We continue to make progress in all areas of the MSI project.

Equipment purchases to date, summarized by project area, include:

Networking		
	Cisco routers (9) for QoS experiments	$234,320
	Intel workstations (24)	$58,969
	Sun workstations (2)	$8,986
	Intel-based portables (9)	$20,556
	Handheld devices (9)	$6,243
	Networking equipment	$15,661
Databases and Storage		
	Sun E450/A1000 database server	$54,950
	Intel-based storage server engine (3)	$19,188
	Dell-based multimedia proxy servers (4)	$25,892
	Intel-based workstations (2)	$19,752
	Intel-based portables (2)	$6,647
	Sun-based workstations (3)	$37,372
Image Capture		
	Sony high resolution camcorders	$7,110
	Futuretel MPEG-2 encoder	$5,520
Image Compression		
	Intel-based workstation	$1,466
	SGI high resolution display	$2,335
Applications		
	Laptop computers (3) for distance learning	$10,146
	Workstation/printer accessories	$5,607
	Videoconferencing stations (3) and streamer	$14,338
Total		$555,058

We have been extremely successful in obtaining research funds and grants from numerous sources. Noted among them is an award from the State of Indiana for supporting faculty and graduate students to conduct research and experimental development of a multimedia system for the telemedicine application. These funds are in direct support of this effort. The total funding from the State is over $1.7 million for two years.

In addition to our funding from Telcordia, we have also established research collaboration in the area of multimedia document delivery with Siemens Corporation. These efforts have been carried out under the leadership of Ahmed Elmagarmid (PI).

We have also received internal funding from the Center for Education and Research in Information Assurance and Security (CERIAS) to support one post doctoral fellow, a visiting scholar, and two graduate research assistants. The CERIAS center at Purdue University, which promotes information security research and education, was started by a multi-million dollar grant from the Eli Lilly Foundation. The leadership for this effort is

provided by Gene Spafford, its director, who is also a co-PI. Gene Spafford has also received separate NSF funding for his security research.

Our group has received two large grants, one from NCR/Wal-Mart and the other from Hewlett Packard (HP). The NCR grant is for a 1.6 terabyte parallel database engine valued at approximately $7 million. This system not only augments the other storage and compute servers we are purchasing through the NSF grant, but also includes a large sample dataset for warehousing and mining experiments. This parallel server has 80 processors and 400 disk drives. The grant from HP allows us to have 20 workstations and one server to use as desktops for the researchers in this project.

We have received a National Science Foundation ITR grant (joint with Ohio State University and the University of Illinois, who are subcontractors) for studying multiple time scale traffic control aspects of the network research component, a project led by Park (co-PI).

Park has been awarded a DARPA grant from the ATO FTN program for studying scalable DDoS protection solutions, a component of the security architecture of the MSI project. Park has also received a gift grant from the Xerox Foundation to study the relationship between quality of service, network security, and fault-tolerance.

We received funding from the Santa Fe Institute (SFI) and National Science Foundation to hold a joint SFI/NSF Workshop "The Internet as a Large-Scale Complex System" (chair: Kihong Park, co-chair: Walter Willinger), March 29-31, 2001. More information is available at http://discuss.santafe.edu/internetcs.

Sunil Prabhakar (senior personnel) has received an NSF Career award to study efficient I/O techniques for a broad range of database applications. The project leverages the equipment that is being acquired through the MSI project. The co-PIs and personnel include several NSF CAREER awardees and a Presidential Young Investigator awardee.

The remainder of this section highlights specific activities in each of the three major areas of the project: multimedia databases and storage management systems, networking, and security.

Multimedia Databases and Storage Management Systems

There are four major research activities that we are currently pursuing:

1. *Enabling content-based access to video data*. We are pursuing activities aimed at developing a comprehensive video database management system that supports content-based retrieval of large-scale video data as a critical component of multimedia document management system. We are developing techniques and tools for content-based retrieval, semantic browsing, and querying of video, including techniques for (1) partitioning video source into meaningful segments to support more effective video representation and indexing, (2) grouping the physical segments into semantically richer units, (3) detecting

and recognizing faces, and (4) developing metadata to support fast and effective search and browsing techniques for large-scale video. We have incorporated the MPEG-7 representation of multimedia content into our system, which has significantly affected our schema representation and query management. We are developing a query manager that parses and executes MPEG-7 queries. We are also developing an MPEG-7 wrapper to import pre-extracted features in the MPEG-7 format and export features from the database in MPEG-7 document format. To access/query video data and experiment/test system functionality, we have developed a query interface that supports different kinds of queries, such as query by example image (matching image features against extracted video frame features or aggregate shot features) and query by motion (pan, tilt, rotate, zoom, etc.). We are building a web client as a web service to web-enable the entire system.

