

MIGRATING AN ENTERPRISE GIS - A SUCCESS STORY

Amy Ramsdell, Dr. Rachel Arulraj, and Dr. Mary Daum

Abstract: Brookhaven National Laboratory (BNL) has used GIS for over a decade. During this time, the system has grown in scope, size and complexity. BNL is a multi-purpose research facility operated for the U.S. Department of Energy by Brookhaven Science Associates. The core GIS was originally developed as part of a comprehensive Environmental Information Management System (EIMS) to support the Lab's environmental cleanup activities. Currently, EIMS-GIS has expanded to include facility operations and maintenance, emergency response, and routine environmental surveillance and compliance. To cater to a more diverse customer base, coupled with increasingly complex demands and decreased resources, BNL carried out the migration of its Enterprise GIS. This paper covers the organizational and technical issues encountered in the migration of BNL's decade old Enterprise GIS. It identifies successes, lessons learned, and ongoing issues for a mature Enterprise GIS.

INTRODUCTION

Brookhaven National Laboratory (BNL) is a multi-purpose research facility operated for the U.S. Department of Energy (DOE) by Brookhaven Science Associates. The site housed Army camps during World Wars I and II. Since 1947 it has been home to world-class research facilities for high-energy physics and other energy-related sciences. An unfortunate legacy of these military and scientific activities has been environmental pollution. In the early 1990s, BNL established an Environmental Restoration (ER) program as part of its cleanup activities under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA). In 1993, an integrated GIS and database system, called the Environmental Information Management System (EIMS), was built to store, analyze and visualize environmental characterization data generated by the ER program.

Over time the EIMS has grown to serve a larger customer base and extended its scope to encompass facility operations, maintenance, emergency response, waste management and routine environmental surveillance and compliance. These will be the primary customers once BNL's CERCLA mission ends, scheduled for completion in late 2005. However, these organizations have very limited resources to spend on GIS-related needs. To prepare for this new phase, the EIMS has focused on becoming more efficient, aiming to provide services to a wide variety of customers with minimal resources. The time was right to shift the GIS from a UNIX/ArcInfo and Oracle relate environment to the more recent geodatabase format on Windows. Moving away from the file-based coverages to the database-oriented geodatabase involved a major migration endeavor.

THE DECISION TO MIGRATE

When one typical thinks of an enterprise GIS, one thinks of a grand or large-scale geographic information system. BNL's enterprise GIS, while similar in architecture to larger systems, supports a comparatively smaller organization of GIS users and customers. Three main criteria drove the decision to migrate the current system to an enterprise GIS.

There was a need to reduce the complexity of data sharing among the different divisions generating data at the Laboratory. The initial source of spatial data for the GIS was the BNL Plant Engineering Division (PE) site map, a series of 131 multi-layer CAD files that used a local coordinate system. However, ER needed to map offsite areas and share data with local agencies, so the GIS staff began maintaining the site map layers as well as adding new environmental layers to a separate dataset of coverages in the State Plane Coordinate System. The ability to share current data between these two groups has been a challenge over the years because of the use of two different coordinate systems, the differences in the software, and the cross-platform steps of the transformation process. An enterprise GIS design with ArcGIS and ArcSDE would provide the foundation and tools necessary to simplify the data sharing process thereby increasing productive use of the data.

A second criterion was the production and revision of maps more efficiently. Demand for system products and services increased rapidly with the growth of the ER program from eight staff members in 1992 to 40 in 1996, to 80 by 2002. To meet this demand, GIS staff had built up an extensive collection of utility programs written in Arc Macro Language (AML) to automate routine GIS tasks and to standardize elements of map production. The ability to generate AMLs in the UNIX environment is a specialized skill. The decision to move new map production to ArcView and subsequently to the ArcGIS Desktop version 8.x and ArcSDE environment allowed more user-friendly access to the data by a wider range of people. The migration offered an opportunity to redesign system elements and processes to manage new data types and output requirements more effectively.

