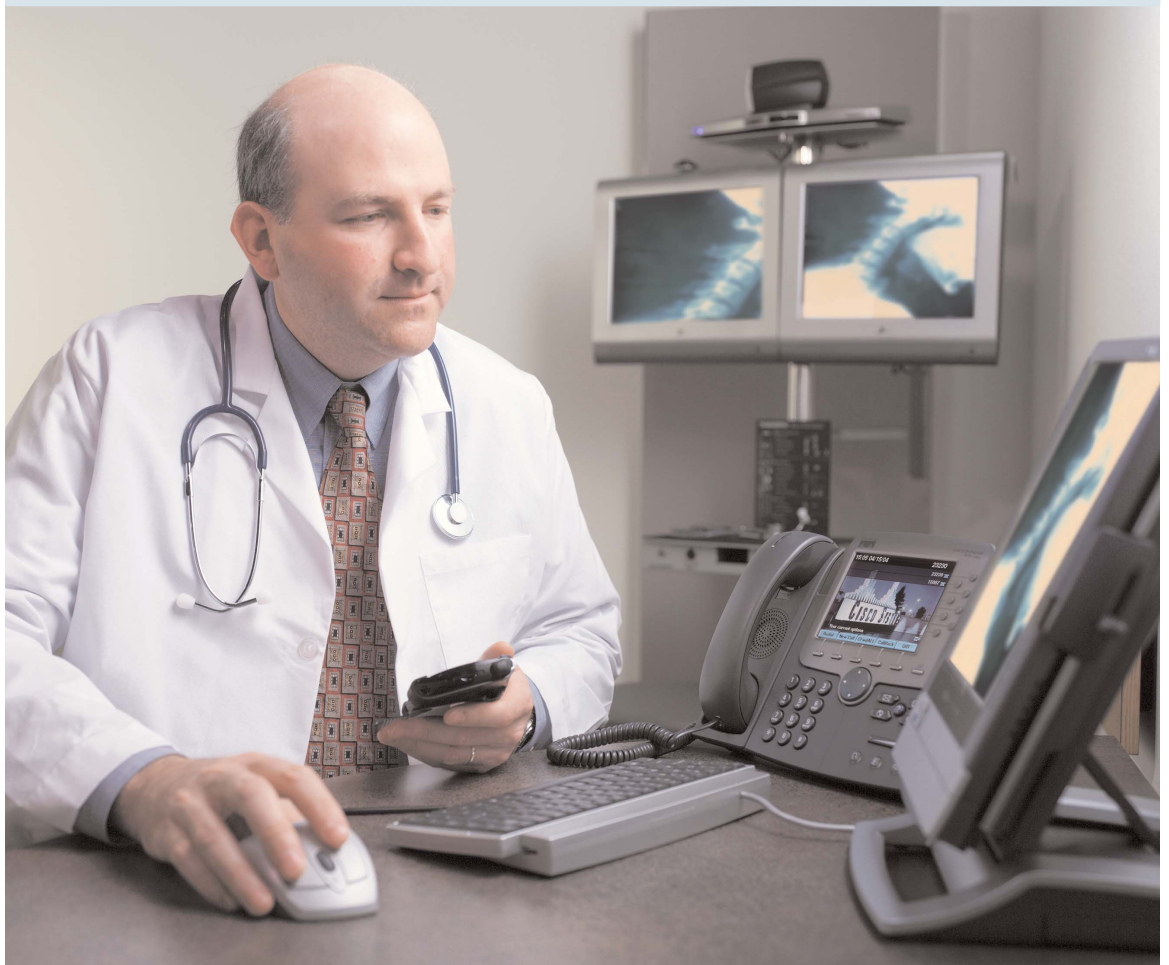


HEALTHGRID – A SUMMARY



A joint **White Paper** from the
Healthgrid Association
and Cisco Systems

CONTENTS

Chapter 1	An introduction to Healthgrids	5
Chapter 2	Creating a compelling business case for Healthgrid	9
Chapter 3	Opportunity 1 - Medical imaging and image processing	12
Chapter 4	Opportunity 2 - Modelling the human body for therapy planning and computer assisted interventions	14
Chapter 5	Opportunity 3 - Grid-enabled pharmaceutical R&D	17
Chapter 6	Opportunity 4 - Grids for epidemiological studies	19
Chapter 7	Opportunity 5 - Genomic medicine grid	22
Chapter 8	From Grid to Healthgrid: confidentiality and ethical issues	24
Chapter 9	Legal approaches of the Healthgrid technology	27

ABSTRACT

Abstract: This document summarises the critical information and discussion points raised in a detailed white paper that has been written by members of the Healthgrid Association. The Summary sets out for senior decision makers the concept, benefits and opportunities offered by applying newly emerging Grid technology in a number of different applications in healthcare. The aim of this document is to inform and enthuse healthcare leaders on the prospects offered by using open, dynamic and highly scalable access to data, information and analysis power - to improve the quality of patient care, ease access to care, and reduce the cost of delivering healthcare. The white paper also sets out a balanced view of the implementation and ethical issues involved in adopting grid-style solutions in the particularly sensitive world of healthcare - security, confidentiality, liability and ownership of data.

A list of authors and section coordinators / reviewers for the full white paper is available in Appendix 1.

The full white paper can be downloaded from <http://whitepaper.healthgrid.org>; the full white paper has a complete reference guide, omitted here.

EXECUTIVE SUMMARY - HEALTHGRIDS

Most health systems face major economic and capacity challenges to maintain a sufficient quality of care in the face of the growing demands of an ageing population and the increasingly sophisticated treatments available. Add to this the desire to improve access to new care methods, and the challenge of delivering care becomes significant. In an attempt to meet these demands, health systems have increasingly looked at deploying information technology to scale resources, to reduce queues, to avoid errors and to provide modern treatments into remote communities, for example.

However, the modernization process faces significant challenges:

- **Connecting and understanding patient records across organization structures and even national borders**
- **Ensuring that information is secure and those accessing it are authenticated**
- **Discovering trustworthy sources of information for comparison**
- **Handling a huge volume of data, especially that involved in genetic medicine for instance**
- **Applying traditional information networks and technology into healthcare.**

However a new weapon, Grid technology, is becoming available to clinicians that addresses several of the high priority issues they face.

A Healthgrid is an innovative use of emerging information technology to support broad access to rapid, cost-effective and high quality healthcare. Grid computing aims at the provision of a global ICT infrastructure that will enable a coordinated, flexible, and secure sharing of diverse resources, including computers, applications, data, storage, networks, and scientific instruments across dynamic and geographically dispersed organizations and communities (sometimes known as Virtual Organizations). Grid technologies promise to change the way organizations tackle complex problems by offering unprecedented opportunities for resource sharing and collaboration. Just as the World Wide Web transformed the way we exchange information, the Grid concept takes parallel and distributed computing to the next level, providing a unified, resilient, and transparent infrastructure, available on demand, in order to solve increasingly complex problems.

In particular the areas of healthcare provision and research that can be beneficially affected by Grid technology include:

- Medical imaging and image processing
- Modelling the human body for therapy planning
- Pharmaceutical Research and Development
- Epidemiological studies
- Genomic research and treatment development

In all these areas, Grid technology can either significantly reduce the cost or time to produce results and evidence, or even provide resources that are able to deliver services that cannot be economically delivered using conventional networked information systems. There are some significant hurdles to overcome before Healthgrid applications become commonplace - including ethical, confidentiality and working practice issues that are extremely challenging.

Two versions of the white paper have been produced; this, the shorter version, is aimed at those wishing to understand the high level issues, benefits and challenges offered. A fuller white paper version is available for those wishing to delve deeper.

1. AN INTRODUCTION TO HEALTHGRIDS

A Healthgrid is an innovative use of information technology to support broad access to rapid, cost-effective and high quality healthcare. As such, Healthgrids are an emerging aspect of a much more developed subject - eHealth. Before examining the potential for Healthgrid technology to address important challenges in healthcare, it is important to understand eHealth itself.

eHealth - new capabilities to care, a new industry

eHealth describes the use of Information and Communication Technologies (ICT) to develop an intelligent, connected environment that enables ubiquitous management of citizens' healthcare, assists health professionals in coping with major collaborative challenges or integrates the advances in health knowledge into clinical practice.

Many eHealth applications have been developed for dealing with information management and the procedural challenges of current healthcare. eHealth is not only a good strategy for improving healthcare quality, but also a good business. The eHealth or Health Telematics sector is becoming the third industrial pillar of healthcare after the pharmaceutical and the medical imaging devices industries. It is estimated that the health expenditure on ICT systems and services will rise from 1% to 5% by 2010; there are more than 1,500 healthcare sites on the Internet today; eHealth retailers, such as online pharmacies predict revenues ranging up to \$348B (US) by the year 2004.

In particular, "evidence-based medicine" requires medical decision-making to be founded on sound knowledge of the patient combined with peer-reviewed scientific evidence, rather than the more traditional informed guesswork, craft knowledge and personal skill. Furthermore, there is a pressing need to move away from paper-based patient information to digital records that can move effortlessly with the patient as they journey through care - connecting that journey for patient, family and clinicians alike. Countries in the EU are investing heavily to establish electronic patient record systems. Technically there are problems of standardisation and ensuring that systems are developed that interface through common 'languages' to enable the sharing of information. Technology to secure the information can also be complex and expensive to deploy.

