

An Overview of the MAGIC Project

1. Introduction

An increasing number of applications utilize powerful computing resources that are distributed over local- and wide-area networks (LANs and WANs). In such an environment, the full potential of these resources can be realized only if the network speeds are sufficient to support the processing demands. Currently available WANs operate at megabit-per-second (Mbps) speeds and therefore can limit the overall performance of these distributed systems. Gigabit-per-second WANs promise the next major advance in computing and communications: powerful, geographically distributed computing resources with high-speed access to remote and time-critical data sources. Such networks will allow researchers to develop distributed, interactive applications with massive and real-time data requirements. Furthermore, they will allow data from multiple sources to be integrated for use by these applications, with neither the data sources nor the users colocated with the computing resources.

Many challenges must be addressed before the benefits of gigabit WANs can be achieved. Some of these include:

- Ensuring that heterogeneous computing and networking devices can interoperate
- Coordinating multiple data streams destined for a single location
- Accommodating bursty as well as steady traffic
- Compensating for the effects of network delays and errors on high-throughput data transmissions

The MAGIC (Multidimensional Applications and Gigabit Internetwork Consortium) project has been established to develop a very high-speed, wide-area networking testbed that will address these challenges and demonstrate real-time, interactive exchange of data at gigabit-per-second (Gbps) rates among multiple distributed servers and clients. Participants in the project include organizations from government, industry, and academia. The Advanced Research Projects Agency (ARPA) is providing funding for some of the MAGIC participants:

- Earth Resources Observation System Data Center, U.S. Geological Survey (EDC)
- Lawrence Berkeley Laboratory, U.S. Department of Energy (LBL)
- Minnesota Supercomputer Center, Inc. (MSCI)
- SRI International (SRI)
- University of Kansas (KU)

In addition, ARPA is funding the MITRE Corporation to help with project coordination.

Other MAGIC participants that are contributing significant equipment, facilities, and/or personnel to the effort include:

- Army High-Performance Computing Research Center (AHPCRC)
- Battle Command Battle Laboratory, U. S. Army Combined Arms Command (BCBL)
- Digital Equipment Corporation (DEC)

- Northern Telecom, Inc./Bell Northern Research (NTI)
- Southwestern Bell Telephone (SWBT)
- Splitrock Telecom (Splitrock)
- Sprint
- U S WEST Communications, Inc. (USW)

The objectives of the MAGIC research project are:

- To establish a standards-based, gigabit internetworking testbed that will support research and development for next generation networking technologies and applications
- To demonstrate the feasibility of utilizing a gigabit WAN to implement a defense-related application that requires massive amounts of remotely stored and processed data in real time
- To explore and resolve the technical challenges and issues associated with establishment and use of such a high-speed, distributed computing environment

2. Description of the MAGIC Testbed

The MAGIC testbed will consist of the following three components:

- An interactive, real-time, terrain visualization application
- A distributed image server system with performance sufficient to support the terrain visualization application
- A high-speed internetwork to link the computing resources required for real-time rendering of the terrain

The terrain visualization application, known as TerraVision, will allow a user to view and navigate through a representation of a landscape. Initially, the landscape will be created from aerial images of the U.S. Army National Training Center (NTC) in Fort Irwin, California. As exercises are conducted, the locations of vehicles will be superimposed on the view of the terrain and updated in real time. TerraVision is of direct interest to the U.S. Army since the ability of a commander to see the battlefield, and to share a common view of the battlefield with his command, is critical to effective command and control. TerraVision requires very large amounts of data in real time, transferred at both very bursty and high steady rates, and has network throughput as its major limiting factor. Steady traffic occurs when a user moves smoothly through the terrain, whereas bursty traffic occurs when the user jumps (“teleports”) to a new position.