As ER work shifted from characterization to active cleanup and construction of treatment systems and now closeout of these work packages, the GIS budget began shrinking and departures of GIS staff were left unfilled. In order to provide the necessary level of support of base services, ArcIMS was used as the architecture to provide a range of data access tools to the ER program and a growing number of customers in other divisions. A subset of coverages was converted to shapefiles and served over the intranet reasonably quickly. The web applications also permitted the integration of other types of information with the spatial data maintained by the GIS. An early success was a tool that gave BNL hydrogeologists access to information that previously existed only in bulky paper records. This phase of development was instrumental in improving access to the otherwise restricted database. However, a second dataset was generated that had to be maintained because coverages were still used for mapping. The third criterion was to have a more efficient way that staff could maintain the datasets for mapping and for ArcIMS.

The EIMS's expansion in scope and imminent reduction in funding have coincided with advancements in data management technology. GIS staff determined that migration of the somewhat loosely connected spatial and attribute data from table- and file-based systems to a geospatial data management environment would help to achieve the needed efficiencies and provide additional flexibility to meet future challenges.

ENTERPRISE GIS MIGRATION

The Migration Process

It is fairly expensive to change system design of an existing database every time technology provides a major advancement. The EIMS manager was able to obtain project-specific funding

for the migration. Fortunately, the EIMS program budget had enabled GIS staff to stay current with software, so most of the project budget could be directed to the necessary design and implementation.

The decision to migrate to ArcSDE and the geodatabase was made over the course of one year spent gathering information on the new technology. Once the software was purchased and installed in 2001, initial testing of the geodatabase began. Select data layers that were used for existing ArcIMS applications were migrated to the database and the web application code was modified. During this testing phase, one GIS staff member and a database administrator attended the ArcSDE training courses. Testing began prior to the training because it was more valuable to attend the training prepared to discuss problems encountered already. BNL then recruited the services of Parsons Brinckerhoff (PB) in 2002 to support the migration of the entire dataset of coverages and to develop custom applications to automate some of the routine tasks associated with data updates to the GIS. The migration effort was complicated by the fact that it was more than just data migration. The system was required to undergo a hardware/software and business process migration as well. The conversion transformed ArcInfo coverages and shapefiles to a geodatabases, Oracle tables to SQL server tables, AMLs to Visual Basic (VB) programs, and moved all from a UNIX to a Windows-based system. The migration progressed from February to September for a total of 78 workdays in a very carefully planned and organized manner. The initial planning was well documented, the process of migration was supported with sufficient documentation and automated code, and the finish was also summarized in a working document. It was carried out within the budgetary and schedule constraints established for the project, but not without a few challenges.

Migration Challenges

Overall, the main challenge was to deal with subtle differences in the functionality of the software involved. The old system, based on UNIX ArcInfo and Oracle, dealt with the business processes of EIMS in a totally different manner compared to the Microsoft SQL Server-based geodatabase on Windows NT. The migration team consisted of members from BNL and PB with collective experience in Oracle and SQL Server, AML programming as well as with Visual Basic/ArcObjects programming, and having strong command over the UNIX operating system as well as with Windows NT.

PB developed a set of physical data models for both cases once for the old database and again for the new. Data modeling for a live enterprise GIS with the need to migrate to a new computing environment made it very challenging. Documentation of every aspect of the old database was critical to the migration project. PB also replicated BNL's system back at their office. This freed up GIS staff to continue working on the old system while experimenting on the new and allowed the PB team to become intimately familiar with the EIMS GIS.

The migration had to be performed for hundreds of data files and system files. Considering just the hundreds of ArcInfo coverages and shapefiles accumulated and maintained over the years by EIMS, it was a serious challenge. The new GIS was configured as a SQL Server-based geodatabase. It was thematically organized following stringent database design methodologies. PB developed a custom-built automated conversion process to populate the geodatabase. The coordinate system requirements for converting the spatial reference parameters to integers were

programmed into the tool. During the migration, it was discovered that some of the field names of Info tables are reserved names in the SQL Server database environment. The conversion application was written to catch and report errors in the field names. The field names were then changed in the coverages prior to migration. It was decided that any maps or AMLs would be fixed as they are used. The benefits of speed, efficiency and access, far outweighed the small sacrifices made for field name modifications. This valuable tool has continued to be used whenever additional layers need to be added to the repository.

