1 Executive Summary

The rapid growth in the generation of digital data is changing computational science in a fundamental way. Traditionally, the scope of computational problems was limited by the available processing power. But today, many problems are extremely data-intensive, and the lack of large-scale storage infrastructure creates a new bottleneck. Thus, modern data-intensive applications need high-performance computational resources *and* a system in which the computational resources are *tightly coupled* with large-scale storage. Currently, three major research projects at Cornell are severely constrained by the lack of such an infrastructure.

In the first project, researchers in the Cornell Program of Computer Graphics are acquiring a light measurement system called the Spherical Gantry. The Spherical Gantry will capture detailed information about the light-scattering properties of objects, which can later be analyzed to create accurate models for graphical rendering. The Spherical Gantry is expected to generate 50 Terabytes of data *for each object* within a span of a few days. In the second project, Cornell researchers in the Department of Computer Science are investigating the evolution of the World Wide Web using snapshots from the Internet Archive (http://www.archive.org) which is growing at about 10TB per month. In the third project, Cornell researchers in the Department of Computer Science and the Department of Astronomy are collaborating on the design and implementation of a data management and analysis infrastructure for the Cornell Arecibo Telescope in Puerto Rico. The anticipated total amount of data generated by just one of the Arecibo celestial surveys is more than one 800 Terabytes!

With funding from this grant, we propose to create a system that tightly couples high-performance computing with a large-scale storage infrastructure. Since online disk storage is relatively expensive (compared to near-line or offline tapes, using current cost estimates), the disk storage infrastructure will be shared among the three projects. This is a suitable arrangement because each project has a temporary need for large amounts of online storage when processing raw data, but has reduced (albeit still large) storage requirements when operating on processed data. Each project will also have access to the required amount of dedicated tape storage.

We will make the storage and computational capabilities available as ubiquitously accessible *web services* over a high-speed, high-bandwidth network so that researchers both inside and outside Cornell can effectively access the stored data and associated data processing services. Such service-oriented interfaces will enable data access from any platform through standardized protocols, and will open up the facility to other scientific research groups world-wide.

1.1 Research Impact

If this proposal is funded, we expect to achieve fundamental results in three different areas of study:

- **Computer Graphics**
  - We will perform accurate measurement of the light scattering properties of objects using the Spherical Gantry. This will be the first study ever undertaken at this level of accuracy, and it is expected to lead to new insights and new models for light reflection.
  - We will develop a new generation of multidimensional, multi-object, rendering algorithms that exploit the scattering data measured using the Spherical Gantry.

- **The World Wide Web**
  - We will design algorithms for measuring, detecting and quantifying the evolution of the Web. In contrast to prior Web research, which has mostly considered the Web to be a static entity, our research will develop new models for how the Web evolves with time.
  - We will develop a framework for formally reasoning about the process of collecting Web measurements. Currently, the sheer scale of the Web implies that most reported studies on the Web are performed exactly once. By using multiple snapshots of the Web generated by independent crawls, we will develop precise models that distinguish measurement-independent properties from those that are influenced by the method of measurement.
- **Astronomy**
  - We will discover a large number of new pulsars, including pulsars with millisecond spin periods and those in binary orbits with other neutron stars and perhaps black holes. This is expected to provide numerous opportunities for subsequent research on the equation of state of nuclear matter, gravitation physics, gravitational waves, stellar evolution, relativistic plasma physics, and the magnetized, ionized gas in the Milky Way.
  - We will develop new high-performance data analysis algorithms that will enable large-scale data mining of datasets from the astronomy research community.

The proposed projects bring together researchers from astronomy, computer science, information science, and the Cornell Program of Computer Graphics. Funding of this proposal will seed a storage infrastructure that will be used and extended by researchers throughout Cornell. We propose to host the system at the Cornell Theory Center, thereby leveraging their expertise in managing high-performance computing systems.

1.2 Broader Impact

The funding of this RI Proposal will have a much broader impact than the sum of the research that it enables. First, there would be the impact of increased collaboration within Cornell. The Arecibo Data Management and Analysis Project brings together researchers from computer science and astronomy, and we believe that this project is only the first of many similar projects possible at Cornell. Service-oriented interfaces and easily accessible web interfaces to data management and analysis tools can revolutionize how scientists interact with data and data processing capabilities. A successful Arecibo prototype system could lead the way to show their benefits to the Astronomy community.

Second, the participating research groups have a strong track record of involving undergraduate and female students into their research. We have an ongoing exchange program with Smith College where juniors and seniors from Smith come to Cornell for a summer internship. We also have a large group of undergraduate researchers working in our groups, including several Cornell Presidential Research Scholars.

Third, all the data that will be collected by the projects funded under this proposal will be available to the world-wide research and teaching community through our proposed Web service infrastructure. For example, the data will be a valuable resource for courses on machine learning and data mining to experiment with new data analysis and learning algorithms, by the astronomy community for further data analysis, and by the computer graphics community for the development of new rendering algorithms.