Security is a very important issue. Personal data (any piece of information in which its owner can be identified, either directly or in combination with information that is available or may be available) is confidential, so access to the information must be restricted to authorized and authenticated persons, and data must be encrypted to guarantee its confidentiality and integrity. Indeed, electronic archiving of personal data is strictly regulated by European and national laws. If this were not enough, healthcare practice works around the clock, and thus requires "always-on" applications.

Medical information is also huge and dispersed. Large resources are needed to store patient records comprising images, bio-signals, plain text, videos, photographs or other forms of computerised data. Moreover, healthcare provision structure is distributed and information is not consolidated among hospitals, primary care and casualty departments. Linking federated databases requires computing effort and complex structures. Medical information is far from 'standard'; formats are usually mutually incompatible; standards are incomplete, and even where adopted they are also locally adapted to suit individual manufacturers.

Emerging technology - Grids

The Healthgrid white paper, summarised only in this version, describes the current status of Grid applied to eHealth opportunities, analyses the challenges, lessons learnt so far, mid-term developments and a path forward towards the successful deployment of Grid technologies for eHealth and ultimately the creation of a Healthgrid. The first item must be a more detailed description of Grid technology and its particular application in Health.

Grid computing aims at the provision of a global ICT infrastructure that will enable a coordinated, flexible, and secure sharing of diverse resources, including computers, applications, data, storage, networks, and scientific instruments across dynamic and geographically dispersed organizations and communities (sometimes known as Virtual Organizations). Grid technologies promise to change the way organizations tackle complex problems by offering unprecedented opportunities for resource sharing and collaboration. Just as the World Wide Web transformed the way we exchange information, the Grid concept takes parallel and distributed computing to the next level, providing a unified, resilient, and transparent infrastructure, available on demand, in order to solve increasingly complex problems.

Innovative research projects are using Grid technology to integrate bio-medical knowledge, make advances in imaging, and develop improved diagnostic and treatment tools. Such research suggests that Grid technology-based systems can make a significant contribution to the goals of accessing many different sources of

medical data, usually geographically distributed, and raising the availability of computer-based tools that can extract knowledge from that data - key requirements of high quality standard healthcare provision. Benefits of improved access are raised to a new level, not merely enhanced by integration over a Grid.

Although Grid technologies have clear potential for many applications (those demanding computing or storage power, dealing with geographically distributed information or requiring ubiquitous access), the take up of Grid is slow. Reasons for this are the lack of adequate infrastructure, lack of users' confidence and, most frequently, the shortage of applications - which is where 'healthcare' comes in. A Healthgrid should be an environment where data of medical interest can be stored, processed and made easily available to the different healthcare participants: researchers, physicians, healthcare centres and administrations, and in the longer term citizens. If such an infrastructure were to offer all necessary guarantees in terms of security, respect for ethics and observance of regulations, it would allow the association of post-genomic information and medical data and open up the possibility for individualized healthcare.

Grids may be classified into:

- Computational Grids - the goal of a computational Grid is to create a virtual supercomputer, which dynamically aggregates the power of a large number of individual computers in order to provide a platform for advanced high-performance and/or high-throughput applications that could not be tackled by a single system.
- Data/Information/Knowledge Grids - focus on the sharing of vast quantities of data; Information and knowledge Grids extend the capabilities of data Grids by providing support for data categorization, information discovery, ontologies, and knowledge sharing and reuse.
- Collaborative Grids - establish a virtual environment which enables geographically dispersed individuals or groups of people to cooperate as they pursue common goals. Collaborative Grid technologies also enable the realization of virtual laboratories or the remote control and management of equipment, sensors and instruments.

From the original experiments investigating possibilities offered by broadband networks, Grid technologies have entered into a phase where production capabilities are available, e.g. NASA's Information Power Grid, CERN's DataGrid, and NSF's TeraGrid. However, the vision of large scale resource sharing has not yet become a reality in many areas. This may be attributed to the lack of commonly accepted standards and to the diversity and fragmentation of available Grid middleware, tools and services.

Future developments of Grid technologies will be characterized by a full adoption of the service-oriented paradigm and Web Services technologies, a complete virtualization of resources and services, and the increased utilization of semantic information and ontologies (cf. Semantic Grid). Significant efforts will have to be undertaken in order to provide appropriate high-level tools and environments that hide the complexity and reduce the costs of Grid application development. The availability and adoption of advanced security standards, support for quality of service measures and the establishment of associated Grid business models and processes will be pre-requisites for large scale adoption of Grid technologies.

Healthgrids

Healthgrids are Grid infrastructures comprising applications, services or middleware components that deal with the specific problems arising in the processing of biomedical data. Resources in Healthgrids are databases, computing power, medical expertise and even medical devices. Healthgrids are thus closely related to eHealth.

Although the ultimate goal for eHealth in Europe would be the creation of a single Healthgrid, i.e. a Grid comprising all eHealth resources, naturally including security and authorisation features to handle subsidiarity of independent nodes of the Healthgrid, the development path will mostly likely include a set of specific Healthgrids with perhaps rudimentary inter-Grid interaction / interoperational capabilities. The following issues have been identified as key features of Healthgrids:

- **Data ownership & security:** Healthgrids are more closely related to data, but many hospitals are reluctant to let the information flow outside the hospital bounds. For a large-scale deployment of Healthgrids, and thus for opening an attractive business, it is important to leverage security up to a trustworthy level of confidence that could release a generalized access to data from the outside. Biomedical information must be carefully managed to avoid privacy leakages. Failure on privacy in biomedical personal information causes irreparable damage, since there is no way to revert to the previous situation. Secure transmission must be complemented with secure storage, higher security mechanisms that could avoid malicious users granting unauthorized access to part of the Grid to be able to decrypt and visualize personal data.
- **Developing new tools:** Distributed databases and data mining capabilities, and their management, are important tools for many biomedical applications in fields such as epidemiology, drug design or even diagnosis. Expert systems services running on the Grid must be able to dig into large distributed databases extracting the knowledge that can lead to the early detection of

new sources of diseases, risk populations, evolution of diseases or suitable proteins to fight against specific diseases.

- **"Always On" medical applications:** Robustness and 'fault tolerance' of Grids is well suited to the needs for 'always on' medical applications. Grid technologies can ease the access to replicated resources and information, just requiring the user to have a permanent Internet connection.
- **Creating and observing standards:** Compliance with medical information standards is necessary for accessing large databases. There are many consolidated and emerging standards that must be taken into account. Complex and multimedia information, including images, signals, videos, etc., is clearly a target for Grid and is more sensitive to data formats.

Today most of the Grid applications for health meet a demand for manipulating large amounts of data to give better understanding of diseases or the design of medical devices - in areas such as radiotherapy, cranio-facial surgery, medical imaging or neurosurgery. As an example, breast cancer imaging has been the focus of several successful Grid projects. The efforts have been concentrated on improving the sensitivity and specificity of breast cancer screening programmes.

Person-centric Grid for health approaches are also explored later in this paper. The main aim of this approach is to make the information available to the whole health community (patient, relatives, physicians, nursery...), considering access rights and language limitations.

Grid development

Grid technology is a moving target. The frenetic evolution of platforms and versions makes it difficult to progress applications beyond research to a production stage. The main concern of industry is to define and exploit business models on the Grid. But industry needs more stability and standardisation on Grid infrastructures before they can develop the business models. Critical factors that must be solved include:

- Grid middleware lacks reliable and comprehensive accounting services; these should clearly identify providers, consumers and resource usage of a wide range of heterogeneous resources, owned and shared by different entities.
- Grid security and privacy models are not adequate for deploying applications that can be certified by end users and health authorities.
- Reliable benchmarking must be performed to certify that the components can perform to the quality of service and robustness requirements of healthcare applications. Middleware certification is even more important in healthcare applications, since legal and ethical factors are implicated and it has the potential to impact patient morbidity and mortality.

- Current software licensing models conflict with Grid deployment; they usually prevent software use in Grid environments in which the computers and the users are not clearly defined. New licence models will appear with the development of new business models. Until then, it is better to focus on the exploitation of current application licence or public licence software.
- Before developing business-relevant applications, there is a clear need for a production infrastructure in which applications can be run. Many services can be implemented and tested and deployed for validation. Validation that any healthcare application requires for its exploitation can be performed on such a platform, although final exploitation can be performed on separate resources.

Solving these problems could lead to effective, wide-scale use of Grid technology by industry in healthcare. The business case for doing so is discussed later, however the prime sources of benefit are the consolidation of resources that Grid-sharing allows, the efficient provision of large peak computing, knowledge and analytical resources, and reducing the cost of research leading to the discovery of easy-to-use, affordable, effective drugs for diseases.

There are two main potential early-adopter industries for Healthgrid - pharmaceuticals and medical information technologies. The challenges in both derive from the same sources as those listed above for the general deployment of Grids, though the issues of security and standards are even more critical. These are exacerbated in the areas of genomics and personal health data. Later, this paper discusses in detail the opportunities and issues arising from some of the exciting application areas for Healthgrid.