Another technical challenge of the migration involved the conversion of the ArcInfo AMLs from the UNIX environment to the Windows environment. It was decided to convert the AMLs to preserve all the old maps in electronic, editable form. Originally only ten ArcPlot AMLs were planned for conversion. These were chosen based on their complexity and use. However, a significant accomplishment was that all the UNIX ArcInfo AMLs were converted through the use of AML scripts and UNIX shell scripts. The ArcPlot AMLs had to be rewritten with the forward slash being replaced by the backward slash to run on NT when referring to coverages. But the slashes representing database relates, which are totally different from the operating system related slashes in functionality, had to be considered separately, so, the script would work only partially. The only way to automate the migration was to make the script go through multiple passes at the hundreds of pages of AML code for different find/replace situations. The litmus test was to run the AML after the data files, relates and views were established in the new computing environment, and to verify the result of the process.

POST MIGRATION

After two years we can make an initial assessment of the new system to determine if it has indeed helped achieve the needed efficiencies and additional flexibility by examining the original criteria.

Data Sharing

The enterprise system has promoted data sharing among the different divisions at the Lab by eliminating some of the previous hindrances. Moving to NT alone has eliminated some cross platform problems. One example is logging in to a session on X-Win32 to access the UNIX drives and programs. X-Win32 allowed only a certain number of windows for all sessions. Often these windows were already being used by the database group and had to be closed out in order for the GIS staff to log in to access ArcInfo Workstation. Mapping to Unix drives in NT also caused frequent lock outs from the NT network when logging in to the BNL domain during computer startup and reboot.

Another benefit is that CAD drawings can be used in ArcMap projects. If there is no time or resources allotted for conversion of the CAD layers to the geodatabase, the drawing can be used in the map document just as if it was another GIS layer with query and symbol formatting capabilities.

While we still have a long way to go integrating the workflow between the users of CAD and ESRI software here at the Lab, a significant advancement was made in the transformation between the two different coordinate systems used that would not have been possible in the old GIS system. A coordinate system expert and one of the programmers at ESRI were able to

define BNL's local coordinate system in a projection file. This has enabled GIS staff to update GIS layers based on CAD site map information in the ArcGIS Desktop environment and eliminated the need for the old AMLs. For example, to update a GIS roads layer, the site map CAD drawing is dropped into an ArcMap edit session with the GIS roads layer using the BNL projection on the fly. The layer is then updated as needed and saved, retaining its original state plane coordinates.

There are many layers in the geodatabase that are also needed by CAD users and offsite contractors who do work for BNL. SDE's CadClient is one way to use the geodatabase in CAD although it has not been very successful because of the incompatibility of the software versions for each product. It was tested briefly in AutoCAD Map 2000 but those versions have since been upgraded to Map 2004 and other CAD users do not have Autodesk product versions tested and supported by ESRI. When sharing data layers with contractors and other divisions in the Lab, shapefiles are used. The shapefiles are updated by exporting the layer from the geodatabase. This is easy enough through ArcCatalog, however the metadata has to be fixed each time because the mapping of required fields in the attributes are different for each. This causes the attribute definitions in the shapefile metadata to be incorrect. This problem is supposed to be fixed in ArcGIS 9 but upgrading may not be an immediate option. Resources are much more limited now and these software upgrades take up a considerable amount.

Map Production

The new system has enabled the production and revision of maps more efficiently. Much of the original customization that was done through AMLs has been superceded by functions that are now built into the current GIS software graphical user interfaces (GUIs). For example, formatted geodatabase layer files are now used to standardize the appearance of data layers and their labels. A VB expression has replaced an ArcInfo relate to an Oracle view which contained the field used for labeling. Spatial views have been constructed in the database to pull some attribute information originally only accessible in Oracle and not through ArcInfo. Now one-to-many relationships are built using the SQL Server views.