The image server system (ISS) will store, organize, and retrieve both the processed images and elevation data required by TerraVision for interactive rendering of the battlefield, and the raw images and supporting information used to compute them. The processed images and elevation data comprise a multi-resolution hierarchy of orthographic (ortho) image- and digital elevation model (DEM)-tiles. The size of the area represented by an individual tile depends on the resolution of that tile: the higher the resolution, the smaller the area represented. The ISS will accept requests for tiles from the application and either return them directly, or send the raw images and supporting information to a supercomputer for processing into tiles and subsequent transmission to the application. The ISS will consist of multiple, coordinated data servers that are designed to be distributed around a wide-area network. This approach to system architecture will

permit location-independent access to databases, and allow for system scalability.

The MAGIC network will include four high-speed LANs interconnected by a gigabit WAN. The WAN will be based on evolving standards, namely Synchronous Optical Network (SONET) and Asynchronous Transfer Mode (ATM). ATM LANs will be installed at BCBL, EDC, and KU; an existing HIPPI (High-Performance Parallel Interface) LAN at MSCI will be interfaced to the WAN. The network will provide trunk speeds of 2.4 Gbps and access speeds of 622 Mbps, allowing an application to use a supercomputer at one location to process data from a database at a second location, and display the results, in real time on a graphics workstation at a third location.

For TerraVision, the supercomputer, a Thinking Machines CM-5, will be located at MSCI; the database system (*i.e.*, the multiple servers of the ISS) will initially consist of three Sun SPARCstations co-located at EDC; and the rendering workstation, a Silicon Graphics Onyx, will be located at BCBL. Real-time position updates of vehicles at the NTC will be obtained via the Global Positioning System. In addition, a workstation at KU will provide a remote “over-the-shoulder” view of the terrain and vehicle overlays as they are rendered. A geographical and functional overview of the MAGIC network is presented in Figure 1.

3. Project Phases

The research and development required to establish the MAGIC testbed, to demonstrate its capabilities, and to investigate and resolve the associated technical challenges will be accomplished in two phases. In Phase I, the high-speed LANs will be interconnected by point-to-point SONET links, and the interactive terrain visualization application will utilize ortho-image and DEM tiles that have been pre-computed and stored on the remote and distributed data servers of the ISS. The network will aggregate and transmit the multiple data streams to the application. In Phase II, the WAN will be converted to a switched ATM backbone by the addition of one or more ATM switches, and the testbed components will be enhanced to permit real-time, remote generation of the ortho-image and DEM tiles. Generation of the tiles and display of the terrain in real time will simulate the processing of satellite imagery as it is received. The major activities for each phase are as follows:

Phase I:

- Establish a gigabit WAN, based on customer premises ATM equipment and SONET point-to-point links, connecting BCBL, EDC, KU, and MSCI
- Implement high-speed ATM LANs at BCBL, EDC, and KU, connect them to the gigabit WAN, and demonstrate interoperability among the heterogeneous technologies used in these networks
- Pre-compute the tiled multi-resolution hierarchy of ortho-image and DEM tiles representing the NTC exercise area from aerial images and elevation data
- Design and develop an ISS to load, organize, and retrieve the tiles for interactive rendering
- Design and develop a terrain visualization application on an advanced graphics workstation that will allow interactive navigation of the NTC exercise area using pre-computed tiles
- Implement a system for real-time tracking of combat vehicles and for incorporating vehicle position data into the view of the terrain

