How Grids are Perceived in Healthcare and the Public Service Sector

Mikko Juhani PITKÄNEN [a,1], Xin ZHOU [b] Miika TUISKU [a] Tapio NIEMI [a]
Ville RYYNÄNEN [a] and Henning MÜLLER [b,c]
[a] *Helsinki Institute of Physics, Technology Programme, Finland*
[b] *Medical Informatics Service, University and Hospitals of Geneva, Switzerland*
[c] *University of Applied Sciences, Sierre, Switzerland*

Abstract.
Computational and data grids have been developed to manage large amounts of data produced in scientific collaborations. While the developed technologies could be employed to help the daily work in the public sector and in healthcare, their wide adaptation has not been seen, yet. In this paper we present a survey, which we conducted to find out how the decision makers and system specialists in the public service sector see the role of Grid technologies in their future work. The respondents of the survey work as decision makers and systems specialists in the healthcare and the public services domain in Switzerland.

Keywords. Grid, Survey, Public Sector

1. Introduction

The concept of Grid computing was introduced by Foster et al. already in the late 1990s [11] but the idea itself is even older. Foster and Kesselman defined the Grid as a system that:

1. coordinates distributed resources,
2. using standard, open, general-purpose protocols and interfaces,
3. to deliver nontrivial qualities of service.

The usage and the development of Grid computing tools has actively continued in numerous computing projects. For example the European Union has financed several large–scale Grid projects. The developed methods and tools have been applied in several fields of science. The application scenarios include e.g. high energy physics [15] and healthcare [4]. In these fields the Grid computing facilitates resource sharing amongst the participants of the scientific collaborations, thus enabling the collaborators to use, access, and share resources across the organizational boundaries. Public service in general have a larger number of desktop machines in common, that are currently not shared for any common goal. As they often use confidential data on these machines the motivation for any resource sharing is often low.

[1]Corresponding Author: Mikko Juhani Pitkanen, Helsinki Institute of Physics, Technology Programme, CERN/PH, CH-1211 Geneva, Switzerland, mikko.pitkanen@hip.fi

The software development in the Grid community has been guided largely by the needs to provide computational power for scientific collaborations. This need was answered by developing Grid middlewares to enable secure resource sharing for processing and storing data. The role of a Grid middleware is to control distributed resources in such a way that needs of the participating organizations can be satisfied without interruptions in a reliable manner. Today, Grid computing is widely adopted in the high energy physics community [15], it is used for resource sharing in bioinformatics[2], and data access in the medical domain [3].

The Service of Medical Informatics [4], is responsible for the design, development, and evaluation of an advanced clinical information system for the Geneva University Hospitals (HUG). The research group participates in the KnowARC[5] project, in which an extension of the advanced resource connector (ARC) middleware is produced to offer foundations for know–how sharing services for business and society. To better understand the needs for Grid technologies in the healthcare and the public services sector that share many legal and regulatory restrictions for information handling a survey was conducted.

We conducted the survey presented in this paper by interviewing systems specialists and decision makers in the Geneva area. The goal of the survey was to identify the current situation and the needs for Information Technology (IT) in diverse public sector organizations including the healthcare providers. The participants of the survey were chosen from diverse organizations comprising two large universities, a large university hospital and several public service sector IT providers such as the Geneva administration. For each participant the same set of questions was presented as a basis of the interview. The results were analyzed using qualitative research methods as the number of interviews remains small.

As an example on IT services on the public sector, we interviewed employee in the IT department of the state of Geneva. The characteristics of the Geneva state IT infrastructure can be described by the need to centralize all development and operation to the state's IT department and bring the benefits from economies of scale. The IT department treats other administration professions as internal clients. As such Geneva state has a large data center of 630 server machines, including mainframe and some 19'000 distributed desktop computers for employees. The challenges faced are mostly linked with the heterogeneity of Software and Hardware. There are 700 different applications to support. The requirements differ greatly if the client comes from education or if the client represents the police or justice department. The strategic initiatives set forth deal with server consolidation, putting services online (eGovernment) and pushing open standards, formats, and source code initiatives forward. There are several good application scenarios possible from pure Virtualisation and Enterprise Grids with service oriented architecture (SOA) to few classical monte carlo HPC scenarios for example in the Tax department.