Many map requests can now be handled in under one hour. New requests from other divisions are able to be completed more quickly and therefore at a lower cost. The ease of making maps in the GIS has enabled an expansion of the customer base to include GIS projects for cultural and natural resources, emergency services, future land use, and building construction while still providing support to the ER program.

Another added benefit of the migration is noticed with the greater reliance on student help. It is much easier for inexperienced users to understand the new software because, for the most part, they are already familiar with the NT environment and the software GUIs. One challenge yet to be overcome deals with the expansion of products and services. In order for the GIS to gain new customers and new mapping projects, the system has to have access to the information stored in other databases. Attempts to make a live connection from the geodatabase using views to a linked server's SQL Server business tables have so far been unsuccessful and unsupported by ESRI. Fortunately there are other ways to approach this such as potentially rebuilding the tables in the geodatabase with nightly updates.

Data Maintenance

Lastly, the enterprise system created a more efficient way to maintain the datasets for production of maps and for ArcIMS. A major customization effort by PB during the migration phase assisted the sample point update procedure to the GIS. Input of new sample data collected by field crews, which previously followed a complicated flow chart of paper documents, logs, data entry operators, multiple spreadsheets, database commands, and “how to” procedures, is now partially encapsulated into a custom built tool on the ArcCatalog interface. This VB program, written to update the geodatabase, along with an update application previously written to update Oracle and the tools in ArcGIS Desktop, together have greatly reduced the flow chart steps and made the staff more accurate and efficient in carrying out their update activities. Having the BNL projection file has also enabled GIS staff to edit data and maintain the most current information on the websites much more easily than previous methods. Also, small VB macros were written and others were downloaded from the ESRI ArcScripts collection to enable all data layer updates and maintenance to be performed in the ArcGIS Desktop instead of using ArcView and Avenue for some processes and AMLs for others.

All but one of the ArcIMS mapping applications have been converted to using geodatabase layers. This makes updates to the data layers possible without interruption of web services and instantly visible in the applications. Now there is really only one dataset, the geodatabase that is maintained in its entirety. The coverages and shapefiles are only updated when needed, where previously all three had to be maintained for the different programs in which they were used.

CONCLUSION

Overall the migration effort was a success by enabling the EIMS GIS to become more efficient while at the same time expanding the customer base in the aforementioned ways. Continued maintenance and use of the GIS enterprise system in Version 8.2 or upgrading to the newest version is a question that still needs to be answered. If an upgrade is possible it will be the next biggest challenge aside from expansion of GIS services from the ER program to the rest of the Laboratory. Diagrams of the old GIS enterprise, the migration effort in 2002, and the post-migration GIS enterprise system appear in the Appendix.

ACKNOWLEDGEMENTS

The authors would like to recognize the following people for their greatly appreciated contributions to the GIS enterprise system migration:

Brookhaven National Laboratory

Alain Domingo, Database Programmer

Daqi Li , former Database Administrator

Frank Tramontano, Network Administrator

Parsons Brinkerhoff

Jules Loweff, Senior Project Manager

Feng Lu, Computer Programmer

ESRI

Melita Kennedy, Product Specialist and technical expert in coordinate systems, Redlands