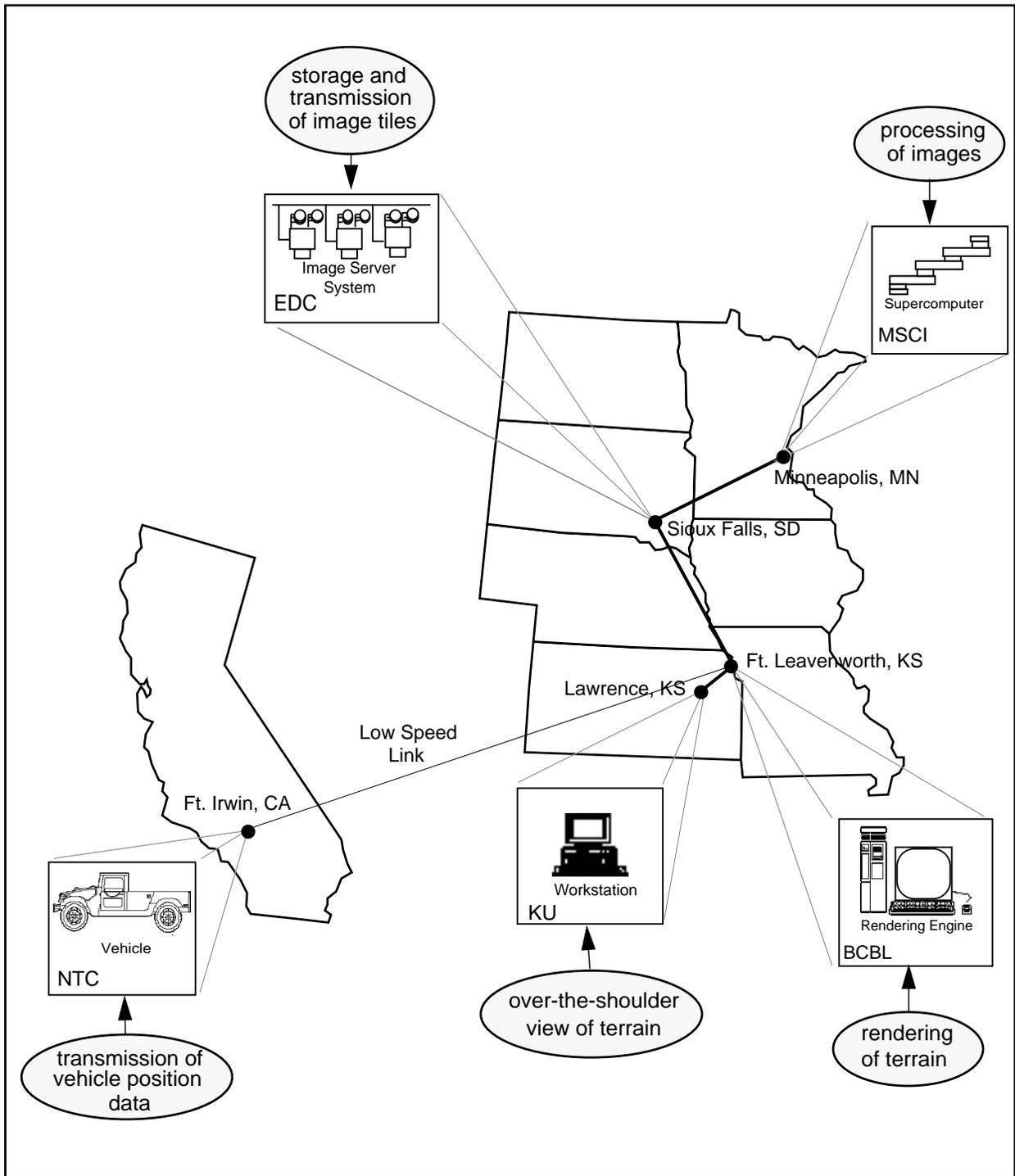


Figure 1. MAGIC network: geographic and functional overview.

- Provide an over-the-shoulder view of the rendered image on a workstation at a remote site

Phase II:

- Interface the MSCI HIPPI LAN to the MAGIC WAN
- Migrate the network to an ATM WAN through installation of one or more ATM switches in the backbone
- Upgrade the gateways and hosts of the LANs to use switched ATM connections
- Measure, characterize, and enhance the performance of the ATM internetwork
- Implement algorithms on a supercomputer to compute the ortho-image and DEM tiles on demand and in real time
- Improve the application's man-machine interface, and enhance the terrain rendering software to increase throughput
- Increase the number of ISS data servers, and enhance the data storage and file placement software as well as the software for coordinating the multiple, high-speed data streams from the individual data servers