Earlier work by Montagnat et al. [5] shows how Grid computing can efficiently analyze a database of medical images. In their work, the database is partitioned over a set of resources, the query image sent to the resources. Then, similar images are returned to the

[2] http://www.bioinfogrid.eu/

[3] http://dev.globus.org/wiki/Incubator/MEDICUS/

[4] http://www.sim.hcuge.ch/

[5] http://www.knowarc.eu/

user. The evaluation is based on using a set of resources within a single site, whereas our study looks on how heterogeneous resources from multiple administrative domains could be used to provide the computing power. One of projects that has managed to access medical data across organization borders is called Globus MEDICUS [7]. The project has designed and implemented an image sharing system that uses Grid tools for efficiently transferring medical images. The system guarantees sufficient security measures for sharing medical data across various administrative domains.

Computational Grids have also been used to harness large amounts of computational power to identify drug candidates for the influenza A virus by Lee et al. [6]. Similarly to our focus area, they employed a Grid infrastructure, which extended over diverse sites. Oliveira et al. [9] have shown how Grid computing can be employed to reduce the computation time in content–based image retrieval (CBIR) systems. The authors used the MyGrid architecture in a hospital network to distribute image analysis to computers on the network.

Breton et al. [8] have proposed use of Grid technology to transfer patient data over organizational boundaries. Their work discusses many of the challenges in distributing medical data over multiple sites. Sloot et al. [10] have presented the Grid–based ViroLab decision support system for various challenges in the medical domain.

2. Methods: Interview of System Specialists and Decision Makers

The experimental part of this study consists of a survey that was conducted in the Geneva area in late 2007. We interviewed information systems specialists and decision makers who work in public service sector and have good knowledge on the information technology. Since the interviewed persons were experts or part of the top management we expected the number of respondents to be fairly small. To gain information from such a small group of specialists, we used a qualitative survey as our method. The interview consisted of 35 questions, for which multiple answers for each question were allowed.

Content validity means that the questions of the survey really measure the subjects meant to be measured. In this survey, a Grid specialist designed the questionnaire to cover aspects that are important for the adaptation of the Grid technology. A medical informatics specialist further helped to focus the questions on the challenges often faced in the healthcare and public service sector. Moreover, literature on the challenges in the field was used for creating the questions, see e.g. [8]. When the survey was conducted, each of the respondents was given the same questions and the interviewer (a grid specialist) was always present when the questionnaire was filled. Assistance was given if some of the questions were unclear for the respondent.

The form of an interview can vary from a tightly structured or formal interview to an unstructured or open interview. A themed interview is the intermediate form of open and formal interviews. It is typical that the topics under discussion are planned and known, but the specific order and form of the questions are not. (Hirsijarvi et al. 2005) A themed interview is closer to an unstructured interview than a structured one. A themed interview is the most used qualitative research method in the research field of society and business economics [16]. This makes it as one of the most valid research methods for qualitative research.

There are some issues affecting a survey's reliability and validity. The results of this survey can not naturally be generalized because the random sample was regional.

However, the participating group of the survey were top professionals and the results are therefore well descriptive for the current situation of Grid technology acceptance/usage in the public service sector in the Geneva area in Western Switzerland [14].

3. Results

In this section we present the results of the survey together with the used questions. The first part of the survey consists of 12 questions, shown in Table 1. These questions were designed to get information on the IT knowledge of the respondent as well as his working environment. The left column of the table shows the question and the right column shows the responses. The number of the same responses is shown in parentheses *()*.