1.3 About the Research Team

The PI and director of the proposal is Alan Demers, Cornell Professor of Computer Science. Alan Demers was an architect at Oracle Corporation and has extensive real-world experience with the design and implementation of Terabyte database applications. The co-PIs and senior personnel are an interdisciplinary team consisting of Cornell professors from the Astronomy and Computer Science Departments, and the Program for Computer Graphics. The investigators have won many awards for work in their respective areas of research.

The researchers participating in this proposal have an established track-record that shows that they can identify and solve critical problems in the proposed research areas.

- **Computer Graphics:** Steve Marschner (Assistant Professor of Computer Science and Program for Computer Graphics) is one of the leading experts in graphical measurements. He has made many influential contributions to light scattering models, including work on reflectance measurement and on translucent materials that has been widely implemented by others in the academia and the industry. Kavita Bala (Assistant Professor of Computer Science and Program for Computer Graphics) is an expert on graphical rendering.

- **The World Wide Web:** Jon Kleinberg (Associate Professor of Computer Science) has developed some of the seminal algorithms for searching the web, modeling connectivity in networks, and for detecting bursts in data streams. William Arms (Professor of Computer Science and Director of Information Science) and Daniel Huttenlocher (Professor of Computer Science and School of Management) have extensive experience in Web data management and analysis. Jayavel Shanmugasundaram (Assistant Professor of Computer Science) is an expert on Internet querying.

- **Large-Scale Data Analysis for Astronomy:** Jim Cordes (Professor of Astronomy) is the chair of the Arecibo Pulsar-ALFA consortium. His work on the discovery of pulsars was featured in PBS's Scientific American Frontiers in 1998 with Alan Alda. Johannes Gehrke (Assistant Professor of Computer Science) has developed some of the fastest scalable data mining algorithms published in the research literature. His data mining algorithms are available as open source at http://himalaya-tools.sourceforge.net. Alan Demers (Professor of Computer Science) has real-world experience in dealing with large data sets, and Jayavel Shanmugasundaram (Assistant Professor of Computer Science) has done some of the seminal work on large-scale data publishing.

2 Research Infrastructure Description

All three projects described in this proposal share a common requirement for high-performance computing tightly coupled with large-scale storage. With support from this grant, we can provide a facility with sufficient computing capacity and storage to meet the peak requirements of all of the projects. We now provide an overview of how we propose to develop this infrastructure, and a timeline for acquiring the desired equipment.

2.1 Overview of the Facility

Over the life of the grant, we will develop a facility consisting of multiple high-performance clusters, tightly coupled to at least 650 TB of fast online disk storage. There will be a comparable amount of lower-cost (based on current prices) near-line robotic tape storage as well. Some near-line tape storage is acceptable for our projects because some fraction of the data produced, such as the raw data, is only periodically processed and does not have to be online. The total storage capacity of the facility will be a little more than one petabyte.

2.1.1 Computation

In its first year, the grant will provide a pair of 16-processor Itanium 2 systems, each with 50 TB of tightly coupled disk storage and a total of 64 GB of memory. This will be sufficient for the initial requirements of the research projects. Over the period of the grant, as the projects' computational needs grow, additional, more powerful clusters will be added, increasing the total computing capacity by more than a factor of five.

2.1.2 Storage

Initially, the facility storage will comprise the 100 TB of disk mentioned above. Over the period of the grant, the available storage will be increased by the addition of online disk and near-line tape storage. These plans are based on an assumption that the prices of disk and tape storage capacity will continue to decline at comparable rates. However, there is some reason to believe that the price of disk storage may fall more rapidly than the price of tape over the next few years. In this case, we would shift towards a higher proportion of disk storage, or possibly even stop acquiring tape altogether. However, based on the current prices, tapes seem to have a reasonable role as part of the overall storage plan. By the final year of the grant, total storage will be more than a petabyte.

2.1.3 End-User Platforms

Currently the PIs and senior personnel are well provided with end-user computers. Kavita Bala and Steve Marshner have sufficient workstations available through the Cornell Program of Computer Graphics. Jon Kleinberg and Daniel Huttenlocher have client workstations provided by individual grants, and William Arms has workstations available through the Cornell Digital Libraries Group. The Cornell Database Group (Alan Demers, Johannes Gehrke, and Jayavel Shanmugasundaram) and the Department of Astronomy (Jim Cordes) also have a sufficient number of end-user workstations available. Upgrades to these equipment will be provided by individual research grants.

2.2 Timeline

Year	Item	Quantity
Year 1	Unisys ES7000 Model 420 with 50TB ATA RAID disk	2
	Development Server: Dell Poweredge 2600 Dual 3GHz Xeon with 8 250GB Serial ATA 7200RPM drives	1
	Gigabit fiber networking cards	4
Year 2	14.7 Terabyte Node: Dell Poweredge 2600 with 48 350GB Serial ATA 7200RPM drives	2
	50TB disk expansion for Unisys ES7000 cluster	1
	LTO-3 drive upgrades for ADIC Scalar 10K	6
	100TB LTO-3 tape expansion for ADIC Scalar 10K	1
	Gigabit networking hardware	
Year 3	100TB expansion to Unisys ES7000 cluster	1
	LTO-3 drive upgrades for ADIC Scalar 10K	6
	150TB LTO-3 tape expansion for ADIC Scalar 10K	
Year 4	100TB Compute/Storage Nodes	2
	150TB tape library expansion	1
	Networking upgrades	
Year 5	100TB Compute/Storage Nodes	2
	Networking upgrades	

The above table shows our proposed timeline for acquiring the equipment. We discuss this timeline in more detail below.