2. *Adapting traditional databases to handle video data*. The large-scale data volume and continuous media properties of video have a significant impact on the design of video database management systems (VDBMS) in the areas of storage, buffering, query processing, indexing and schema representation. We are developing a prototype video database management system that is capable of handling raw video data as well as its associated metadata. The system uses SHORE, a system developed at the University of Wisconsin, as a storage manager, and PREDATOR, which is a value-added server for SHORE. We continue to address the issue of indexing video feature data to support similarity searching. Our approach is to develop dimension reduction techniques in combination with semantic clustering by exploiting several types of features and the semantics of the video data. We are enhancing the query processing capability of the system to support query by multiple examples and multiple features, and we are developing global similarity ranking techniques to rank results based on any number and combination of features. We have developed mechanisms for executing online queries on multiple video streams for the purposes of video editing, analysis and monitoring. In the query engine we are implementing novel join techniques for online sensor data in the query engine.

3. *Delivering Distributed Multimedia Documents with Support for End-to-End Quality-of-Service*. Currently, we are exploring different approaches that will allow mapping of the user-specified Quality of Presentation (QoP) parameters to Quality of Service (QoS) requirements for different system components of the overall MSI architecture, including storage, servers, networking, and security subsystems. The implementation of the translation mechanisms will be an integral part of the QoS-based resource scheduling modules that will be implemented using several dynamic and static approaches. We have designed an architecture for a real-time distributed multimedia database system (RTDMDS) for managing multimedia documents with Quality of Service (QoS) guarantees. This system is one of the key components of MSI. It allows distributed users to author, store, query, and retrieve multimedia documents over a broadband network. We have developed several modules of RTDMDS and continue the development of the entire system using the equipment acquired through the NSF grant. We are also investigating QoS-sensitive storage of video data.

We are currently developing an end-to-end resource management framework using proxy servers for a large-scale distributed multimedia database system. The system uses a suite of proxies to provide support for interactive function on multimedia presentation. A quality of service (QoS) routing technique has been developed to allow dynamic mapping of QoS connections to network resources, including end-to-end link capacities and buffering capabilities of the proxy servers. The primary objective is to minimize the cost of transmission while simultaneously balancing the load among proxies. Load balancing among proxies is critical because when a proxy becomes heavily loaded as a result of severe load imbalance, the proxy may not be able to meet its QoS commitments to clients. The proposed algorithm is based on a Lagrangian relaxation and a Rounding Breadth First Search algorithm. The method provides a near-optimal solutions in real-time. It is scalable to a large-scale, proxy-based networking infrastructure, and has computation time within the time scale of network state updates of existing network services.

4. *Developing techniques for the large-scale storage of video data*. For physical storage management of multimedia documents, we have designed several novel data placement and scheduling schemes. These schemes are currently being implemented on a Sun E450 server and a Sun A1000 Raid array acquired through this grant. Also, due to the unique temporal nature of video data, we are analyzing several techniques for real-time disk scheduling. Managing large volumes of data necessitates the use of cheap tertiary storage. Given the very high random access cost of tertiary storage, efficient management of data is critical for performance. We are therefore developing data placement, migration, prefetching, caching, and scheduling schemes for the effective retrieval of video from secondary and tertiary storage.

As part of this project, we are currently developing an XML-based multimedia application using the Oracle XML DB which is a high-performance, native XML storage and retrieval technology available with Oracle9*i* Database Release 2. This development framework complies fully with the W3C XML data model, and absorbs the XML document model into the database for navigation and query. In addition to the native XML repository to the database, the Oracle XML DB encompasses both SQL and XML in a highly interoperable manner. With this combination, the application can store both structured and unstructured data that can be accessed using either SQL or XML operations interchangeably.

Networking

1. We have set forth a new unified theory of differentiated services that is implementable on IP networks. It improves on our earlier work on QoS scheduling, facilitating a theory of aggregate-flow QoS control.

2. We have benchmarked the new architecture and theory using QSim, our ns-based WAN QoS simulator, confirming the theoretical predictions. Our work shows that differentiated services, following our architecture, can provide scalable and efficient user-specified services.