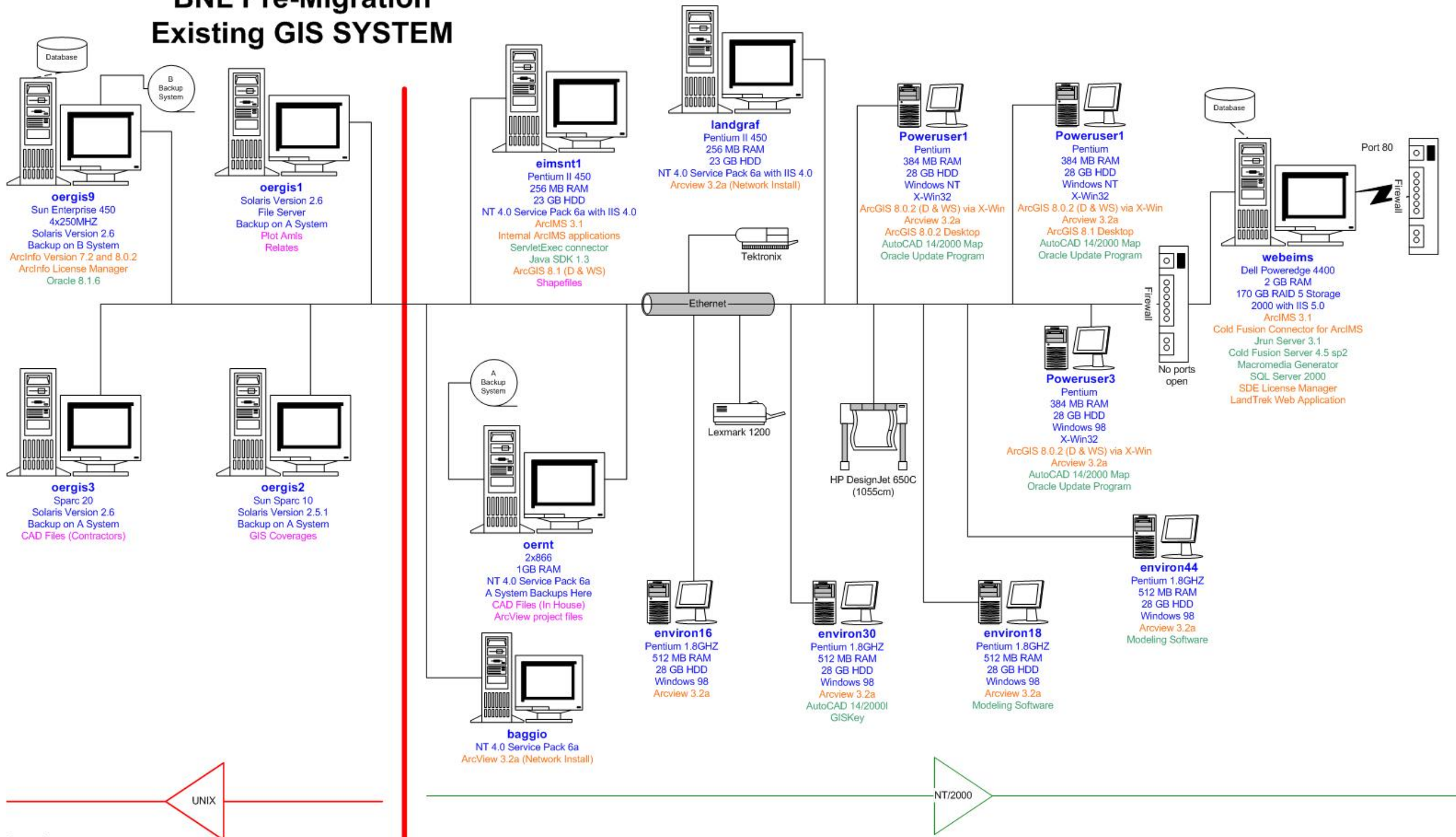
Matt Kennedy, Regional Representative, NYC office