2.2.1 Year 1

Based in part on a major equipment donation from Unisys and Intel that we will receive if this grant is funded, the infrastructure requested for the first year will include 100TB of online disk together with a total of 32 Intel Itanium 2 processor nodes with 4GB of memory per node. It specifically consists of two Unisys ES7000 Model 420 systems. Each contains 16 Intel Itanium 2 processors, 64GB of memory, and 50 TB of online disk storage. This initial 100TB of storage, coupled with very high-performance computation, will provide sufficient capacity for the initial research on physically accurate imagery, Web evolution, and data mining in astronomical surveys.

Selecting the appropriate online disk storage with the right price-performance ratio is a challenging task. At the low end, systems with IDE disk and no RAID can currently be configured for roughly $1000/terabyte for the bare disk drives, plus another $800 for the computer system to support them, for a total of about $1800/terabyte for a system with four 250GB disks or six 160GB disks. For software RAID, the cost would go up by another $200 for a total cost of $2000/terabyte. With hardware RAID, there would be another increment of $200-$300 for the RAID card. However, this system provides relatively low throughput and relatively limited local RAM and processing power. For our proposed applications, we need higher data rates, and a machine with greater I/O capability and much larger RAM (such as the Unisys ES7000) is needed. With hardware RAID, disk systems that can connect significant amounts of information (tens of terabytes or more) to large memory/processor computing systems currently run around $3000/terabyte. This is the configuration that we have chosen for Year 1, and expanded on in future years.

We have also requested a development server, with 2TB of storage, to support development and test of production data-intensive applications. This allows the research groups to avoid tying up the very high-performance production storage/analysis facility with development and test activities. To support the required gigabit/second connections between the Program of Computer Graphics (location of the Spherical Gantry), the Computer Science Department (location of the development server) and the storage/computation nodes (located in the Cornell Theory Center), we have also requested gigabit fiber hardware as part of the initial year request.

2.2.2 Year 2

During the second year of grant, we plan to expand the facility in three dimensions. We will add another 50TB of online disk storage to the ES7000 systems. This will continue to grow the resources needed to support the largest computation/storage efforts, such as the astronomy data-mining effort. We will also add two 15TB compute/storage nodes. These will directly support work on Web feature sets, as well as other data-intensive projects that are not large enough to require the main ES7000 systems. Finally, we will add 100TB of LTO-3 robotic tape storage. This will make use of the Cornell Theory Center's existing ADIC Scalar 10K robotic tape library system, allowing us to add significant storage at relatively low cost. The tape library will allow near-line storage of Web snapshot information and Spherical Gantry scans that are not being actively processed.

We plan to acquire tape storage because tapes are currently less expensive than disk. Specifically, 200GB LTO-2 tapes are available at roughly $90, holding 200GB uncompressed, or an estimated 400GB compressed. At the uncompressed end, the cost of tape is $450/terabyte, compared to the $1000/terabyte of bare hard drives. Tapes are also much easier to manage, store, and access than are offline disk drives, and do not have to be actively monitored like online disks.

It is possible that over the course of the grant, the cost per terabyte of disk will drop much more dramatically than the cost per terabyte of tape. In this case we will certainly revisit the need for tape at all. However, based on the current prices and the research application needs, tape seems to have a reasonable role as part of the overall storage plan.

2.2.3 Year 3

In the third year, we plan to acquire major additional storage capacity. The Unisys ES7000 cluster will be augmented with another 100TB of online disk storage, and the robotic near-line tape library, with another 150TB of tape storage. These additions will support continuing growth in the data being generated by all the projects, supporting larger online analyses and greater near-line storage of partially processed data sets. Again, these plans are based on an assumption that the prices of disk and tape storage capacity will continue to decline at comparable rates, and will be modified based on actual price ratios.

2.2.4 Year 4

By the fourth year of the grant, we expect that the Unisys ES7000 systems will no longer be providing leading edge computational power. Assuming continued growth in processing power and decline in storage costs, we are projecting being able to acquire two new 100TB compute/storage nodes with clustered/multiprocessor systems capable of significantly outperforming the initial Unisys systems. We are estimating the cost of these systems at $200,000/system and will configure them to maximize storage and processing power based on what is available at that time. We are also budgeting a further 150TB expansion in the tape library to increase our near-line storage capacity, subject to the caveats mentioned above. To support networking expansion we have included networking upgrades to add multiple gigabit links or new 10 gigabit networking if it is available.

2.2.5 Year 5

Finally, in year five, we propose acquiring two more compute/storage nodes to continue increasing the range of data sets that we can process, together with further networking upgrades. As in year four, we will maximize the storage/computation of these systems based on available hardware.

