: User-level Support for Building Distributed Shared Abstractions", to appear in *Concurrency: Practice and Experience*, 1996.

3. Christian Clemencon, Bodhisattwa Mukherjee, and Karsten Schwan, "Distributed Shared Abstractions (DSA) on Multiprocessors", *IEEE Transactions on Software Engineering*, February 1996.

4. Prince Kohli, Mustaque Ahamad, and Karsten Schwan,"Indigo: User-level Support for Building Distributed Shared Abstractions" , *Fourth IEEE International Symposium on High-Performance Distributed Computing (HPDC-4)*, August 1995.

5. Ahmed Gheith, Bodhi Mukherjee, Dilma Silva, and Karsten Schwan, "KTK: Kernel Support for Configurable Objects and Invocations" , *Second International Workshop on Configurable Distributed Systems*, IEEE, ACM, March 1994.

COBS: Coarse-Grain Distributed Objects.

Coarse-grain distributed objects that enable interaction among application components executing on heterogeneous nodes are needed for meeting the sharing needs of distributed laboratories subsystems such as the collaboration toolkit. We have developed a distributed object system based on the CORBA-like Fresco toolkit. We have explored several new techniques to provide high performance for such object systems. These include the implementation of a light-weight object request broker as well as the investigation of a new service that allows clients to cache objects. Object caching can improve system performance because when there is locality in access of objects, latency and communication costs can be avoided when the object exists in the client cache. Caching coupled with replication of servers also provides high availability and scalability for such object systems. However, both replication and caching give rise to the problem of maintaining consistency among multiple copies of object state. We have explored several consistency levels and their scalable implementations. These multiple consistency level address the various sharing needs of applications in distributed laboratories. For example, causal consistency can be used when users interact asynchronously whereas strong consistency can be used for those objects that support synchronous collaboration among users. We have developed a prototype distributed object system and are using it in an application that supports collaboration among atmospheric scientists. The following publications provide details of the research results. In the coming year, we expect to produce a more robust version of the prototype and will investigate scalable sharing algorithms for systems that allows applications to simultaneously use objects with different consistency levels and when such levels can be changed dynamically based on application needs and/or resource availability in the system.

1. M. Ahamad, S. Bhola, R. Kordale and F. Torres, Scalable Information "Sharing in Large Scale Distributed Systems", *Proc. of ACM SIGOPS Workshop*, September 1996.

2. R. Kordale, M. Ahamad and M. Devarkonda, "Object Caching in a CORBA Compliant System", *Proc. of Second Conference on Object-oriented Technologies*, June 1996.

3. R. Kordale and M. Ahamad, "A Scalable Technique for Implementing Multiple Consistency Levels for Distributed Objects", *Proc. of International Conference on Distributed Computing*, May 1996.

4. M. Raynal, G. Thia-Kime and M. Ahamad, "From Serializable to Causal Transactions", *Proc. of Symposium on Principles of Distributed Computing* (abstract), May 1996.

5. F. Torres and M. Ahamad, "Compact Clocks: Scalable Implementations for Logical Clocks", submitted for publication, April 1996.

The high performance ORB being constructed 'underneath' the Fresco and Indigo work described here is funded by a grant from DARPA. The work performed by our IBM partners will concern collaboration system support in more loosely coupled environments.

3.5 Networking Research

Work in telecommunications support for distributed laboratories is proceeding along several fronts. First, multicast mechanisms are important for many distributed laboratory applications. Techniques and tools for characterizing the dynamic behavior of MBone sessions are being developed. We are addressing the problem of fairness in a feedback-controlled multicast video distribution scheme, and a scheme called destination set grouping (DSG) has been developed that significantly improves fairness at a small bandwidth cost. Dynamic multicast routing techniques are also being developed. In a highly dynamic multicast group (e.g., this often arises in distributed simulations), the identity and location of the group members changes with the lifetime of the communication. We have been developing routing algorithms to support these types of groups, and are developing mechanisms to improve performance.

Replication of services within a network leads to the problem of selecting the best server from those that are equivalent in some sense. Anycasting is a paradigm that allows applications to easily select and communicate with the best server, even when network conditions and availability of servers are changing. We have designed and partially implemented an anycasting service.

Third, our research is also addressing how network- or protocol-level overheads, performance constraints, and behavior might be made available to implementors. Toward this end, we are investigating the runtime configuration of communication protocols, first directly driven by application characteristics or network feedback and second, driven by exposing selected protocol and network information via the COBS object layer to application programs.

The applications used to evaluate protocol research include the scientific and multimedia applications described above and additional applications developed outside the scope of this research, including interactive distributed games, embedded real-time applications, and multimedia systems like Video on demand and interactive jukeboxes. Additional details of this work are described in the following publications:

1. Robin Kravets, Ken Calvert, and Karsten Schwan, "Dynamically Configurable Communication Protocols and Distributed Applications: Motivation and Experience", GIT-CC-96-16, conference submission.

2. R. J. Clark, M. H. Ammar, "Providing Scalable Web Service Using Multicast Communication," submitted to *Computer Networks and ISDN Systems.*

3. S. Bhatacharjee, M. H. Ammar, E. Zegura, V. Shah, Z. Fei, "Application-Layer Anycasting Service," submitted to *INFOCOM 97*, Kobe, Jpan.

4. M. Donahoo, E. Zegura, "Core Migration for Dynamic Multicast Routing," to appear in *Proceedings of IC3N 96.*

5. P. Schneck, E. Zegura, K. Schwan, "DRRM: Dynamic Resource Reservation Manager," to appear in *Proceedings of IC3N 96.*

6. Daniela Ivan and Karsten Schwan, "Improving Protocol Performance by Dynamic Control of Communication Resources", to appear in *Proceedings of ICECCS 96*, Oct. 1996.

7. K. Almeroth, M. H. Ammar, "On the Use of Multicast Delivery to Provide a Scalable and Interactive Video-on-Demand Service," *IEEE Journal on Selected Areas in Communications*, Vol. 14, Number 6, August 1996, pp. 1110-1122.

8. S. Y. Cheung, M. H. Ammar, X. Li, "On the Use of Destination Set Grouping to improve Fairness in Multicast Video Distribution," *Proceedings of IEEE INFOCOM '96 Conference*, San Francisco, California, April 1996.

9. K. Almeroth, M. Ammar, "Collection and Modeling of the Join/Leave Behavior of Multicast Group Members in the Mbone," in *Proceedings of the High Performance Distributed Computing Symposium (HPDC-4)*, August 1996, Syracuse, NY.

10. X. Li, M. Ammar, "Bandwidth Control for Replicated-Stream Multicast Video Distribution," in *Proceedings of the High Performance Distributed Computing Symposium (HPDC-4)*, August 1996, Syracuse, NY.

11. E. Zegura, K. Calvert, S. Bhattacharjee, "Tera-op networking: Local adaptation to congestion," *Gigabit Networking Workshop*, 1996.

4 Equipment Usage

The computing equipment is heavily utilized, and has become the "workhorse" for much of the 'systems' research in the College, also including some collaborators in other departments such as Computer Engineering. Fourteen faculty (including three women) and approximately 50 graduate students (including eight women and four minority students) have been utilizing the equipment for research projects, with priority access given to the five faculty members acting as PIs and their 30 students utilizing this equipment.