What is the size of your organization?	80-8000 employees, median 300, 500
How many years have you worked in your current environment	less than 1 (1), 1-3 (1), 3-5 (1), more than five (5)
How many years have you been using computers in your work?	more than five years (all)
What is the computer situation in your organization?	only in hospital less than a computer per employee (but servers used), otherwise all have personal computers and shared servers
Do you have a personal email for work or do you share email with your colleagues	personal, except also shared in organization that provides services
To which part does using a computer contribute to your daily tasks	25-50% (1), 50-75% (2), almost all (5)
In your work, do you need to call by phone or visit other people to perform computer operations for you, e.g. install software?	rarely (7+1), other people call me (1)
How is your work computer connected to the network (Internet)?	with limitations (1), unlimited (7)
Do you need to sign-in (username and password) to multiple services to perform your work tasks?	several passwords (5), single–sign–on (3), both (1)
Do you use computing devices other than desktop or mobile phone in your work?	digital badge (5), PDA (3), laptop (5)
Do you believe that the work in your office could be made easier by introducing new computer systems or services?	current is sufficient (4), too many computers already (1), more services needed (3), need for fast centrally managed servers (1)
Which of the following tools you use to collaborate with your colleagues?	email (8), skype (3), instant messaging (2), Intranet (7), Groupware (4), wiki (7), phone (all), shared folders (6)

Table 1. Questions and answers on respondent background and the computer as a tool.

As Table 1 shows, all of the respondents come from organizations that have more personnel than many private sector companies. The respondents have a long professional background on IT related tasks. The work tasks are mostly carried out using computers, except management duties performed in meetings. Most have non–limited Internet connectivity, which enables easy use of external services. Although some places have adapted a single–sign–on (SSO) system to reduce the number of authentication mechanisms, most of the users still need several passwords. One of the mentioned reasons was that external systems could not yet be integrated to the organizations' SSO systems. Moreover, the respondents used several digital devices as part of their daily work.

In general the amount of computing devices was seen sufficient, but some of the computers used for running services could be combined to run in a single centralized resource. Only the system administrators responded that there are too many computers and services and it is challenging to keep all systems up–to–date. Finally, several different group–communication applications were used to communicate and share documents. A part of the group–communication applications were proprietary solutions.

Security Related Questions

Grids use strong security mechanisms both to guarantee fair use of resources and to secure the information processed, stored, or accessed by the Grid middleware. For strong security, a public key infrastructure (PKI) is deployed to enable secure cross–organizational access. Once the PKI and the related security mechanisms are deployed, the benefits of the Grid security infrastructure could be extended to common office tasks such as digitally signing documents. The questions and answers in Table 2 show security challenges that the respondents face in their work.

Do you work with confidential documents using your computer?	yes(8)
When you need to get documents signed in your work by some authority, are they signed on paper or digitally?	always on paper (5), rarely digitally (3),always digitally (0)
Which of the following security concepts are you familiar with?	digital signature (8), integrity hashing (6), authentication (8), accounting (7), cryptography (7), group based access management (6), authorization (8)
When starting to use digital security systems do you find that the instruction was sufficient to make people realize the legal aspects?	sufficient instructions (4), more instructions on technology needed (1), more instructions on legal aspects needed (4)
Is there a person responsible for information confidentiality in your organization? Do you think the responsible has expertise on computer systems?	security policy and technology well understood (5), security policy is well understood, but IT not well (3), lack in understanding both IT and security policy (1)
If you have severe computer problems that prevent you from working, how long do you expect (are used) to wait until the problem gets fixed?	within one hour (4), several hours (2), even more than a day (2)
What of the following difficulties you think could be faced with the use of Grid in your organization?	very limiting security policy (4), people not used to work with technology (7)

Table 2. Questions and answers on security perspectives.