2. CREATING A COMPELLING BUSINESS CASE FOR HEALTHGRID

Although both healthcare in general, and the use of IT to support the development of effective treatment, delivery and management of healthcare, are top priorities in many countries, there are many competing areas of investment. The benefits of using even basic IT to provide high quality information and decision support to clinicians and patients are intuitively very significant. In other industries - airlines, automotive, banking, defence, and manufacturing - IT has underpinned productivity, quality, security and improved product performance for many years. However, progress in even basic IT has been patchy and slow in the healthcare industry; there are few high quality, well documented business cases with results and very few for IT implementation at large scale. There are even fewer cases that demonstrate the benefits of dramatically new IT technologies (like Grid) or in innovative areas of healthcare such as genetics, imaging, or bioinformatics. Therefore in applying for funding and prioritisation of resources to continue to develop Healthgrid applications, it is vital that a clear and highly compelling business case is created that acts on all the benefits levers of healthcare.

The advent of Healthgrid applications, even at the research stage, coincides with a crucial period of investment and experimentation in IT for healthcare. The main drivers for this shift in the pace and levels of investment include:

- Increased understanding of the impact of medical errors on patient safety and the resulting cost of care. IT's basic value proposition includes the ability to regulate processes and scale information "written" once to many uses and contexts;
- Demand for healthcare outstripping resources at all levels, driven by an ageing population in most countries, living longer but with access to an increasingly sophisticated armoury of tests, surgical interventions, medications etc. IT has the power both to add to the armoury of clinical tools and to reduce costs through efficient operation with reduced numbers of process steps, wasted activity (tests, prescribing unnecessarily etc.) and better utilisation of disparate resources

The coincidence of growing capability in Grid technology with this increase in investment has its drawbacks. First, there are many strategic and investment plans being made at local, regional and national levels that take no account of emerging technologies like Grid. Even if the first truly useful Healthgrid applications will not be ready for several years, this is within the planning and budgeting horizons of the public sector. Second, as IT is introduced into everyday healthcare, custom and practice are changing on how care is delivered. Such change in the clinical world is very significant - for instance, rationalising the outpatients' process to a single series of steps, supported by sharing of electronic data in all hospitals within a region, is a considerable change. Overlaying such serious changes with the completely new capabilities of Grid will simply add to the challenges. And in healthcare, change can take time to embed - a recent study in the USA showed an average 17 year delay in adopting widely proven practices in healthcare.

Nor are Healthgrid technologies being anticipated in the many eHealth strategies being created around Europe. Very few senior health managers in Europe understand the potential or the practicalities of Healthgrid; in general they are certainly not embedding their strategies with even link points to take advantage of Grid in the future. The risk therefore is that it will be even harder than it should be to take advantage of Healthgrid capabilities over the next 5-10 years - unless the potential is understood quickly and strategies adapted accordingly.

Quality, access and cost provide the means to measure success. If the business case for Healthgrid is to be demonstrated conclusively, it must be articulated in terms that senior health managers can understand. One suggestion, based on the work of the European Commission eHealth Unit is to define the benefits across three categories, specifically the impact on:

- Raising the quality of care. Here factors include the ability to make faster decisions or interventions; fewer medical errors; more informed decisions or diagnoses.
- Improving the access of patients to care. Sources of benefit might include the extension of specialised diagnostic and therapeutic procedures to large numbers of patients through increased capacity; the provision of new tests or diagnoses that simply could not be made using traditional approaches (at reasonable cost).
- Reducing the cost of care. A complex issue for Healthgrid since it is an emerging technology creating opportunities for new procedures and test for instance that may actually add to the short term budgets; however there may be sources of benefit from such short term investments against the long term reductions in cost of care as disease is identified earlier and prevented for instance.

It is important to recognise that rarely do these three factors appear independently - for instance it may be that improving the access to care via new tests also impacts the long term cost of treating either chronic diseases or immediate palliative care.

Casting the benefits of Healthgrid applications against these three factors has a great advantage in creating compelling business cases for senior health management - and politicians - because it allows them to see the benefit in the terms that they are managing day to day healthcare outcomes and budgets against. Creating such resonance is critical to gaining priority and share of resources / budgets.

Why Healthgrid?

Not only does the modern healthcare management team have many choices for investment of their time and money in traditional sources of patient care improvement but they also have a bewildering array of IT support that can be purchased. So why, in such an already complex, packed marketplace, should the relatively new, often untried, Grid technologies be given any priority at all? Of course, not all Healthcare informatics problems will be suitable for a Grid solution. There are problems where the advantages of Grid approaches will outweigh the potential drawbacks of a relatively untried and new technology. The characteristics of clinical problems that could have significant advantage from distributed computing-type solutions include:

- Analyses that require dynamically assembled data-sets and investigation routines; for instance genetic-related investigations where the initial analysis may raise the need for further data sets to be added to give better, more representative results from analysis.
- Processes of analysis and data assembly that cross organisational boundaries, where the ability to distribute both the data and analysis without recall to the normal "data process" flows is key. Again medical research and future patient-centric analyses are probably two areas where the utility of the Grid will be highest.
- Huge scale analysis, that requires a scalable infrastructure to deal with the potentially massive quantities of data both to be assembled and analysed. This leads us again to imaging and genetic analysis as potential opportunities.
- Dynamic grouping of healthcare professionals for review / analysis of diagnosis or research results, such that different "expert teams" can be assembled without a formal organisation structure (indeed across organisation structures). Feedback from clinicians on existing Grid health projects indicates a strong need to enhance collaboration on a daily basis between communities, removing their reliance on conferences to achieve this.
- Further benefits may be realised through the pooling of resources, whether it be the sharing of training cases to enable smaller clinics to benefit from the knowledge available in larger hospitals, or the sharing of compute resource to reduce the local investment on IT.

Therefore, in summary, there seems to be an advantage available from using Grid approaches where the clinical problem requires a scaleable, flexible infrastructure that can work across normal organisation and process boundaries.

Barriers to economic, rapid implementation

While there may be some very serious advantages from applying Healthgrid technologies to suitable problems, there remain significant barriers to implementation. They can be summarised under three main headings:

- **Governance and accountability.** On many levels, the Healthgrid does not match current governance models and tried and tested processes. As one example, research conducted using Grid approaches does not necessarily have the same degree of independent scrutiny and open accountability to which traditional research, reviewed by peers, is routinely subjected. In fact, the very nature of dynamically assembled, self-discoverable data sets and analyses means that such scrutiny is probably impossible, although it may be linked to the trend towards online publication of research results. Secondly, the entire area of trust (particularly in data) is critical to the wide-spread acceptance of Grid approaches in health. This trust issue ranges from building diagnoses or clinical evidence on data collected, maintained and shared by organisations or individuals outside of the originator's span of control; to the acceptance of Grid applications being shared across organisations' infrastructures
- **Quality of service and speed.** With any distributed system, where the infrastructure (computing devices, data stores, networks) is not all under a single span of control, the issue of the availability of resources, and the maintenance / reliability of such resources, is critical. Add to this reliability issue the potential contention for resources that massive data manipulation could experience, and the quality of service (guaranteed speed of response) could be frequently compromised. There are approaches for managing this problem, but most increase the cost or require heavy structured governance processes
- **Incomplete business model and technologies.** Much of the Grid technology has only been applied in research fields where human lives do not directly depend on it or the decisions made on its output. Before life-critical applications can be trusted, many more examples, pilots and controlled trials will be necessary. Whilst there have been significant advances in standards for the integration of healthcare systems (IHE), it is evident that further work is needed in order to take this to the dimension of the big "joined-up healthcare" approach.

The Healthgrid is potentially a significant addition to the armoury of tools health professionals and researchers can use to increase quality and access and reduce the cost of healthcare. However, significant progress is required on the governance, quality of service and operational models for Grid technology before it can become a widespread tool in daily use.

3. OPPORTUNITY 1 - MEDICAL IMAGING AND IMAGE PROCESSING

Among the broad range of opportunities where Grid technology could offer a significant benefit to healthcare, imaging presents one of the most compelling cases. Medical diagnosis and intervention increasingly rely upon images, of which there is a growing range available to the clinician: X-ray (increasingly digital, though still overwhelmingly film-based), ultrasound, MRI, CT, PET scans etc. This trend will increase as high bandwidth systems for picture archiving and communications are installed in large numbers of hospitals (currently, primarily in large teaching hospitals). More than patient records, medical images represent a far larger volume of medical data.

However, medical images are not sufficient by themselves: they need to be interpreted and analysed in the context of the patient's medical record, which may be regarded as part of the metadata associated with the images. Medical data are used in diagnosis, continuing care, and therapy planning - which means that images need to be available in multiple locations as the patient journeys through care.

Acquisition and diagnosis

There are a number of factors that make patient management, including providing accurate diagnosis based on medical images, particularly difficult. Medical data are naturally distributed over a number of acquisition sites. Physicians as a rule have no way to access all the medical records across all of their patients - and often a single image or series of images does not help the clinician form the full diagnosis. Images are high volume data, especially if they are complex (such as 3-D images, time sequences, multiple imaging protocols) and require annotation; clinically and epidemiologically significant signs are subtle, with interpretation depending on many variables, such as, patient age, diet, lifestyle and clinical history, image acquisition parameters, and anatomical/physiological variations.