3. We have installed the Purdue Infobahn comprised of four Cisco 7206 VXR routers that form an IP-over-SONET QoS backbone. We have implemented initial signaling and end-to-end controls using both RSVP (Int-Serv) and AS (Diff-Serv) to ascertain the routers' operating capabilities.

4. We have performed LAN-scale testing and benchmarking over the IP-over-SONET IP routers differentiated service provisioning, guaranteed service provisioning, and multiple time scale traffic control. Twelve PCs have dedicated connections to the testbed and are engaged as QoS flows and background traffic generators.

5. The present QoS testbed has been renamed to Q-Bahn (QoS Infobahn) to highlight its broad, comprehensive nature with respect to providing a complete solution for scalable QoS provisioning in IP internets. Q-Bahn consists of 9 Cisco 7206 VXR routers connected as an IP-over-SONET backbone, with 50+ PCs and workstations directly connected to the backbone for benchmarking. The physical topology of Q-Bahn is the same as that of Internet2/Abilene, albeit completely dedicated for advanced QoS and security research, which Internet2/Abilene (due to its operational demands) is not. An even more distinctive feature of Q-Bahn is the operating system running inside the Cisco 7206 IP routers, called purdue-ios, which is a custom IOS (Cisco's router operating system) implementing the QoS switching algorithms developed in the Network Systems Lab. This unique collaborative effort between academia and industry led by Park (Directory of Network Systems Lab and co-PI of MSI project) allows advanced QoS research to be prototyped in a state-of-the-art production environment, a characteristic strength of the QoS component of the MSI project.

6. The Q-Bahn QoS testbed has been extended to incorporate a wireless mobile component, where a private six access point WLAN cellular network covering the three floors of the CS building is directly connected to the Q-Bahn testbed. The wireless mobile extension has been deemed necessary given the explosion, and expected dominance, of WLAN based local access technology, which injects additional complexities into provisioning effective end-to-end services that must be explicitly incorporated in the global system architecture. The wireless mobile access network is driven by a collection of in-house QoS-enabled applications developed in the Network Systems Lab, one of which, called QVI, a VoIP application incorporating QoS amplification mechanisms, runs over both Windows XP (PCs and laptops) and Windows CE (Compaq iPAQ Pocket PC). In addition to providing QoS-sensitive real-time VoIP and CD quality audio communication over integrated wired/wireless environments, the application QoS amplification infrastructure provides completely transparent QoS support for legacy applications via a QoS module, called Q-Driver, which resides in NDIS of the Windows XP and CE microkernels. Thus, our QoS platform is truly portable to environments characteristic of a typical end user.

7. We have generalized the unified differentiated services framework to a queuing framework (m-class G/G/1 queuing system), which will provide the most comprehensive

theory of aggregate-flow scheduling yet and a significant advancement of scheduling theory.

Security

1. We are focusing our efforts on building security-based access control mechanisms for video and multimedia databases. For video databases, we are currently developing an access control system on top of a video database system. At the higher levels, we are developing an access control model that specifies the users' credentials and qualifications as well as the content description of the underlying video. At the lower levels, we are building a toolbox for extracting desired features from the underlying video streams. For multimedia document systems, we are developing a security framework that allows the integration of heterogeneous access control policies in a distributed environment.

2. We have advanced and analyzed a denial-of-service (DoS) attack prevention framework based on probabilistic packet marking (PPM), and shown its effectiveness in the presence of single-source DoS attacks. We have used an adversarial framework to derive solutions to optimal decision making in a two-player environment comprised of the attacker and victim/target.

3. We have investigated the distributed DoS (DDos) attack problem under PPM and general network topologies where the attacker can pick the location and number of attack hosts. We have shown that uncertainty factor amplification --- a measure of the attacker's ability to hide his true location(s) --- is achievable by reducing the attack volume at each individual attack site.

4. Our new work on scalable DDoS attack protection based on route-based distributed packet filtering (DPF) is being recognized as one of the very few effective and deployable DDoS protection mechanisms. One of our future goals is to incorporate the packet filtering mechanisms in the switching elements of the Q-Bahn testbed, thus providing further integration of QoS and network security. Two avenues we seek to explore are extension of the collaborative work with Cisco to include network security and exploration of Intel's IXP 1200 network processor as a programmable IP router platform wherein both QoS and security mechanisms can be implemented, tested, and benchmarked as part of the Q-Bahn testbed.

5. We have investigated the network security architecture of Cisco's router operating system IOS with the aim of incorporating the adaptive security architecture AdSec into its structure.