4. Research Issues

Some of the research issues that will be addressed in the MAGIC project are:

Terrain visualization software. Development of two components of the terrain visualization software will present significant challenges: the path prediction and tile pre-fetching algorithms, and the terrain rendering software. The rendering engine must maintain a cache of the terrain imagery tiles that are currently visible on the screen as well as those that are likely to be visible in the near future. The size of the cache is a function of end-to-end latency. That is, the longer the time required for a requested tile to arrive at the rendering engine, the larger the cache required. However, by accurately predicting the user's path and viewpoint, the size of this cache can be reduced, and the number of unnecessary tile requests can be minimized, thereby reducing traffic across the network. The challenges in developing the rendering software include development of a fast algorithm for selecting appropriate resolution for each tile as a function of both its distance from the viewplane and its projected velocity in the viewplane, elimination of the boundaries between adjacent ortho-image tiles at different resolutions, and dealing with tiles that are required but that are not currently in memory.

Architecture for high-speed, scalable, parallel-distributed storage systems. In order to meet the throughput requirements of the application, data requests will be satisfied by several ISS disk servers operating in parallel. These disk servers, in turn, consist of semi-independent sub-systems (e.g., disks, disk controllers, memory, network interfaces) that can also operate with some degree of parallelism. In order to optimize server performance, those server sub-systems that are capable of operating concurrently must do so since the server will operate fastest when its sub-systems are running in a parallel, pipelined mode. Issues related to parallelism in these sub-systems, and to the impact of geographic distribution of servers will be addressed. Furthermore, in order to take advantage of the parallelism of the components, the data must be distributed among the servers, and among the disks of the servers, so that the tile requests made by the terrain visualization application are satisfied with data from a large number of disks and servers. The spatial data management algorithms required for the distributed placement of data will be developed for the

general case of multiple users and multiple areas of terrain.

Testing and monitoring ATM networks. Data flow and time monitoring mechanisms are needed to facilitate an understanding of how the terrain visualization application, the ISS, and the network interact under both steady and bursty traffic. Monitoring facilities will be designed so that the components of the ISS, the network, and the application can be characterized. All of the ISS modules involved in accepting data requests, sending the data, and coordinating the data will be instrumented to generate and collect precise time and data volume statistics. These monitoring facilities will also be incorporated into the application, and into any application-like test programs. The nodes of the MAGIC testbed will be equipped with precision time sources to synchronize the servers and clients so that they can provide the time stamping needed for high-speed network characterization.

Development of a gigabit LAN/WAN gateway with dynamic bandwidth allocation. A gateway that connects a DEC AN2 experimental gigabit LAN to the WAN will be developed. The gateway will provide (1) high-capacity and reliable network services to hosts on the LAN, and (2) advanced functions for managing bandwidth on a high-speed LAN/WAN. The goal for the advanced gateway functions is to match the allocation of WAN resources (*i.e.*, capacity) to the requirements of the data sources and destinations for these resources. To achieve this goal, techniques will be developed to estimate data-source requirements for WAN bandwidth, and to signal the WAN for the need to change its resource allocation accordingly. When the wide-area ATM switch(es) is installed, LAN/ATM switch control, supervisory, management, and signalling algorithms will be developed, and the gateway development work described above will be extended to allow the gateway to work with the switch. To evaluate performance, the gateway will be able to generate/monitor per-virtual-circuit ATM traffic in real time.

Enhancement of network performance. Traffic measurements from the MAGIC internetwork will be obtained and will form the basis for the development of characterizations of bursty gigabit network traffic. Network performance measurements will also be obtained and will be correlated with the traffic measurements. These measurements will be used to identify performance deficits and will be the basis for developing performance enhancements. In addition, traffic measurements will be used to characterize the data sources, and performance measurements will be correlated with source characteristics.

The network will be stressed by a combination of traffic from the terrain visualization application, and from a cosmology simulation application that is being developed independently at KU, and that also uses a supercomputer at MSCI. Simultaneous operation of these two applications will cause large volumes of bursty traffic to impinge on the network.