'Picture Archiving and Communication Systems' (PACS) deployed in hospitals today address some of the challenges related to medical data management. However they suffer many limitations:

- Often they are disconnected from the 'Radiological Information System' (RIS) carrying the medical records.
- They are often proprietary solutions of medical imaging companies and no open standards exist to ease communication between different PACS.
- They are usually limited to data management inside one health unit (one hospital or at best a federation of hospitals) and are not scalable on a national or international scale.

Manipulating medical data on a large scale also raises the problems of security and confidentiality of personal data. Grid technologies are expected to ease the design of distributed medical information systems in a secure environment. Although Grids cannot by themselves resolve the problem of heterogeneity in data formats and communication protocols, they are expected to motivate the establishment of standards in this field.

Medical data storage and archiving

Medical images represent enormous amounts of data: a single image can range from a few megabytes to one gigabyte or more. The total amount of digital images produced in Europe thus probably exceeds 1000 petabytes each year. The legal aspects concerning medical data archiving vary from country to country in the European Union but the actual trend is towards long term archiving of medical data (about 20 years for any data, up to 70 years for some specific data) and to make the patient the owner of his or her own data.

Grids provide support for distributed mass storage of data. Several Grid middlewares propose distributed and transparent file systems aggregating many storage resources to offer extensive storage capacity. Several aspects of Grids that are still under investigation concern the implementation of data access control and security of data. While remaining internal to the hospitals, data security problems are relatively easy to solve. However enabling data exchange between hospitals over wide area networks makes this matter much more complex. Medical data should always be considered as sensitive in general and identifying data should remain strictly confidential. In particular this means that data should only be accessible by authorised users (for sensitive data) or accredited users (for identifying data), often excluding service providers and system managers. Encryption (and thus anonymisation) of data on disk and during network transmission is therefore mandatory; the access to decryption keys being strictly controlled.

Medical Image processing

Computerised medical image analysis algorithms have been developed for two decades or so. The aim is to assist clinicians to handle the amount of data by providing reliable and reproducible assistance to diagnosis and therapy. Indeed, the manual processing of 3D images is very fastidious and often error prone. Moreover, 3D medical image interpretation requires mental reconstruction on the part of the physician and is subject to large inter-operator variations.

Grid technologies will enable image processing communities to share common datasets for algorithm comparison and validation. They will offer an access to large processing power suited to processing full datasets in reasonable time, compatible with the needs for experiencing new algorithms. They will also ease the sharing of algorithms developed by different research groups thus encouraging comparative studies. For all these reasons, Grid technologies are expected to boost the production of medical image analysis algorithms and to facilitate their quality improvement.

A further advantage is that Grid technologies will make the comparison of large numbers of images much more efficient. This comparison, whether comparing a series of images taken over time for the same patient, or comparing patients with similar conditions, is a complex process; Grid technology can provide faster more efficient means of finding and comparing the large amounts of data involved.

4. OPPORTUNITY 2 - MODELLING THE HUMAN BODY FOR THERAPY PLANNING AND COMPUTER ASSISTED INTERVENTIONS

Beyond medical data acquisition and analysis, modelling of the human body enables specific medical treatments. The key distinguishing factor compared with image processing or image reconstruction in the same applications arena is the use of computational methods for predictive purposes - providing physically accurate information (within modelling accuracy) that is not included in medical images themselves.

Enormous progress has been made in recent years (aided by the increases in performance of computing platforms) and numerical modelling is now able to provide realistic (and validated) predictions of very complex phenomena. However, there is a real need for the continued development of numerical modelling and simulation technology to address the future challenges of multi-scale, multi-physics problems that arise naturally and automatically in virtual human modelling.

Given the complexity and the computing cost of most human body models, Grid technologies are a good candidate to face computation challenges arising in this area.

Atlases - mapping the human body digitally

Atlases have long been used in medicine for anatomy and physiology studies. For centuries, atlases have been produced manually by experts from their knowledge of the human body. Atlases attempt to provide a 'standard' description of the human body or parts of it. They are very dependent on the designer and have been incrementally refined with the progress of medicine. They tend to be general and hardly take into account infrequent parameters.

With the advent of digital images and image registration algorithms, the production of digital atlases has become possible. Digital atlases are assembled by registering large training sets in a common frame and averaging the registered images by different means. Digital atlases prove to be much easier to produce than manual atlases. They have been very successful and have led to significant research progress, especially in the domain of brain imaging. The production of atlases requires the availability of training datasets large enough to be statistically representative and of sufficient computational power to accomplish the atlas computations. Grid technologies promise to cover both aspects and should therefore boost the production of anatomical and functional atlases of the human body. Given a wide scale medical information system and considerable computing power, one can even imagine producing on-the-fly individualised atlases. For example a physician may want to study the brain of a 50 year-old male subject to multiple sclerosis; he could ask for the production of an atlas from all the data available on other patients with matching criteria. Such an individualised atlas would prove to be much more specific and precise than a generic atlas.

Simulating the human body

Providing a simulation of the human body has many quality and cost of care advantages, both in research and patient treatment. A number of initiatives (such as the Virtual Human project and the Living human Project) are attempting to produce simulations of the human body, but are challenged by the amount of data and modelling that can be reliably accomplished using conventional computational techniques and databases.

Many areas of development in numerical human modelling are already at the stage where they can be used by medical researchers as investigative tools into causes of medical problems and effectiveness of treatment procedures. Research into cardiovascular disease in particular is an area where high performance computing simulation software is widely used, for example to improve understanding of processes leading to illness or to failure of implants such as artificial heart-valves or stents.

The interest of the Grid approach is to provide services to medical or clinical users, removing any need for them to have to handle the details of the computing systems or simulation methods. Grid technologies are also required to provide high-bandwidth to large collections of coarse-grained, distributed, non-textual, multidimensional time-varying resources. Web services technologies are required to cope with the dynamic aspects of a digital library that provides, not only data, but also simulation services, collaborative work services, interactive visualisation services, etc.

Broadening the term "medical supplier" to include pharmaceutical industries, the acceptance of the potential benefits of using numerical simulation tools (i.e. actual use or willingness to investigate use) is already well established within the R&D divisions of companies. For large companies, Grid offers the possibility to deploy simulation software across their own distributed resources. There are also established SME's supplying services and consultancy based on numerical simulation. Future Grid developments will allow them to enter into virtual organisations with their customers (including controlled access to data sources) and to have access to external computational resources when needed.

Planning and executing therapy

Many human body models have been developed for therapy planning. Examples of numerical simulation used by health practitioners include radio-surgery/radio-therapy planning, electromagnetic source localisation (an inverse procedure to identify areas of disorder within the brain based on external EEG/MEG measurements), reconstructive maxillo-facial surgery etc. Today, most developments are in the transition between research use and clinical use. Grid technology can be used to provide access to appropriate computational services and deliver these to medical users. The Healthgrid requirements in such cases are for larger scale deployment studies allowing evaluation of a larger range of requirements, including local deployment aspects, and practical experience with production Grid use. The major challenges will be to ensure that services can be delivered into the user's workplace in an appropriate, ergonomic manner and that security, policy and legal constraints related to the use of patient data are fulfilled. Some medical applications such as surgery simulation are more demanding and require real-time computations. Real-time is a challenging problem for Grid infrastructures today. Although Grids can provide additional computing power, deporting computations on remote resources is often done at the cost of an initialisation pay-off that can be rather important (from minutes to hours in common batch oriented scheduling systems). To empower real-time applications, a Grid middleware will need to ensure immediate execution of real-time code. Strong network requirements are also dictated by real-time constraints. Grid services dealing with jobs as sensitive as surgery simulation and computer assisted intervention should also take into account advance reservation of resources and emergency situations: the requested computation and networking resources must be allocated when the surgery starts and it should be possible to submit prioritised jobs in case of emergency with resource requisitions if needed.

Example Application - A Grid scenario for radiotherapy planning and treatment

From a technology point of view, radiotherapy is a highly complex procedure, involving a variety of computational operations for data gathering, processing and control. The modularity of the treatment process and the need of large data sets from different sources and nature (physics, mathematics, bio-statistics, biology, and medicine) make it a privileged candidate for Healthgrid applications.

In an enlarged Europe sharing data, expertise and computational resources will be a significant factor for successful cost containment and improved access to a high overall quality of care in radiotherapy. Healthgrid is an ideal tool for harmonising the cancer treatment as well as providing a common base for research collaboration.

Presently patients are treated with standardised radiation doses. Gene profiling may enable an individualised adjustment of the dose so as to achieve tumour control in patients with low radiosensitivity and avoid severe side effects in patients with above average sensitivity to radiation. In a first step a Grid structure should allow research groups, each focusing on different molecular mechanisms, to access data in the distributed infrastructure for comparative studies. As a possible next step users should be able to submit the results of predictive tests for analysis to a shared software and expert platform for radiosensitivity grading.

A similar approach can be followed for other aspects clinical decision making such as the assessment of a tumour's capacity for metastatic spread. For rapidly metastasising tumours, systemic (chemotherapy) treatment needs to be associated to the locally delivered radiotherapy. New tests now under development, predicting on the basis of gene profiling which tumours are most likely to metastasise, can make 60% of the chemotherapy currently administered e.g. for breast cancer, redundant. However, it takes a highly specialised team to interpret the results of these tests correctly. Grid-supported consultation of libraries of gene profiles or, alternatively, tele-consulting services offer excellent prospects also in these cases.