Gregg Ziak, Technical Support, Redlands

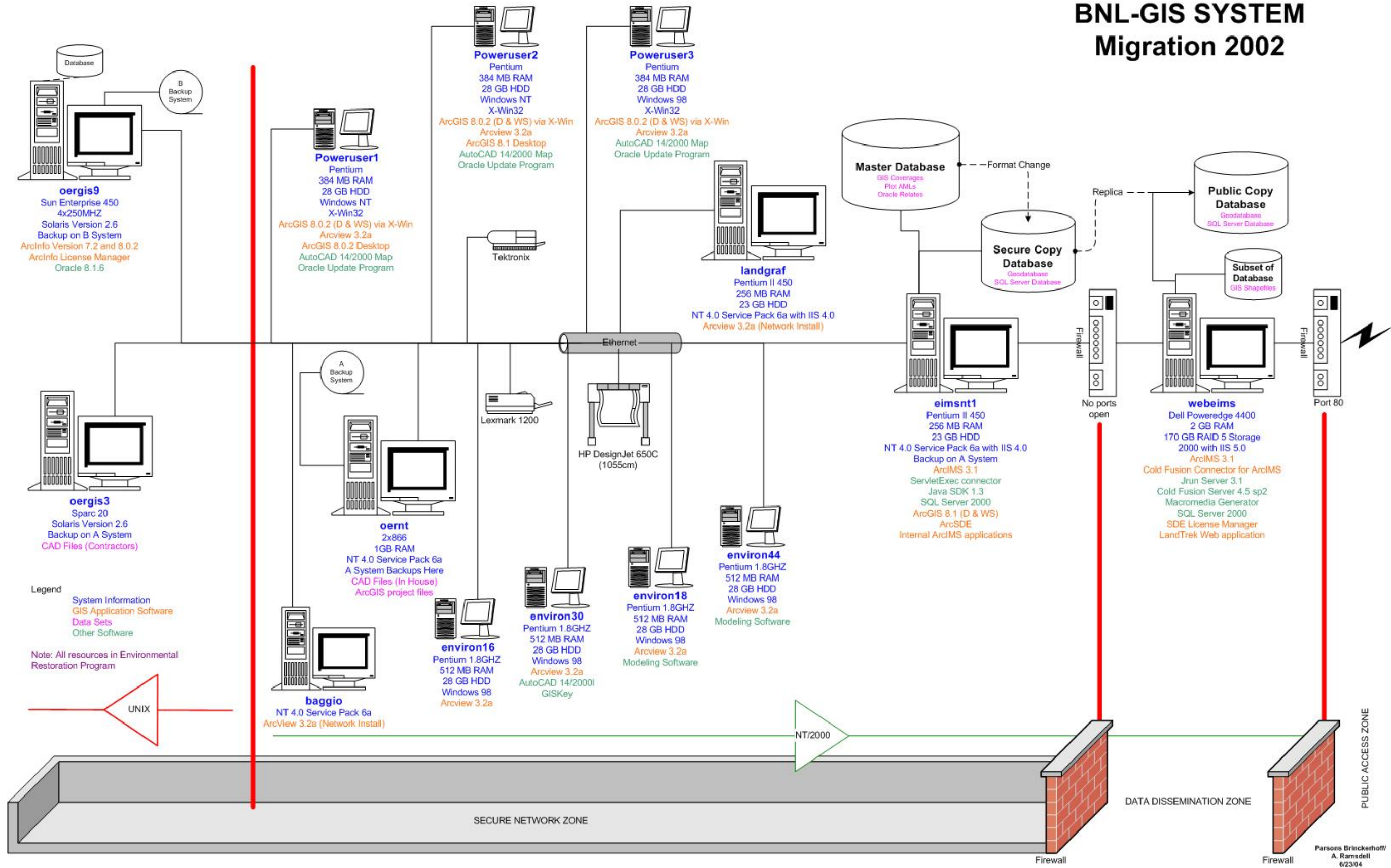
As well as countless other technical support representatives manning the support phone lines.