High-speed transport protocols for imagery applications. Timer-driven UDP transport protocols will be investigated for supporting a steady flow request from the ISS control, and for allowing application control over re-transmission. It will also be necessary to include a mechanism for the network to put some back pressure on the servers to cause them to throttle back when the network becomes congested.

The work on these and other MAGIC research and development issues is expected to demonstrate a number of innovations, including the following:

- Implementing alternative architectures for connecting LANs to an ATM WAN
- Interoperating diverse, high-speed LANs and heterogeneous ATM devices across a broadband ATM WAN

- Allocating network resources through approaches that deal with bursty as well as steady communications at gigabit speeds
- Storing and retrieving data at aggregate gigabit speeds from distributed mass storage devices
- Using a high-speed network in combination with supercomputers and advanced graphics workstations to provide real-time, multi-source imagery modeling and visualization
- Using a multi-resolution hierarchy of tiled data from a multi-gigabyte database to provide both broad and close-up views of terrain

5. Roles and Responsibilities

The MAGIC testbed represents a collaborative effort among members of the academic, government, and industrial research communities to enhance the state-of-the-art of high performance, high-speed networking. The specific roles and responsibilities of the organizations participating in the MAGIC research project are presented below.

- AHPCRC: Provide access to a CM-5 supercomputer, and associated technical support
- BCBL: Provide facilities for the terrain visualization rendering engine
Provide guidance for the development of the user interface to the application
- DEC: Provide an AN2 LAN to KU, and provide technical support for development of a MAGIC WAN to AN2 LAN gateway
- EDC: Prepare the artifact-free raw images and elevation data and provide all supporting data required to compute the multi-resolution hierarchy of ortho-image and DEM tiles
Prepare the ortho-images and the multi-resolution hierarchy of ortho-image and DEM tiles
Host and administer the ISS
- KU: Develop a MAGIC WAN to DEC AN2 LAN gateway, and its associated resource allocation and dynamic control mechanisms
Characterize the network through network performance and traffic measurements
Demonstrate the feasibility of multi-location viewing of the screens generated by a user of the terrain visualization application
- LBL: Design, implement, and test the performance of the ISS
Co-develop a tile service protocol for communication between the ISS and the terrain visualization application (with SRI)
Develop data flow and time monitoring mechanisms for testbed components
- MITRE: Support ARPA in project coordination and oversight
- MSCI: Develop the MAGIC network architecture
Select and acquire the hardware and software required to implement the LANs at EDC and MSCI, and to connect these LANs and the rendering engine at BCBL to the WAN
Perform pre-deployment testing and integration of the components of the LANs
Deploy the equipment, perform point-to-point testing and integration to ensure end-to-end IP connectivity, and evaluate and enhance protocol performance

- NTI: Provide broadband communications equipment and technical support for backbone integration and operation
- Splitrock: Provide the fiber link between EDC and Sioux Falls, SD
- Sprint: Design, install, and service the SONET backbone for the MAGIC gigabit WAN
 Coordinate backbone integration with local exchange carriers and equipment suppliers
 Upgrade the SONET backbone with one or more ATM switches
 Provide advice and assistance in development, deployment, and integration of MAGIC gigabit LAN/WAN technologies
- SRI: Develop and/or provide the software for preparing the artifact-free DEM, ortho-images, and the multi-resolution hierarchy of ortho-image and DEM tiles required by the terrain visualization application
 Design and develop the interactive terrain visualization application, including the capability for real-time display of moving vehicles
 Co-develop a tile service protocol for communication between the ISS and the terrain visualization application (with LBL)
 Provide the systems engineering and integration support required at BCBL to implement and demonstrate the terrain visualization application
- SWBT: Provide facilities and technical support for backbone integration and operation
- USW: Provide facilities and technical support for backbone integration and operation, including interoperability testing of SONET equipment
 Test operation and performance of application traffic through an ATM switch