Tissue density provided by CT scanning is still needed to calculate the dose delivered by photon and electron beams. To define the planning target volume (PTV) and organs-at-risk (OAR), new imaging modalities based on MR-imaging, MR-spectroscopy and PET are far superior and become a requirement for high-precision high-dose radiotherapy. In contrast to CT scanning, the latter imaging modalities are available only in reference centres for reasons of cost and expertise. To secure access for all patients to optimal imaging for radiotherapy planning, the coordinating centre could perform a Grid-mediated selection of an imaging centre, and the resulting complementary image acquisitions could be sent back through the Grid. To reproduce the patient positioning and perform the complementary imaging in treatment-relevant conditions, the patient-individual immobilisation devices could be physically sent to the imaging centre. Alternatively, a retrospective registration Grid service could be used to realign all the images in the relevant coordinate system.

Many tools have been developed for computer-aided definition of PTV and OAR including anatomical atlases that can be warped to the patient-individual anatomy. A Grid could make such tools and their upgrades in due time available to all groups involved in PTV and OAR definition. Nodes on the Grid that provide expert help for patient-related problems in defining PTV and OAR are needed.

Accuracy of Monte Carlo (MC) dose computation is exceptional, provided the computing power is sufficient to allow for enough runs to reduce the statistical noise. The Grid is a natural alternative to costly parallel computers. In this way, MC dose computations could become standard for radiotherapy quality assurance (QA), planning, and plan optimisation years before individual departments could afford a local investment that is capable to support MC. Requirements needed for such deployment include the existence of a service level agreement between the departments and the Grid providers by which the Grid level of performances in terms of security, stability and response time is guaranteed.

Each delivery centre manages the commissioning of its own treatment units and incorporates both mechanical-physical and dosimetric parameters, including uncertainty flags, into an identity card that is accessible through the Grid. This identity card will allow treatment-planning providers and computation services to establish, refine or fit their computational model of the linear accelerator. The identity card also contains the reference data so that periodical quality assurance (QA) procedures could make sure that the machine performs accordingly. One might expect that the cooperation through the Grid between QA providers and delivery centres will streamline the QA procedures and harmonise the identity cards over the different accelerator types.

The quality assurance of the treatment can also benefit from the Grid, even if it is patient specific: once a treatment plan has been designed, some locations are selected to measure the dose level in a physical phantom that replaces the patient during the first treatment session. In parallel, the coordinating centre consults the Grid for an independent dose computation service to compute the dose in the same set of points in the phantom. The comparison of the measured dose to the computed fractional dose is performed automatically at the delivery centre and will be submitted to the coordinating centre. In case of violation of tolerances, the treatment plan will be recomputed in patient and phantom by a second dose computation service in the Grid. Alternatively, the coordinating centre may consult the Grid for a virtual treatment at another delivery centre.

5. OPPORTUNITY 3 - GRID-ENABLED PHARMACEUTICAL R&D

The pharmaceutical ('pharma') R&D enterprise presents unique challenges for information technologists and computer scientists. The diversity and complexity of the information required to arrive at well-founded decisions based on both scientific and business criteria is remarkable and well-recognised in the industry. The decisions can form the basis for multi-year, multi-person, multi-million Euro investments and can create new scientific territory and intellectual property. Thus all aspects of managing, sharing and understanding this information is critical to the R&D process and subject to substantial investment and exploration of new informatics approaches.

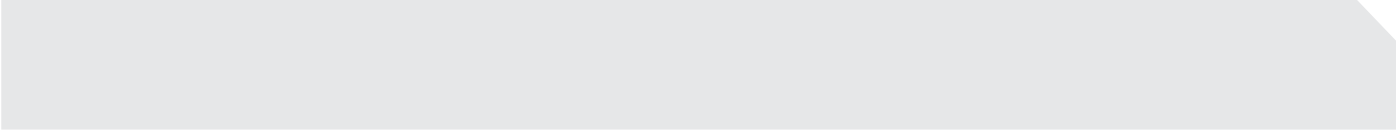
Pharma R&D information includes a large variety of scientific data as well as sources of critical organisational information, such as project and financial management data and competitor intelligence information. This data is often held in a unique format. Examples are images, models, sequences, full text scientific reports, records of prescriptions and physician encounter re-imbursments. These sources of information consist of internal proprietary, external commercial and open-source data.

The problems range from knowledge-representation and integration, to distributed systems search and access control, to data mining and knowledge management, to real-time modelling and simulations, to algorithm development and computational complexity.

Grid technology holds out the promise of more effective means to manage information and enhance knowledge-based processes in just the sort of environment that is well established in pharma R&D.

A pharmaceutical Grid should be a shared *in silico* resource to guarantee and preserve knowledge in the areas of discovery, development, manufacturing, marketing and sales of new drug therapies. It would cover three dimensions:

- a resource that provides extremely large CPU power to perform computing intense tasks in a transparent way by means of an automated job submission and distribution facility;
- a resource that provides transparent and secure access to storage and archiving of large amounts of data in an automated and self-organized mode; and
- a resource that connects, analyses and structures data and information in a transparent mode according to pre-defined rules (science or business process based).



Pharmaceutical Grids open the door to cheaper and faster drug development. Pharmaceutical Grids should enable parallel processes in drug development, away from the traditional approach where target discovery, target validation, lead discovery, lead optimization and transition to development take on average 12 years. These parallel processes would take advantage of *in silico* science platforms for target identification and validation, compounds screening and optimization, clinical trial simulation for detection of deficiencies in drug absorption, distribution, metabolism and elimination, including any harmful interactions between metabolites.

For competitive and intellectual property protection reasons, pharmaceutical Grids will predominantly be private enterprise-wide internal Grids with strict control and standards. At least this will most likely be the case in the near-term while R&D organisations explore and become comfortable with this technology and its potential.

6. OPPORTUNITY 4 - GRIDS FOR EPIDEMIOLOGICAL STUDIES

Conventional epidemiology requires extensive collections of data concerning populations, health and disease patterns, as well as environmental factors such as diet, climate and social conditions. A study may focus on a particular region or a particular outbreak, or it may take as its theme the epidemiology of a condition across a wide area. The range of data required will, therefore, vary with the type of study, but certain elements persist: a degree of trust in the data is essential, so its 'provenance' has to be assured and the standards of clinical practice under which it was obtained have to be above a certain threshold. Where the data has been gathered under different clinical regimes, it must be possible to establish their semantic equivalence, to ensure that aggregation or comparison of datasets is legitimate. Ethical issues may also arise if data collected in the first place in the course of individual healthcare is to be used for research.

The analysis of aggregated data requires the construction of complex models and the use of sophisticated statistical tools. This has necessitated collaboration between physicians and statisticians, and the rise of epidemiology as a discipline. The impact of genomic analysis will extend the kinds of variable under study and the range of expertise to be applied.

The technology to allow federation of databases stored locally in hospitals has existed for some time. It is possible for these databases to be queried for epidemiological purposes while preserving patient anonymity. Such distributed queries may be managed and supervised by the hospitals with primary responsibility for the data, ensuring compliance with ethical and legal regulatory frameworks. Nonetheless, the political difficulties inherent in the integration of information systems are well known and this has plainly not happened to the degree that is possible despite major government efforts.

Examples of epidemiology Grids are:

- Genetic epidemiology Grids for the identification of genes involved in complex diseases
- Statistical studies: work on populations of patients. One example is the tracking of resistance to therapeutic agents. This is most notable in relation to antibiotic resistance in common bacteria in nosocomial and community settings
- Drug assessment: drug impact evaluation through population analysis
- Pathology follow-up: pathologies evolution in longitudinal studies
- Grids for humanitarian development: Grid technology opens new perspectives for preparation and follow-up of medical missions in developing countries as well as support to local medical centres in terms of tele-consulting, tele-diagnosis, patient follow-up and e-learning.

Grids overcome the problems of the mere integration of databases. They can enforce the interoperability of tools and analysis services and they may also enforce common standards and semantic clarity about database content and tool input / output. Indeed, the Grid-based federation of retrieval systems provides a significant alternative to federation of databases. We may not see the latter for quite some time: federation of databases requires - in case the databases should be interoperable - clear semantics and standards based on conventions about semantics. Attempts to use semantics-based mediators have not been particularly successful so far. One road to go for the integration of medical data would be to adopt Grid strategies for data integration developed for bioinformatics. If such an approach is adopted, the cost and effort of establishing completely new databases in the field of clinical research / genetic epidemiology would be significantly limited, thus paving the way for smooth and rapid implementation of first demonstrators.

Opportunities in Image-oriented Epidemiology

Patient management (diagnosis, treatment, continuing care, post-treatment assessment) is rarely straightforward; but there are a number of factors that make patient management based on medical images particularly difficult. Often very large quantities of data, with complex structure, are involved (such as 3-D images, time sequences, multiple imaging protocols). In most cases, no single imaging modality suffices, since there are many parameters that affect the appearance of an image and because clinically and epidemiologically significant signs are subtle. Among the many relevant factors are patient age, diet, lifestyle and clinical history, image acquisition parameters, and anatomical and physiological variations. Thus any database of images developed at a single site - no matter how large - is unlikely to contain a large enough set of exemplars in response to any given query to be statistically significant. Overcoming this problem implies constructing a very large, federated database, while controlling for statistical biases such as lifestyle and diet almost certainly leads to a database that must transcend national boundaries. Realizing such a geographically distributed (pan-European) database necessitates so-called Grid technology, and the construction of a prototype would push emerging Grid technology to its limits.