2.3 Broader Impact

The funding of this RI Proposal will have a much broader impact than the sum of the research that it enables. First, the research will have impact within Cornell. In the Arecibo Data Management and Analysis Project, researchers from the Department of Computer Science and the Department of Astronomy are working together on the analysis of the data from the Arecibo Observatory. This collaboration is exactly in the spirit of Cornell's unique vision of "Computing and Information Science" (CIS), where Cornell has recognized that computing and information science has emerged as a key enabling discipline vital to nearly all of its scholarly and scientific pursuits. Cornell has recognized the need for new programs of study in this field that span the campus. The Arecibo Data Management and Analysis Project has the potential to become a showcase of such collaboration, and it could give a significant boost to the still-nascent CIS structure. We believe that this project is only the first of many similar projects possible at Cornell, and that results from this work will demonstrate the value of such interdisciplinary interactions to the Cornell community at large.

Second, the research will have impact outside of Cornell. All the data that will be accumulated by the projects that this proposal would fund will be available to the world-wide research community through our proposed Web service infrastructure. For example, the data from the Arecibo Observatory will be made available through a data analysis infrastructure (running on the resources funded by this proposal) to the Astronomy community at large. With this new infrastructure, we hope to change the way the Astronomy community thinks about data interaction — from writing Perl scripts to select relevant subsets of the data to writing declarative queries for data retrieval (or to specify relevant data subsets through a graphicsal user interface). We urge the reviewers to read the support letter from Dr. Robert Brown, the Director of the National Astronomy and Ionosphere Center at the Arecibo Telescope, who expresses his enthusiasm about the proposed work. Similarly, funding of this proposal would allow us to make the datasets created by the Spherical Gantry available to the computer graphics community; we urge the reviewers to read the support letter from Dr. Pat Hanrahan from Stanford University who expresses enthusiasm about the availability of these datasets.

Third, the participating research groups have a strong track record of involving undergraduate and female students into their research. We have an ongoing exchange program with Smith College where juniors and seniors from Smith come to Cornell for a summer internship. For example, during the summer of 2003, a junior from Smith College, Melody Donoso, worked as an intern in the database group with PIs Johannes Gehrke and Jayavel Shanmugasundaram where she gained valuable experience into the workings of a database systems research group. We also have a large group of undergraduate researchers working in our groups. As another example of the attraction of our research, there are currently four undergraduate Cornell Presidential Research Scholars (CPRSs) working with the personnel in this proposal. The CPRS program offers outstanding Cornell students the opportunity to work with a faculty mentor; together they design and plan an individualized research program that can draw on funding from a small research account that the student can use. For example, during the fall of 2002, CPRS student Jay Ayres attended the 2002 ACM SIGKDD International Conference on Knowledge Discovery and Data mining to present a paper on a project that he worked on as a sophomore! We believe that the proposed research would further expand our ability to excite smart undergraduate students into the field of computer science — students who will be able to work hands-on in the exciting research projects that would be enabled with funding from this grant.

3 Resource Allocation

3.1 Current equipment

The following list includes all computing equipment in the Computer Science Department owned either by Cornell or by the Federal government.

- **Desktop Systems:** 114 Intel Celeron-based laptop PCs, 80 Intel 1.5GHz Celeron-based laptop PCs, 280 Intel Pentium III desktop PCs, 97 Intel Pentium III laptop PC, 345 Intel Pentium 4 desktop PCs, 18 Intel Pentium 4 laptop PCs, 47 Compaq 1GHz Transmeta Tablet PCs, 2 Sun UltraSparc 5, 14 Sun UltraSparc 10, 2 SunBlade 100, 19 SunRay 1, 4 Apple Power Mac G4 desktop, 2 Apple Mac iBook, 1 Apple iMac, 13 Mac Powerbook G4

- **Back-End Resources.** 15 TB server-based online storage, 9 Sun Ultra Enterprise 420/450 quad-processor servers, 2 SUN Ultra Enterprise 250 dual-processor servers, 8 SUN Sunfire 280R quad-processor servers, 1 3.8TB SUN tape library, 1 12-node AMD 1.2GHzAthlon Beowulf cluster, 1 45-node Linux PIII dual-processor cluster, 166 Dell dual-processor Pentium III/Xeon servers, 5 Dell quad-processor Pentium III servers, 4 Dell 8-way Pentium III servers, 40 Dell Xeon dual-processor servers.

The following resources at the Cornell Theory Center are also available to the Department of Computer Science for specific research projects.

- **Computational Biology Service Unit Clusters:** There are two clusters available for computational biology research. The first cluster has 192 nodes with 2.4GHz P4 dual processors, 2 GB RAM/Node, 72GB Disk/Node, 128HP Gigabit, and 512KB Cache/Processor (SMP). The second cluster has 64 Nodes with Pentium III 1GHz dual processors, 2 GB Ram/Node, 54 GB Disk/Node, and 256 KB Cache/Processor (SMP).

- **Virtual Reality Studies:** CAVE is a 3-wall 3-D immersive virtual reality environment available for virtual reality studies of complex scientific problems such as macroscopic fracture simulations and protein folding.