Table 2 shows that all of the respondents handle confidential documents in their work. The signing of documents is still mostly done on paper and only three of respondents had used digital signing. The respondents showed good general knowledge on the central digital security concepts used in the Grid world. The good knowledge on IT made the technical adaptation of new systems easy, but more information on legal aspects could have been provided. The respondents saw that in general the security technology and the related security policies were well understood. However, sometimes problems existed with understanding how the IT should be used to implement the chosen security policies. The availability of the systems was mostly high, but sometimes systems were still seen to have unacceptable down–time. Many organizations still had not used the Grid technolo-

gies and they saw that the technical challenges would be a major barrier of adaptation. Also half of the organizations had such limiting security policies that the deployment of the (Grid–like) multi–domain infrastructures was seen as problematic.

Questions Related to the Grid Technology and IT Services

Grids are used to share resources across organizational boundaries. However, in the public services sector many organizations store and manage highly confidential data, varying from medical health records to income and tax information or legal documents. The confidentiality often forces strong limitations on where the data can be processed and what services are allowed within the organization's network. The questions and answers in Table 3 show what could be the challenges when the Grid technology is used in the organizations of the respondents.

Table 3 shows that working with the respondents' computer systems is often dependent on the network services, which were more often hosted by the own organization rather than an external service operator. Most of the respondents were open to the idea of running a desktop Grid on the organization's computers. Most of the organizations had evaluated feasible or already purchased external computing services, whereas two respondents did not find it an interesting idea. The availability of the organization's computing resources during the past year was generally seen good. The possible benefits of Grid technology could mostly be achieved by getting access to computing resources. Specific databases and data storage was seen as an interesting service but less appealing than general computing. The respondents had good knowledge on some of the most general Grid concepts, indicating their familiarity with Grid computing. Most of the organizations could send processed data out and thus use the external service providers, whereas for two the idea was impossible. If a Grid system was taken into use, the largest challenges were seen in administering the security and IT of the system and educating the end–users. All of the respondents had previously discussed with Grid specialists and half of them had some real usage experience. Half of the respondents had donated their resources for desktop computing already, and some others were thinking to do so in the future. The virtualization technologies were used surprisingly widely in the organizations but this is a fairly recent development.

Computer System Investments and Education

Before conducting the survey we assumed that most of the interviewed persons had not actively used or deployed Grid technologies in their organizations. To find out possible non–technical barriers of adaptation we presented questions on the future plans of computing system investments and the adaptation challenges faced by the users.

Table 4 illustrates that the amount of money invested in computer systems has mostly stayed constant during the last year, and the constant investment trend will continue. Most of the respondents were from organizations that provide computing services for external users and organizations. The questions on education revealed that it often takes in the scale of several years before the new computer systems are found helpful. Often, when the computer systems were delivered sufficient help was provided. However, the lack of further sources of information was seen insufficient. This has led to the situation where people are reluctant to start using the new systems. Online tutorials were seen as

If the computer system does not work, is the corrective measure done usually at your desk, or somewhere else?	at my desk (3), external service operator (2), local machine room (5)
Do you think that other people could use computers/computing power of your office when they are not in use, if system provider would ensure security guarantees?	impossible because of security (1), can think about making an effort if we get enough resources (7), already allowed (1)
Do you think that the working environment would benefit from having the possibility to use stronger computing resources (e.g. computing cluster) outside your office?	already using (2), evaluated recently and found feasible (3), no additional resources needed (2)
Has the work in your organization been interrupted because of computer system problems during the last year?	no notable problems (1), sometimes short problems (6), office computers caused several interruptions (1)
What Grid services would you find most useful as a user?	access to computing resources (8), access to large data storages (5), secure and fine grained access to databases (6)
What following concepts are you familiar with?	parallel processing (6), cluster computing (8), job monitoring (7), grid job submission (4), distributed data management (6), batch processing (8), job brokering services (5), load balancing (6)
Can your data be sent outside your institution for processing? If not, how challenging would you judge changing the situation?	is being already sent (3), possible with sufficient security policies (4), impossible because of confidentiality (2)
What of the following you could think to be problematic in your organization with the use of the Grid?	naming person to: be responsible for security (3), install and update the systems (5), present organization in meetings (1), educate users (5)
What of the following experience do you have with Grids ?	read on the Web (2), attended a presentation (6), discussed with grid developers (8), used grid clients (3), installed grid services (1), used a parallel computing system(3)
What of the following is true?	familiar with distributed desktop computing (8), would like to contribute resources for such use (6), have already participated (4)
Virtualization, we have planned to use?	we run only Java services (2), we have installed virtual machines (6), use of virtualization has been discussed (1), I do not know (1)