Building population-based datasets

A European Grid for genetic epidemiology would open completely new perspectives for gathering data on large populations and - as a consequence - would allow stratification of large cohorts for large scale European genetic epidemiology studies. One possible problem that can be foreseen in this context is that there are regional, legal and cultural differences that may obstruct the building of pan-European, population-based datasets. As a

consequence, it would be necessary to complement any type of Healthgrid activity that could possibly encounter problems of this type by research activities in the field of ethical, legal, and cultural aspects that might impact future Healthgrids.

The current situation in Europe is quite heterogeneous. Initiatives to build large population-based datasets have been started in Iceland, the UK, and in one Baltic state, Estonia. These national initiatives are driven by different rationales. In how far commercial aspects will interfere with the goals of a pan-European initiative to build population-based datasets remains unclear, however, it is clear that large population-based datasets (and associated sample collections) are not only interesting for basic science but also for the pharmaceutical industry.

Even though there could be problems of privacy as discussed above, the chances that come with large scale studies and pan-European population-based datasets will exceed the risks of potential abuse of genetic information. Currently, genetic epidemiology studies suffer from low numbers of samples, inconsistent acquisition of bio-parameters and complex genetics.

Built on population-based datasets statistical studies on the influence of allelic predisposition, behavioural aspects, nutrition habits, regional or national healthcare management and many other parameters will be possible. A central task for a Grid project for genetic epidemiology would be to enable and to promote interoperability of statistical analysis tools. Similar to initiatives in the field of systems biology an exchange service for statistical models based on a common understanding and classification scheme of statistical approaches would be needed. A point to start with would be a "tool box" of statistical models including relevant meta-information on algorithms, modelling strategies and constraints, application scenarios and possible equivalence or variations of statistical models. As a Grid service this tool box would allow easy exchange of methods and improve interoperability of statistical models and data mining capabilities on the side of the users of the genetic epidemiology Grid.

The study of pathologies follow-up would include information related to regular hospital visits, home-care monitoring of signs and symptoms, recording of interventions and drug effects, environmental issues, etc. However, these studies are usually fragmented and are certainly not uniform, so that they cannot result in common conclusions. One can see this issue from two points of view: (a) how pathology follow-up or the setup of clinical trials can be supported, and (b) how the results of clinical trials can be better utilized in a manner that feeds medical knowledge and clinical practice.

Drug Assessment

On the biological and pharmacological side, the determination of allelic frequencies of drug target genes in European population is one important application field for a genetic epidemiology Grid with large population-based datasets. A second application scenario concerns aspects of drug safety; again an aspect that is highly relevant for public health and the pharmaceutical industry. Adverse drug effects depend - amongst other factors - on cytochrome gene polymorphisms and one of the first large scale studies done on a Grid for genetic epidemiology could be a project on cytochrome allelic variability in patients with e.g. resistance to a certain class of compounds.

A third application scenario could strive to unravel the genetic basis of drug insensitivity which is not based on allelic variation of acute response detoxification genes. As an example we might think of the insensitivity of a huge percentage of multiple sclerosis patients to treatment with interferons. Another scenario would concern the insensitivity of a significant portion of the European population to treatment with glucocorticoids.

From the Grid research perspective, drug related epidemiological studies require a tight integration of knowledge coming from heterogeneous disciplines, namely pharmacology and genetics. Currently, knowledge representations (ontologies) for pharmacology are missing by and large; we therefore expect that a Grid on genetic epidemiology that addresses aspects of drug action will have to include an activity on ontology construction for the domain of pharmacology. A "pharmacology-ontology" would also help to formalise and to standardise the description of clinical parameters measured in the course of large scale studies. As drug assessment comprises all aspects of pharmacodynamics, special attention will have to be paid to appropriate representation of dynamic processes (e.g. changes of drug serum concentration over time); sharing of mathematical / statistical models for the analysis of drug effects and drug stability will be essential for pan-European studies.

Genetic Epidemiology

The genetic basis of complex diseases provides a real challenge to any information system for genetic epidemiology and for a Grid for genetic epidemiology in particular. Complex diseases are characterized by the high number of parameters to be recorded and by an "intrinsic fuzziness" in the conceptual definition of clinical phenotypes (e.g. "depression"). Genetic epidemiology studies in this field require much larger cohorts of patients to produce significant results. The use of Grid technology could help with the acquisition, analysis and modelling of data at, say, national or pan-european levels.

Genetic epidemiology studies try to establish links between genetic variation (polymorphisms / allelic variance) and individual risk that have an impact on the quality of life (including major diseases).

Genetic epidemiology studies have a direct impact on decisions on health quality standards, disease management and risk assessment. Unfortunately, the prospects of Europe-wide genetic epidemiology studies have not yet been fully explored; even though significant effort has been undertaken in the course of national projects, data from different studies are not easily comparable and data access is very limited.

7. OPPORTUNITY 5 - GENOMIC MEDICINE GRID

The full realisation of the concept of 'genomic medicine', in which genomics and proteomics are used to empower healthcare, requires the integration of knowledge from worlds traditionally apart, biology and medicine. To harness effectively the wealth of information available in research centres and care facilities, a new framework of computer methods and tools must be in place, bridging medical and bioinformatics. In such an approach, all levels of information (from the molecule to the population, through the cell, the tissue, the organ and the patient), as well as the most appropriate techniques and methods, would be used.

The completion of the Human Genome Project (HGP) is seen for medicine as a source of new knowledge to understand the relationships between the structure of human genes, environmental factors and physio-pathological processes. In the post-genomic era, the possibility of studying all the genes, all the proteins, or a high number of mutations in human cells, paves the way to new research possibilities not feasible until now. These will help us understand the molecular basis of complex diseases, thus facilitating the development of new diagnostic and therapeutic solutions.

Biomedical resources are often structurally unrelated, though the contents they hold are strongly and semantically connected. Bringing together such knowledge is a complex task, since it is difficult to automatically uncover the semantic associations. Recent advanced in Grid technology are in line with semantic integration needs:

- Web Services allow the searching, calling and execution of distributed services, and could be used to implement some basic biomedical services and applications.
- Grid-based databases and metadata management systems provide a secure, efficient, and automatic data source management.
- Support for Virtual Organization clusters through basic Grid services, such as security, and tools and platforms for cooperation.

Semantic integration involves both modelling and technical aspects. While the former allow for the deployment of high level semantic services and applications, the latter can enhance performance and efficiency on distributed and Grid environments.

Biomedical Grids

Expected contributions of Grid technology to genomic medicine realization include:

- Computational genomics and proteomics in the identification of genes and proteins; technologies to store large amounts of phenotype, genotype and prototype.
- Support for the development of clinical tests.
- Provide personalized healthcare services following: (a) the genetic profile of each patient, (b) epidemiological studies, (c) heredity, (d) statistical analysis results, and (e) clinical observations.
- Development of models and digital simulations of cells and diseases.
- Providing tools to support physicians' training, and improve biomedical knowledge management.
- Molecular imaging.
- Developing tools that support clinical decision making.
- Integrating databases and knowledge between the clinical world and that of genomic research.

8. FROM GRID TO HEALTHGRID: CONFIDENTIALITY AND ETHICAL ISSUES

Healthcare information systems hold, process and transmit sensitive personal data. This implies a need for strict confidentiality and enforcement of privacy protection. These requirements have not been dealt with in early applications of Grid technology; in domains such as high energy physics, elementary particles need no privacy nor do physicists want to shield information from each other.

Medical data is usually very sensitive and even where it is used for the benefit of the community, this information is prone to abuse. Thus there is appropriate concern about the proper treatment of sensitive data. Incidents of abuse have been reported in the media and the public reaction has always shown that the threat is considered genuine. There would be a strong impact on society if banks, insurance companies and employers, were able to access healthcare data about their customers or employees, revealing past, current and (with genomics) probable future health. Indeed, abuse of medical data can affect all of us, as at some point in life practically everyone has to make a loan, insurance or job application.

Grid Security

Since the very start, the Grid community has put a lot of effort into the design of security measures. Authentication and authorisation mechanisms are the main point of focus of these developments, as they are the most basic of security measures. Integration at the level of the lower middleware allows security mechanisms to be uniform and interoperable. However, implementation is still at an early stage. It is important to realise that the further development of security technology is key to the acceptance of the Healthgrid concept.

Current Grid security technology is sufficient to address computational problems in healthcare. But Healthgrid does not intend to restrict to the use of Grid technology for distributed computing only. Eventually, Healthgrid should offer a generic platform for all eHealth actors. Sharing of large amounts of distributed heterogeneous (on various levels) data is therefore an important point of attention.

It is clear that the linkage of several distributed data sources bound to a single individual on a data Grid opens up a range of privacy risks. The (virtual) federation of a large amount of personal medical data is not the only risk at hand. Grid technology will undoubtedly further stimulate the use of genomic data in research. However, this particular type of data has a number of specific characteristics related to privacy which are not found in any other type of (medical) information:

- Genetic data concerns not only individuals, but also their relatives. A person's consent to release his or her genetic information constitutes a de facto release of information about other individuals, i.e. his or her relatives. In the case of genomic medicine, there is a complex interaction between individual rights, collective requirements and legislation.
- Medical data deal with past and current health statuses of persons, whereas genetic information can also give indications about future health or disease susceptibility or other predispositions.
- An individual person's genotype is almost unique and stable, hence it can become the source of an increasing amount of information.
- The full extent of the information included in the genomic data is not known yet, hence it is difficult to assess the full extent of disclosure.
- Genomic data is easily wrongly interpreted by non-professionals, "susceptibility" to diseases can easily be mistaken with certainty of illness.