3.2 Requested Equipment

The table below lists the requested items by year. For year one, exact pricing, manufacturers, and model numbers are given for the equipment requested. In subsequent years, the equipment costs are estimated. Specifically, we have assumed that the cost of tapes, and of RAID 5 disk storage that can connect to systems with large RAM and multiple processors, drops roughly 20% per year. The figure of 20% is very conservative (some estimates show a drop of 45% or more per year). We wanted to ensure that the budget would provide at least the minimum storage needed to begin to achieve our research goals. If more storage is available in the later grant years, we will definitely have the data to take advantage of it. We anticipate that we would adjust the exact ratio of disks to tapes, and the actual processing power and RAM in the computation nodes appropriately to most effectively meet the needs of the research projects.

We would like to note that the budget numbers on the next page include employee benefits, but no indirect costs. The Cornell cost sharing in the amount of $917,245 is explained in more detail in Section 3.7 and does not appear in this budget.

Year	Item	Qty	Cost/Item	Total
Year 1	Unisys ES7000 Model 420 computer systems including 16 Intel Itanium 2 processors, 64GB memory, and 50 Terabytes ATA RAID disk	2	$238,000	$476,000
	Development Server: Dell Poweredge 2600, Dual 3GHz Xeon, 6GB memory, single 36GB SCSI, 3Ware Escalade 8506-8 Serial ATA RAID card, 8 bay Serial ATA disk enclosure, 8 250GB Serial ATA 7200RPM drives	1	$8,042	$8,042
	Gigabit fiber networking cards	4	$500	$2,000
	Technical Support personnel			$13,100
	Total Cost for Year 1			**$499,142**
Year 2	14.7 Terabyte Nodes, consisting of Dell Poweredge 2600, Dual 3GHz Xeon, 6GB memory, single 36GB SCSI ($4272), 6 3Ware Escalade 8506-8 Serial ATA RAID cards ($490 each for total of $2,940), 3 16-bay Serial ATA disk enclosures ($3000 each for total of $9,000), 48 350GB Serial ATA 7200RPM drives ($285 each for total of $13,680)	2	$29,892	$59,784
	50TB disk expansion for Unisys ES7000 cluster	1	$150,000	$150,000
	LTO-3 drive upgrades for ADIC Scalar 10K	6	$8,450	$50,700
	100TB LTO-3 tape expansion for ADIC Scalar 10K	1	$40,750	$40,750
	Gigabit networking hardware			$4,000
	Technical support personnel			$10,316
	Total Cost for Year 2			**$315,550**
Year 3	100TB expansion to Unisys ES7000 cluster			$250,000
	LTO-3 drive upgrades for ADIC Scalar 10K	6	$7,500	$45,000
	150TB LTO-3 tape expansion for ADIC Scalar 10K			$45,000
	Technical support personnel			$7,221
	Total Cost for Year 3			**$347,221**
Year 4	100TB Compute/Storage Nodes	2	$200,000	$400,000
	150TB tape library expansion			$35,000
	Networking upgrades			$15,000
	Technical support personnel			$3,791
	Total Cost for Year 4			**$453,791**
Year 5	100TB Compute/Storage Nodes	2	$175,000	$350,000
	Networking upgrades			$10,000
	Total Cost for Year 5			**$360,000**
	Total Request			**$1,975,704**

3.3 Rationale for Requested Equipment

Funding of this proposal will enable three research projects in the areas of computer graphics, the World Wide Web, and astronomy. These three projects have the following common characteristics: (a) they are extremely data-intensive, requiring large-scale data storage and data manipulation capabilities, and (b) they require a tight coupling of storage and computational infrastructure in order to perform complex operations on the data sets.

We propose to use the resources from this grant to develop the required storage and networking infrastructure. Since online disk storage is relatively expensive (compared to near-line or offline tapes), there will be 300 TB of shared disk storage. In addition, each project will have up to 100-150 TB of dedicated disk space. This works well for our projects because they each have a temporary need for large amounts of online storage when converting raw data into its processed form, but have reduced (albeit still large) storage requirements when operating on processed data. Each project will have access to the required amount of tape storage.

We now describe the storage and computational needs of each project in more detail.

3.3.1 Physically Accurate Rendering in Computer Graphics

In the first research project, Cornell graphics researchers are developing techniques for physically accurate rendering. The goal here is to start with a model of a three-dimensional scene, and predict with quantitative accuracy what would be seen if that same scene were viewed from a particular vantage point in reality. This requires the development of accurate reflectance models of real objects, and also new rendering algorithms for physically accurate renderings.

Storage needs: In order to develop accurate reflectance models of real objects, Cornell researchers are acquiring an instrument called the Spherical Gantry. The Spherical Gantry is a computer-controlled device that can illuminate an object from any direction using a projector, and measure the light reflected from the object in any other direction using a camera. The storage requirements for the Spherical Gantry are quite large. For example, if we were to capture light reflection measurements at 7 degree spacings over half the sphere using moderate resolution in all dimensions, this will result in about 50 terabytes of data *for each object*. Capturing such measurements for any reasonable collection of objects thus leads to significant storage requirements.

Computational needs: In order to efficiently render complex scenes, one of the directions being explored is to precompute extremely large data sets of object appearance using measurements from the Spherical Gantry. Besides the storage requirements for the precomputed objects, which is expected to run into tens of terabytes, it will require processing power equivalent to hundreds of current commodity processors in order to compute this data.