Table 3. Questions and answers related to the Grid and service oriented technology.

the most desired source of information, and it was mentioned that the existence of tutorials is not enough but time has to be allocated for personnel to go through the material. Also expert help and online forums and helpdesks were seen as possible aid in the learning process.

4. Important Trends and Conclusions

In this paper describe a survey with the goal to find out more about the needs and problems of Grid technology in the public services and healthcare domain to reveal possible barriers of adaptation. Most of these institutions have a very large number of desktop computers that are most often centrally managed and could be reused in a local desktop

How has the amount of computer system investments developed within your organization?	during past year: decreased (5), stayed constant (3), next year: will increase (1), stay constant (5), will decrease (2)
Does your organization provide computer services to other organizations or people?	we provide: for other organizations (6), for individuals (3), do not provide services (2)
How long do you estimate that new computer systems had been in use before they start to be accepted and are found helpful in your organization?	within half a year (3), around one year (4), more than two years (3)
Is sufficient assistance and education delivered when new computer systems are brought into use in your office?	sufficient help is provided (5), lack of information is often problematic (3)
What kind of help would you like to have when new systems are introduced?	outside expert on site (6), online tutorials (8), online forums (5), help–desks (4)

Table 4. Questions and answers on computer systems investments and education.

grid system. On the other hand security issues are important as on many computers confidential documents are at least viewed. We have observed the following trends in our survey:

- *The important role of IT* in the work of the public service sector was a clear although not very surprising trend. All of the respondents had used their computers for a long time and the IT related tasks formed a major part of their daily activity. The IT was not only seen as a part of the daily work but also as an important part of group and collaboration activities.
- *Security challenges* in the public service domain arise from frequently working with confidential data. These vary from student records in universities, financial records in the tax department, to patient data handled in the hospital. The *digitalization* of the security mechanisms, e.g. by the means of employing digital signatures, has started but the wide adaptation is still under way as paper signatures are most often still required. We believe that many security mechanisms developed in Grid research [13] will also help to manage digital credentials in the public service sector. Moreover, the availability of the systems could be increased by adapting the high–availability technologies to the office computing systems.
- *Knowledge on Grid technologies* amongst the decision makers and the IT specialists in the public service sector is on a good level. The central concepts were well know and understood. Daily work involved using networked services, typical to the enterprise Grids implemented based on service oriented architecture (SOA) principles. Many of the respondents had also participated in Grid activities at some point by trying to use software related to Grid computing. Several had contributed resources to community computing projects, such as SETI@HOME or BOINC, or were planning to contribute for such use. Virtualization technology, which is being increasingly used within Grid projects, was well known amongst the respondents and is by used in the majority of the organization.
- *Adaptation of new technologies* went without major problems in most organizations. However, more education and resources such as time allocated for the learning was seen essential to make the new systems adapted more rapidly.

After all we were surprised about the generally high knowledge on most technologies related to grid computing, although grid computing itself was not in routine use in any of the institutions surveyed. Some institutions have first pilot experiments on using desktop

computers as a grid resource. Computing power is still the main reason for implementing grid computing and it seems possible to implement grid solutions based on the large number of desktop computer available in many organizations.