The above clearly indicates that the challenge at hand is the reconciliation of two seemingly conflicting objectives: on one hand, maximization of medical research productivity and efficiency in data handling; on the other, the protection of the human (privacy) rights.

A couple of basic approaches to safeguarding confidentiality have been identified in the past in healthcare practice. The first approach focuses on the creators and maintainers of the information, prohibiting them from disclosing the information to inappropriate parties. Basically, this comes down to the deployment of classical security measures (access control, authorisation). A Healthgrid initiative is ideal for the further development (and actual implementation) of Grid security technology, because of the strict requirements in healthcare. A first task within the Healthgrid context could thus be performing an in depth analysis of the new and specific risks and threats that arise.

An emerging technology that may assist in addressing confidentiality issues in Grid-based healthcare applications is 'privacy enhancing technologies' (PETs) - a way of eliminating or reducing personal data. Mainly used for protection of the privacy of persons involved in medical data collection, the goal of these PETs is to guarantee anonymity of data subjects while making information available for clinical practice and research. For the Healthgrid goal, access to large amounts of useful, personal information can be unlocked though the use of privacy protection techniques, mainly de-identification methods.

The Grid promises access to heterogeneous resources, which means that in a Healthgrid environment remote resources will store and process sensitive personal data. These resources should thus be

trusted by the end-user. But how can one know? Who can be the judge of "trustworthiness" of a Grid resource? A simple and straightforward solution is to use "closed" systems, which means that any resource in the Grid is well known and specified in advance. This however conflicts with the vision of dynamisms and the "ad hoc" nature of Grid technology.

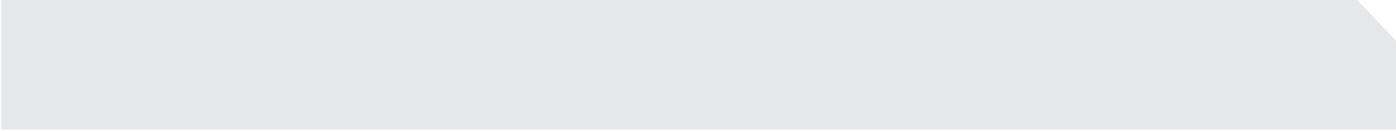
Solutions should rather be searched in the area of policy advertising and negotiation. Resources should be able to inform a candidate user on how the data dealt with will be treated, which policies are applied, what PETs are used, who can have access to the data, etc. These methods are sometimes referred to as not being genuine Privacy Enhancing Techniques, since, strictly speaking, they do not actually limit collection of personal identifiable data and do not give any guarantees about the actual processing. A resource can claim to adhere to strict rules, but in practice this cannot be verified.

The first steps in the direction of policy management have already been taken by Grid developers. The development of standards such as WS-Privacy, WS-Policy and EPAL (Enterprise Privacy Authorisation Language) is an effort in the right direction. However, implementation to date is rather limited, and the full possibilities of the technology will not be researched unless effort is spent here from the healthcare area - the main application domain. A Healthgrid would be the ideal environment where such PETs can be tested and further developed.

The above directly impacts typical Grid mechanisms such as data replication. Replication mechanisms automatically copy data on a resource in order to increase efficiency (e.g. to avoid transfer delays). With medical data, this might however not be allowed. The site on which the data will be replicated should at least be as trustworthy as the data source and should adhere to the same strict policies. A Healthgrid should be able to handle such cases autonomously in order not to lose its dynamic nature and efficiency.

A further example is delegation. Delegation of rights is fundamental in a Grid environment; however in the medical world, this is far from acceptable. If one passes on rights to others (resources), one becomes liable for actions performed on one's behalf. In a healthcare environment this has serious implications on liability. Restricted proxy certificates offer a path to a solution suitable for medical applications, but clearly need to be extended.

Policy management will be an important topic in Healthgrid, both for security (e.g. authorisation policies) and for data protection (privacy policies). A number of difficult problems arise in this context in relation to policy bridging, policy enforcement and assuring that a certain policy is followed.



Equally important and closely related to this subject, is the implementation of auditing mechanisms. All actions in a medical context should be logged in a trustworthy way. Non-repudiation combined with a legal framework could help solve liability issues in healthcare.

Next to the areas of interest mentioned in this text, several other healthcare needs for Grid applications exist which could be developed at, e.g., upper middleware level for the benefit of a large community within a Healthgrid context. Among these are encrypted storage for medical data (a far from obvious problem) and trustworthy federation of research databases - virtual federation of small "cells" of de-identified data (e.g. geographical area, hospital, and so on) can decrease the re-identification risk (by increasing the anonymity set). Finally a range of PETs which are well suited for distributed environments is emerging ('private information retrieval and storage, privacy preserving data mining, processing of encrypted data, and others). However the road to an advanced generic privacy preserving framework for eHealth is still long and littered with technical difficulties which must be tackled one at a time.

It is, however, clear that Grid technology can only be successful in a biomedical environment if the ethical guidelines and legal requirements are adequately met by technological solutions which are continuously evaluated and updated as new needs arise.

9. LEGAL APPROACHES OF THE HEALTHGRID TECHNOLOGY

The introduction of Grid technology in the healthcare sector might seem insignificant and in any event without any legal relevance. It would only concern a new computing technology contributing to the provision of healthcare services or to scientific research, mostly by providing huge computing and memory resources, possibly Internet-based. The first projects deal with medical imaging, medical tele-assistance, medical or pharmaceutical research, human genomic studies, creation of databases for therapeutic, scientific, statistical or epidemiological purposes.

However these projects are ruled by radically different legal contexts. Indeed, distinct legal rules govern the practice of medicine, scientific or pharmaceutical research, epidemiological studies, etc., even if those disciplines may contribute to medical progress. Hence there is no unique answer to the determination of the legal framework in which Healthgrid technology may be implemented and used. In reality, the answers are multiple and depend on the context of each project as well as on the relevant legal viewpoints. Healthgrid technology must conform to the legal context specific to each project.

Nevertheless describing the different legal contexts in which Healthgrid technology might be implemented is not sufficient. The adequacy of the legal context in relation to the characteristics of this technology should also be evaluated. In other words, one should question whether certain rules should not be adapted with respect to Healthgrid technology.

Healthgrid Technology Status

Technologies must frequently comply with precise technical norms in view of their legal utilisation. The same assertion is also valid for the healthcare sector. It is therefore important to define the content of the technical norms relevant to each project.

In this matter, some technical norms have been harmonized at an international or European level. It is useful to note that the European Committee for Standardization has issued a very interesting study entitled "European Standardization of Health Informatics - Results of the mandated work by CEN/TC 252" (CEN TC 251/N01-024 - 2001-06-17).

The European Union has also adopted several rules concerning medical devices, including a number of Council Directives:

- Council Directive 90/385/EEC of 20 June 1990 on the harmonisation of the laws of the Member States relating to active implantable medical devices.
- Council Directive 93/42/EEC of 14 June 1993 concerning medical devices.
- Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices.

The Council of Europe states that the improvement of human life quality and the respect of human rights should prevail when dealing with new technologies. It recommends in this regard that the precise evaluation of any technology should as much as possible rely on the following criteria (cf. Recommendation (90) 8 of 29 March 1990 on the impact of new technologies on health services, particularly primary healthcare):

- Validity of outputs;
- Validity of data capture;
- Ability to fit within the framework of primary healthcare;
- Social acceptability;
- Ethical acceptability;
- Professional acceptability;
- Reliability;
- Capacity for continuous assessment;
- Safety for providers; consumers and the environment;
- Cost effectiveness compared to older technologies;
- Availability of full information on the technology and experience in implementing it;
- Protection of confidentiality;
- Ability to be integrated smoothly into existing systems;
- Availability of adequate resources.

This evaluation should consist of appropriate studies giving conclusive results, and should be carried out prior to the general introduction of any new technology.

Personal Data Processing

Most of Healthgrid technology related projects imply personal data processing for therapeutic purposes or scientific research (e.g. medical imaging, tele-assistance, medical or scientific research, Human genomic studies, creation of Healthgrid databases).

However personal data processing is subject to numerous regulations. Indeed, these data are particularly sensitive and consequently require high protection. Furthermore, because of the therapeutic or scientific stakes, personal data processing must be reliable, or it may lead to medical errors or erroneous scientific works.

On the international level many norms govern personal data processing (including the processing of personal data related to health). Article 8 of the Convention for the Protection of Human Rights and Fundamental Freedoms is imperatively worth attention in that matter. Article 7 of the Charter of Fundamental Rights of the European Union similarly confirms everybody's right to privacy while Article 8 establishes the right to the protection of personal data. The Council of Europe has also issued important norms relative to personal data processing. The European Group on

Ethics has adopted an important opinion concerning the processing of personal data related to health (cf. Opinion of the European Group on Ethics in Science and New Technologies to the European Commission, Ethical issues of healthcare in the information society, n° 13, 30 July 1999). Similarly, the World Medical Association has issued several documents of interest to some Healthgrid projects. (The precise details of the guidance, ruling and directives are discussed at length in the full white paper).