3.3.2 The Structure and Evolution of the World Wide Web

In the second research project, Cornell researchers are studying the evolution of the World Wide Web, and are also developing new models of the Web.

Storage needs: The storage needs for a single static snapshot of the World Wide Web containing about a billion pages are already quite large: about 1 TB of uncompressed data, or 100GB of compressed data. However, a principled study of how the Web evolves over time requires access to *multiple* snapshots of the Web. Cornell researchers have an ongoing relationship with the Internet Archive (http://www.archive.org), which has archives of multiple snapshots of the Web since 1996. The size of this collection is about 500 TB and it is growing at the rate of 10 TB per month. Even if the proposed research is limited to a time window of 2-3 years, it still requires several tens to hundreds of terabytes of storage. Other valuable information such as Web access logs also requires additional tens of terabytes of storage.

Computational needs: Web research is not done on raw Web data, but on features extracted from the data, such as term sequences, term vectors, and hyperlinks. Given the scale of the Web, writing the code to extract these features is a non-trivial exercise. Specifically, for efficient large-scale processing, feature extraction requires massively parallel execution and tens to hundreds of terabytes of temporary storage during computation. Besides tightly coupled storage and computation, the infrastructure needs to be available for research purposes. For example, when we have attempted to run some of our computations on the Internet Archive computers, we have had to ensure that we do not perform large compute or I/O intensive operations that could interfere with their production service. Despite our care, there have been many cases where our programs have been killed by the system operators because of such interference. Thus a large data store tightly coupled with high capacity computing specifically for research is of fundamental importance to advancing our research.

3.3.3 Large-Scale Astronomical Surveys using the Arecibo Telescope

In the third research project, Cornell researchers are analyzing data from the Arecibo Telescope to find pulsars and other exotic objects.

Storage needs: This project requires storage for the raw data from the Telescope and for the processed and derived data. Our estimates of the size of the raw data are as follows: Assuming an overall number of $25,000$ pointings of the Telescope with a 300 second dwell time per pointing and an anticipated data feed rate of 100 MB per second, the total size of the raw data will be around 800 TB. The raw data will be processed (as explained in Section 6.3), and the derived data will be reduced by a factor of 8. The raw data can be stored on tape with the assumption that from time to time a large chunk of the raw data can be brought to secondary storage for processing. The processed data will require about 100-200TB of online storage by the time the survey has finished, and it is the processed data that we plan to make accessible online to the Astronomy community through a web service interface.

Computational needs: The raw data will be subjected to a novel multistep data analysis for identifying real celestial signals with expected (and hopefully new) signatures. Efficient analysis and processing of the raw data is a very challenging data mining problem, and requires research into novel data mining methods that efficiently search large parameter spaces for signatures of interesting celestial objects. This large-scale processing requires a system that tightly integrates large-scale storage with high processing power as data mining methods are data intensive but also pose a very high computational load and require large amounts of main memory.

3.4 Equipment and Maintenance Cost

The Unisys ES7000 systems come with one-year warranty and maintenance (24/7 - 4 hour response). After the first year, any required maintenance support will be obtained by the Cornell Theory Center, and these machines will be maintained and supported by the Cornell Theory Center staff. There will be no direct charge to the project for the maintenance of these systems.

The tape library (ADIC Scalar 10K) is supported and maintained by the Cornell Theory Center. This project will only be using a portion of this shared resource, and there will be no direct charges to the project for the maintenance of the tape library.

Much of the other equipment (Dell systems, networking hardware, and disk drives) comes with three to five year warranties. This covers direct hardware replacement costs for these components. Local hardware support costs are budgeted in the grant as "Technical Support Personnel". Typically, these warranties cover most of the useful research life of the equipment.

Operating systems and some applications software and software updates for these systems are being donated by Microsoft. Other application software will either be developed by the researchers themselves, or is available as part of a campus licensing arrangement at no direct cost to the project. Staff costs to support software installation and maintenance on the systems are included in the budget as "Technical Support Personnel", with an increasing portion being taken on by Cornell over the period of the grant.

3.5 Access to Equipment

The proposed equipment will be connected to both the Cornell Theory Center network and the Computer Science network via Gigabit Ethernet. Access to applications and development tools will be available to Cornell researchers over those networks and, as required, through Cornell's external Internet2 and commodity Internet connections.

In addition, dedicated Gigabit optical fiber connections will run from the Program of Computer Graphics and the Computer Science Department to the Unisys ES7000 systems in the Cornell Theory Center machine room. This will enable a high-speed connection between our existing computational resources and the requested grant equipment.

Researchers do not require physical access to the proposed infrastructure. Thus physical access will be limited to support personnel in the Computer Science Department and the Cornell Theory Center for maintenance and operations.

3.6 Space renovation

The equipment will be housed in the existing machine rooms of the Cornell Theory Center (in Rhodes Hall) and the Computer Science Department (in Upson Hall). No renovations to the existing spaces are required to support this equipment.