APPENDIX – SYSTEM OVERVIEW

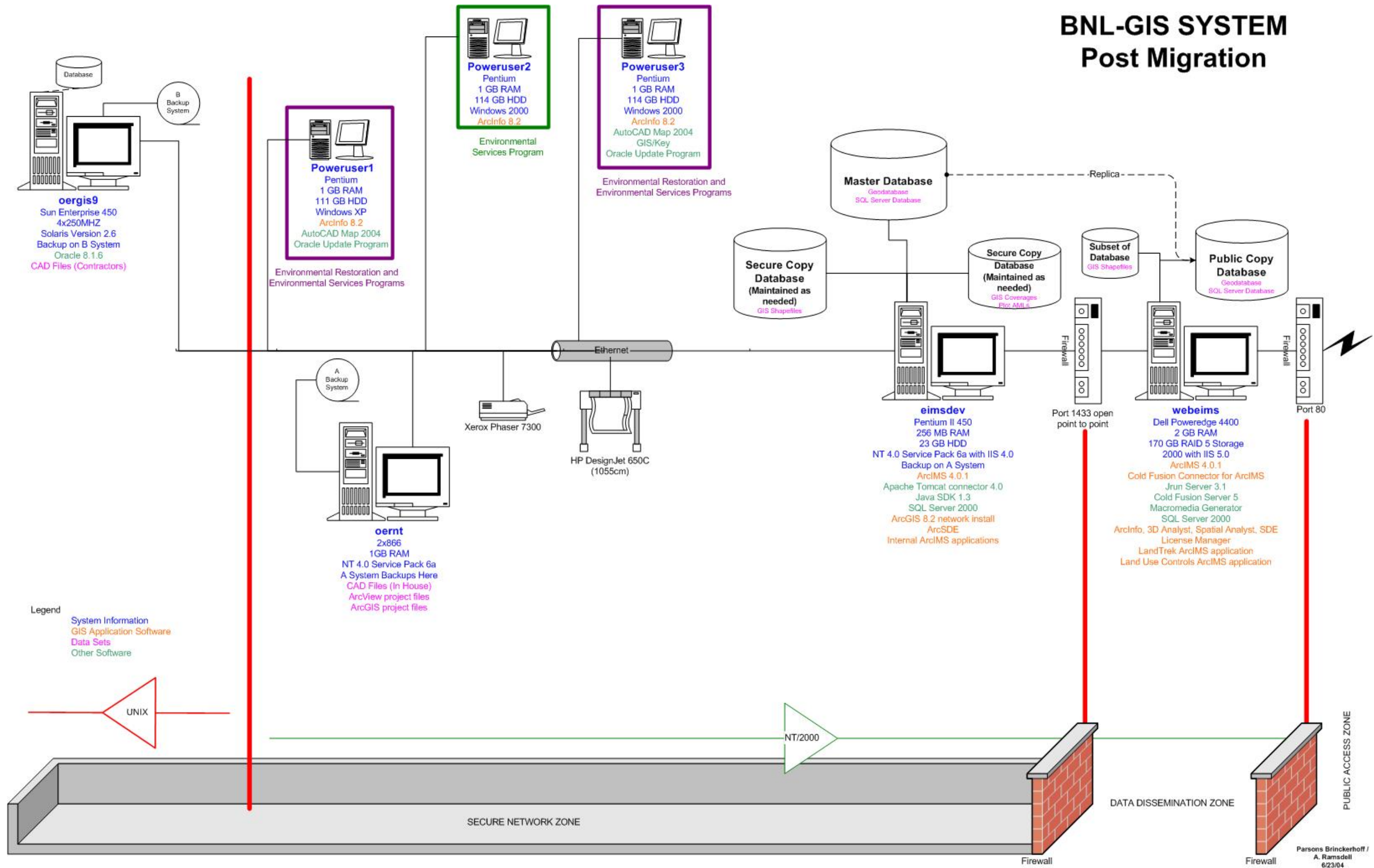
BNL Pre-Migration Existing GIS SYSTEM



BNL-GIS SYSTEM Migration 2002



BNL-GIS SYSTEM Post Migration



REFERENCES

Arulraj, Rachel K. and Mary L.Daum. "Challenges in Migrating an Enterprise GIS – The BNL Experience", Proceedings of URISA 2003 Annual Conference, Atlanta, Georgia, October 2003.

AUTHORS

Amy Ramsdell
GIS Manager
Bldg. 51, Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973
Phone: 631-344-3135
Fax: 631-344-7776
Email: ramsdell@bnl.gov

Dr. Rachel Arulraj
Parsons Brinckerhoff Quade & Douglas, Inc.
One Penn Plaza
New York, NY
Phone: 212-465-5116
Fax: 212-465-5595
Email: arulraj@pbworld.com

Dr. Mary Daum
Environmental Information Management System
Bldg. 51, Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973
Phone: 631-344-2066
Fax: 631-344-7776
Email: daum@bnl.gov