National norms relative to personal data processing must comply with this international framework, although a certain margin is generally recognized to the member states in view of their local implementation. However, this may in itself cause a slight disparity in national norms, adding to the problem of those for which no international rules exist and upon which Member States are free to decide.

In any case it is of prime interest to qualify correctly the operations performed on personal data when using Healthgrid technology and to define the role of each person involved (healthcare practitioners, service providers, patient, etc.).

From a technical point of view, 'privacy enhancing technologies' (PETs) offer strong support to the security and the confidentiality of processed personal data. They aim to reduce the processing of personal data and to suggest appropriate measures to secure data processing.

Healthgrid Services

Some projects aim at providing services to healthcare professionals or to scientists. These services must be qualified according to the norms applicable to 'information society' services. An information society service is any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services.

'Information society services' also encompasses services consisting of the transmission of information via a communication network, in providing access to a communication network, or in hosting information provided by a recipient of the service. Activities which by their very nature cannot be carried out at a distance and by electronic means, such as medical advice requiring the physical examination of a patient, are not information society services.

Information society service providers are subject to prior authorization or some other requirement having equivalent effect (cf. Internal Market - Directive on Electronic Commerce). The service provider must therefore comply with a number of special rules when offering information society services.

This provision of services may result from a contractual relationship. The latter must be analysed on an individual basis in each project. In case of an international situation when providing information society services, one should carry out a preliminary examination of competent jurisdictions before defining the law applicable to the contractual obligations of the various parties.

Several international instruments can be mentioned in this regard:

- Convention on the law applicable to contractual obligations opened for signature in Rome on 19 June 1980 (80/934/EEC);
- Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures;
- Directive 2000/35/EC of the European Parliament and of the Council of 29 June 2000 on combating late payment in commercial transactions.

The End-User

The use of Healthgrid technology by healthcare professionals raises special questions. On one hand, is the end-user legally authorised to use the Healthgrid technology? Is the use of Healthgrid technology permitted in medical practice or in scientific research? The answer lies in the rules governing the professional activities of the end-user.

Concerning some projects, it is useful to be reminded that the European Union has adopted the Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the harmonization of the laws, regulations and administrative provisions of the member states relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use.

On the other hand, in case of medical tele-expertise, medical tele-consultancy, or medical tele-assistance, involving healthcare practitioners from different member states, the question is to establish if the healthcare practitioner in charge of the patient is legally authorised to seek the assistance of a foreign healthcare practitioner, and, if so, under which conditions.

At the same time, this foreign healthcare practitioner should also establish whether he or she is legally authorised to provide assistance to a healthcare practitioner located in another country.

Beyond the determination of the persons liable in case of medical accident or fault, one must define the status of the healthcare practitioner participating in the provision of healthcare in another member state, and the status of the healthcare practitioner having asked his assistance. This problem is far beyond the simple question of the equivalence of medical qualifications. In the same way, cooperation between healthcare practitioners within the same

member state or from different member states raises the very delicate question of the legal framework of this cooperation.

The Patient

Implicitly or explicitly all Healthgrid projects aim to contribute to medical progress as well as in its preventive and curative aspects. Hence the patient is at the very heart of the implementation of Healthgrid technology. The Council of Europe has asserted its interest in promoting the active participation of the patient in his or her own treatment (cf. Recommendation R (80) 4).

The legal qualification of the parties involved in the processing of the patient's personal data, including the location of the patient, is likely to highlight some tensions underlying the medical relationship.

Liability Issues

The question of the determination of the persons liable in case of medical accident or fault relating to the use of Healthgrid technology when providing healthcare to a patient, is crucial but delicate. In case of an international situation, the question is far more complex. With respect to this, one should take into account several factors which are not necessarily likely to be under complete control.

The first element of uncertainty results from the determination of the possible jurisdictions likely to know the case. With respect to this, the European Union has recently adopted the Council Regulation (EC) No 44/2001 of 22 December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters. Determining the jurisdiction is essential to identify which laws are applicable in each case.

(The full white paper discusses the application of specific rulings made by the European Union on determining jurisdiction and the application of special cases).

Intellectual Property Rights and Competition issues

The creation and the use of Healthgrid technologies may raise important intellectual property rights (IPR) questions. Indeed, Healthgrid technologies are sometimes created like patchworks. This poses the question of the IPR relative to the constitutive elements of such "patchwork".

The European Union has adopted several directives concerning IPR issues, listed in the full white paper. Usually projects aiming at implementing Healthgrid technology comprise several partners who have formed themselves into a consortium. Their behaviour must then comply with competition law (where issues of monopolistic positions, abuse of dominant position and concerted practices are controlled).

The Healthgrid Association & contact details

The goal of the Association is to:

- contribute to the structuring of the European Research Area for health, by, for example, favouring the use of Grid technologies;
- offer active support and guidelines for the extension of the association's members contribution to society;
- supply information and other services relevant to the members of the association;
- encourage co-operation between the members of the association, notably in the creation of effective networks of collaboration;
- create partnerships of benefit to both higher education and scientific research in the field of the health in its broadest sense, both within Europe and in the rest of the world.

www.healthgrid.org

Lead contact: yannick.legre@healthgrid.org

Cisco Systems Internet Business Solutions Group & contact details

As a long term trusted business and technology advisor, Cisco Internet Business Solutions Group (IBSG) helps customers maximize their return from technology investment. IBSG business and industry experts bring customers the latest industry trends and technological innovations, sharing Cisco and industry best practices. The group engages globally with Cisco's largest customers to help refine their business processes to increase productivity, reduce costs and create new revenue streams. IBSG offers business and internet expertise across seven vertical industries, including the public sector. It works with more than half the largest organisations across each industry and with all of the top ten global telecoms service providers.

www.cisco.com

Lead contact: kevin.dean@cisco.com

APPENDIX 1 - AUTHORS OF THE WHITE PAPER

Full white paper (available from <http://whitepaper.healthgrid.org>)

Chapter	Section authors	Coordinators	Reviewers
Chapt.1	I. Blanquer, V. Hernandez, E. Medico, N. Maglaveras, S. Benkner, G. Lonsdale	Ignacio Blanquer Espert, Vicente Hernandez, Guy Lonsdale	Nikolay Tverdokhlebov,
Chapt.2	K. Dean, S. Lloyd	Kevin Dean, Sharon Lloyd	Siegfried Benkner,
Chapt.3	K. Hassan, R. McClatchey, S. Miguet, J. Montagnat, X. Pennec	Richard McClatchey, Johan Montagnat, Mike Brady	Siegfried Benkner, Irina Strizh
Chapt.4	V. Breton, W. De Neve, C. De Wagter, G. Heeren, G. Lonsdale, L. Maigne, J. Montagnat, K. Nozaki, X. Pennec, M. Taillet	Johan Montagnat, Xavier Pennec	
Chapt.5	H. Bilofsky, R. Ziegler, M. Hofmann, V. Breton, C. Jones	Howard Bilofsky, Chris Jones	Irina Strizh
Chapt.6	M. Hofmann, T. Solomonides, N. Maglaveras	Martin Hofmann, Tony Solomonides	Clive Tristram, Irina Strizh,
Chapt.7	M. Cannataro, P. Veltri, G. Aloisio, S. Fiore, M. Mirto, N. Maglaveras, I. Chouvarda, V. Koutkias, A. Malousi, V. Lopez, I. Oliveira, J. P. Sanchez, F. Martin-Sanchez	Ilídio C. Oliveira, Juan Pedro Sanchez, Victoria López	Irina Strizh
Chapt.8	G. De Moor, B. Claerhout	Georges De Moor, Brecht Claerhout	
Chapt.9	J. A.M. Herveg	Jean A.M. Herveg	Yves Poulet

Summary White Paper edited by: Kevin Dean, Tony Solomonides



Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 317 7777
Fax: +65 317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the **Cisco.com Web site at www.cisco.com/go/offices.**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic
Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy
Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal
Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2004, Cisco Systems, Inc. All rights reserved. CCIP, the Cisco Arrow logo, the Cisco *Powered* Network mark, the Cisco Systems Verified logo, Cisco Unity, Follow Me Browsing, FormShare, iQ Breakthrough, iQ Expertise, iQ FastTrack, the iQ logo, iQ Net Readiness Scorecard, Networking Academy, ScriptShare, SMARTnet, TransPath, and Voice LAN are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, Discover All That's Possible, The Fastest Way to Increase Your Internet Quotient, and iQuick Study are service marks of Cisco Systems, Inc.; and Aironet, ASIST, BPX, Catalyst, CCDA, CCDP, CCIE, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, the Cisco IOS logo, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Empowering the Internet Generation, Enterprise/Solver, EtherChannel, EtherSwitch, Fast Step, GigaStack, Internet Quotient, IOS, IP/TV, LightStream, MGX, MICA, the Networkers logo, Network Registrar, *Packet*, PIX, Post-Routing, Pre-Routing, RateMUX, Registrar, SlideCast, StrataView Plus, Stratm, SwitchProbe, TeleRouter, and VCO are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0208R)