3.7 Cost sharing

Cost sharing for the proposal is in three components. The first is a donation of $599,628 from Unisys Corporation. This represents the difference between the normal educationally-discounted price of the Unisys ES7000 package (see budget) and the amount Cornell is paying for these systems. The second component is a $40,000 donation of Intel Itanium processors from Intel Corporation. The third component is a Cornell contribution of $277,617 in technical staff support from the Cornell Theory Center. This staff will provide software installation, maintenance, and systems operations support for the Unisys compute/storage servers. The total cost-sharing is $917,245.

4 Management Structure

The PI and Director, Alan Demers, Professor of Computer Science, will oversee the development of the proposed infrastructure. As part of overseeing the development, the PI, the Co-PIs and Senior Personnel will hold quarterly reviews. As all the investigators are at Cornell, no travel is required.

The proposed infrastructure will be managed partly by the Cornell Theory Center, and partly by the Department of Computer Science. The Cornell Theory Center, which reports to the Vice Provost for Research, is responsible for faciliating computational science at Cornell. The Cornell Theory Center is headed by Professor Thomas F. Coleman of the Department of Computer Science. A 12-person faculty advisory committee provides periodic input to the Director and Vice Provost. The Center has considerable experience integrating a variety of resources, and currently operates resources for projects funded by the NSF, NIH, and USDA. A large part of the proposed new infrastructure (the two Unisys servers, the disk storage, and the tape library) would be installed and maintained by the Cornell Theory Center, and we can leverage their expertise in this area.

The Department of Computer Science has a central facilities group that is responsible for supporting all the research and teaching computers and networking infrastructure in the department. The part of the proposed infrastructure that will be hosted in the Department of Computer Science will be installed and maintained by this group. The facilities group currently consists of seven staff members: four programmers, one hardware technician, and two user-support staff. They install and support software on the departmental computing systems, perform hardware installation and maintenance, maintain the computer network, and handle administrative tasks like account maintenance and backup.

The facilities group is supervised by our Director of Computing Facilities, Dean Krafft, and is overseen by the Computer Facilities Executive Committee. This committee consists of three Computer Science faculty members and Dean Krafft. The faculty members are appointed by the chair of the Department of Computer Science, who also names the chair of the committee. The committee meets periodically to set policy for the facility and to discuss and approve all major purchases. The chairman of the committee meets at least weekly with Dean Krafft to provide regular guidance regarding faculty input to the management of the facility. The PI of this proposal (Alan Demers) has previously served as chair of this committee, and will again be appointed to this committee if this proposal is funded.

5 Results from Prior Awards

In 1997, the Cornell Computer Science Department received an award for developing technologies for “A Next Generation Computing and Communications Substrate” (97-03470). This award ended on July 31, 2003. The project developed technologies for a new computing and communications substrate that provides anything-anywhere-anytime (a3) access to information. This new learning environment features: (a) resources for creating and disseminating information, (b) decoupling of location and function to enable project collaboration across campus, to allow lecture access from dormitories and to provide library search capabilities from the classroom or office, (c) qualitatively better information resources that blur the boundary between documents and applications, and (d) support for multiple styles of learning.

In terms of infrastructural outcomes, with significant assistance from Microsoft and Intel, we created a software development laboratory and network infrastructure to support the creation of the next-generation computing and communications infrastructure we had proposed. We leveraged this project and the infrastructure it provided through a successful $300,000 proposal to Intel for support to investigate the student use of continuously networked nomadic computing on laptop computers both in and out of the classroom. We further leveraged funding from NSF and Intel to create a campus-wide wireless network service to continue the creation of an anytime, anywhere computing environment for students and researchers at Cornell University.

The following are specific research results that resulted from the award.

- The Nomad project, which was a joint research project between the the digital libraries group in Department of Computer Science and the computer science and the human computer interaction group in the Department of Communication, explored the use of laptop computers and wireless networks in education. An extensive wireless network was installed covering many academic buildings, classrooms and libraries. With this infrastructure, the usage patterns of some 200 students in two classes were studied. Specific observations were made on the use of online readings, tools for collaborative work and even the use of wireless laptops in examinations. Results from this research have been reported in both the popular press and a series of research articles.

- Support for networking protocol development, including a new system called ”Spinglass”. The Bimodal Multicast protocol underlying Spinglass appears to be a real step forward, because it not only provides a guarantee of scalability and stable throughput under stress, but also has a unique form of reliability - a bimodal probabilistic delivery distribution, which can be formally derived and has been confirmed in practice.

- Security research, including COCA, a fault-tolerant and secure online certification authority, which has been completed and deployed both in a local area network and in the Internet. Replication is used in COCA to achieve availability; proactive recovery with threshold cryptography is used for digitally signing certificates in a way that defends against mobile adversaries which attack, compromise, and control one replica for a limited period of time before moving on to another.

- Programming languages, tools and applications, including adapting advanced type systems for use in low-level languages, so that we may simultaneously achieve safety and performance. Towards this end, we focused on general-purpose type structure for both Intel IA32 (x86) Assembly Language and type structure for C-like programming languages. Though these languages are often used directly by programmers, they are increasingly used as target languages by compilers. We extended the type system and type-checker for TALx86 – a typed variant of Intel's IA32 (x86) assembly language. We studied general type constructors, including alias types, capability types, and dependent types, that allow more language constructs to be encoded and that allow a compiler to perform more optimizations.