Goal of this survey for us was to obtain more knowledge on possible problems of grid computing in the public service sector. A second goal was also to inform the persons responsible for the network in the University and Hospitals of Geneva on the possibilities and requirements for reusing the existing computing infrastructure. Based on the survey we were able to continue discussions and work on a pilot installation of using standard hospital PCs with a virtual Linux machine as an intra–hospital grid. The most important aspect of this discussion was the real need of computing power for a specific application and the benefit of the diagnostic aid application for the institution itself.

References

[1] GRIDSTART Consortium, "Achievements of EU Grid Projects As funded by the EU's Fifth Framework Programme (FP5)", European Commission, (2005)

[2] Müller H., Squire D. M. G., Pun T., "Learning From User Behaviour in Image Retrieval: Application of the Market Basket Analysis, International Journal on Computer Vision, volume **56**(1-2), (2004) 65–77

[3] Müller H., Rosset A., Garcia A., Vallee J.–P., Geissbuhler A., "Benefits of content-based visual data access in Radiology", Radiographics, **25** (3), (2005) 849–858

[4] Müller H., Pitkänen M., Zhou X., Depeursinge A., Iavindrasana J., Geissbuhler A., "KnowARC: Enabling Grid networks for the biomedical research community", HealhGrid 2007, (2007) 261–269

[5] Montagnat J., Breton V., Magnin I., "Partitioning medical image databases for content–based queries on a grid" in Methods of Information in Medicine, **44**(2), (2005) 154–160

[6] Lee H.–C., Salzemann J., Jacq N., Chen H.–Y., Ho L.–Y., Merelli I., Milanesi L., Breton V., Lin S. C., Wu Y.–T., "Grid-enabled high-throughput in silico screening against influenza A neuraminidase", IEEE Trans Nanobioscience. **5**(4) (2005)

[7] Erberich S. G., Silverstein J. C., Chervenak A., Schuler R., Nelson M. D., Kesselman C., "Globus MEDICUS - Federation of DICOM Medical Imaging Devices into Healthcare Grids", Studies in Health Technology and Informatics, IOS Press, **126**, (2007) 269–278

[8] Breton V., Blanquer I., Hernandez V., Legre Y., Solomonidés, T., "Proposing a roadmap for Healthgrids", Stud Health Technol Inform, **120** (2006)

[9] Costa Oliveira M., Cirne, W., de Azevedo Marques P., "Towards applying content–based image retrieval in clinical routine", Future Generation Computer Systems, **23** (2007)

[10] Sloot P., Tirado–Ramos A., Altintas, I., Bibak M., and Boucher C., "From Molecules to man: Decision support in individualized e-health", IEEE Computer, **39**(11), (2006)

[11] Foster I., Kesselman C., "Globus: A Metacomputing Infrastructure Toolkit", The International Journal of Supercomputer Applications and High Performance Computing", **11**(2), (1997)

[12] Müller H., Michoux N., Bandon D., Geissbuhler A., "A review of content–based image retrieval systems in medical applications — clinical benefits and future directions", International Journal of Medical Informatics, **73**(1), (2004) 1–23

[13] Welch V., Siebenlist F., Foster I., Bresnahan J., Czajkowski K., Gawor J., Kesselman C., Meder S., Pearlman L., Tuecke S., "Security for Grid Services", Proceedings of IEEE HPDC-12 (2003)

[14] Hair J. F., Anderson R. H., Tatham R. L., Black W. C., "Multivariate data analysis", New Jersey, Prentice Hall (1998)

[15] Ellert M., Gronager M., Konstantinov A., Konya B., Lindemann J., Livenson I., Nielsen J. L., Niinimaki, M., Smirnova O., Waananen, A., "Advanced Resource Connector middleware for lightweight computational Grids", Future Generation Computer Systems, **23** (2007) 219–240

[16] Koskinen J., Lintinen H., Tilus T., Sivula H., Kankaanp I., Ahonen J. J., Juutilainen P. "Industrial case study of software maintenance evaluations". Proceedings of the IASTED Conference on Software Engineering (2006)