Over 200 Ph.D., Masters of Engineering, and undergraduate students were indirectly supported under the grant. Some pertinent publications that resulted from this grant are:

- Michael Jones, Robert Rieger, Paul Treadwell, Geri Gay, "Live from the Stacks: User Feedback on Mobile Computers and Wireless Tools for Library Patrons", *ACM Digital Libraries*, June 2000.
- Robbert van Renesse, Kenneth Birman, and Werner Vogels. "Astrolable: A Robust and Scalable Technology for Distributed System Monitoring, Management, and Data Mining", *ACM ACM Transactions on Computer Systems*, 21(2), May 2003.
- Lidong Zhou, Fred Schneider, Robbert van Renesse. "COCA: A Secure Distributed On-line Certification Authority", *ACM Transactions on Computer Systems*, 20(4), November 2002.
- Trevor Jim, Greg Morrisett, Dan Grossman, Michael Hicks, James Cheney, and Yanling Wang, "Cyclone: A Safe Dialect of C", *Proceedings of the USENIX Annual Technical Conference*, Monterey, CA, June 2002.

The Cornell Theory Center, along with Iowa State, Notre Dame University, and the University of Delaware, currently has an NSF RI award (99-72853) for a project entitled "A Two-Tier Computation and Visualization Facility for Multiscale Problems." The goal of this work is to develop and implement key enabling Computer Science technologies for multiscale scientific computing simulations, such as crack propagation in solids, which require an enormous amount of computing power and advanced visualization engines for rapid prototyping. The project aims to develop, (a) finite-element methods (FEM's) used in macroscopic fracture simulation that require 3-dimensional meshes of high quality; (b) accurate simulations of protein dynamics and folding; (c) virtual reality platforms for visualization of complex scientific problems.

The equipment requested in the proposal consists of a two-tiered architecture: The entry-level tier, which consists of a number of small-scale NT-based symmetric multiprocessors, is to be used in the development and prototype stages. The second production-level tier will consist of a large, shared NT-based computing cluster built out of symmetric multi-processors and a CAVE virtual reality system. The initial CAVE facility will be an upgrade of the existing Onyx-based CAVE at the Cornell Theory Center.

A number of significant research results have been produced using equipment purchased under this RI award. For example, we have developed a new method for parallelizing 3D constrained Delaunay mesh generation that will improve the accuracy of FEM-based simulations and enable efficient geometric adaptivity. We have also developed a new kind of "consensus shape" for protein families that is based upon on the Unit-Vector Root Mean Square (URMS) distance and that provides a compact summary of significant information for a family of proteins. There are many more example.

Clusters based on commodity off-the-shelf technology, like CTC's V1/V+, deliver the necessary performance and reliability to solve large-scale FEM problems in computational fracture mechanics and biology. For instance, we have found that, with the efficiency and capacity of our parallel FEM-based system, we can more closely match detailed geometry, complex boundary conditions, can use much smaller time steps, and are therefore getting closer matches between predicted and observed crack shapes and fatigue lives. Also, our work in Protein Folding has been dramatically improved by the Velocity complex. For instance, the matrices produced by our methods very often have 10's and even 100's of millions of rows and columns. It is simply infeasible to process such matrices on uniprocessor workstations.

Some pertinent publications that resulted from this grant are:

- B. Carter, et al., "Parallel FEM Simulation of Crack Propagation - Challenges, Status, and Perspectives", *Proceedings of Irregular 2000*, LNCS vol. 1800, Springer-Verlag, 2000.
- Veaceslav Zaloj and Ron Elber, "Parallel Computations of Molecular Dynamics Trajectories Using the Stochastic Path Approach", *Computer Physics Communication*, vol. 128, 2000.
- Paul Chew and Klara Kedem, "Finding the Consensus Shape for a Protein Family", *Proceedings of the 18th ACM Symposium on Computational Geometry*, 2002.
- Demian Nave, Nikos Chrisochoides, and Paul Chew, "Guaranteed-Quality Parallel Delaunay Refinement for Restricted Polyhedral Domains", *Proceedings of the 18th ACM Symposium on Computational Geometry*, 2002.

In 1999, Cornell University received an NSF CISE Educational Innovation Award (98-12630) to support the development of formally-grounded courseware for teaching both mathematics and computing. The project is creating easily used and novel educational technology by modifying existing research tools and incorporating ideas from the practice of machine assisted reasoning to make plain the conventional logical structure of mathematics text.

In 2000, Cornell University (specifically, the Cornell Nanofabrication Facility) received an NSF MRI award "Equipment for Remote Usage, Access, and Learning at NNUN" (00-79443). Cornell acted as the prime awardee for this contract which supported computer and telecommunications infrastructure at the five sites of the National Nanofabrication Users Network (NNUN) located at Cornell, Stanford, Howard, Penn State, and UC Santa Barbara. The award provided compatible video conference equipment across the network to support more efficient interaction and collaboration with users. It provided compatible video processing systems across the 5 network sites to support the development of interactive training media, and finally it provided additional computer aided design resources for the network. The other 4 NNUN sites participated as subcontractors to Cornell. These technology improvements have been implemented and have significantly enhanced the interaction of the networks 2000 annual users with the host facilities.