07-05-2010

Deliverable D.N4.3.1,1: New User Communities and Application Areas

Deliverable D.N4.3.1v1

Contractual Date:	30-03-2010
Actual Date:	07-05-2010
Grant Agreement No.:	238875
Activity:	NA4
Task Item:	T3
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	DANTE
Document Code:	GN3-10-096
Authors:	R. Hughes-Jones (DANTE) , D. Davies (DANTE), W. Singer (DANTE)

Abstract

This document reports on the work done by NA4 Task 3 in liaising with new and emerging user projects and communities with the aim to understand their computing and hence networking requirements, inform them of the connectivity and network services available in the GÉANT portfolio, and help them use the network to their best advantage.

Table of Contents

Table of Tables

Executive Summary

This document reports on the work done by NA4 Task 3 in liaising with new and existing user groups with the objective of understanding their networking requirements, informing them of the connectivity and network services available to them in the GÉANT portfolio, and helping them to use the network to the best advantage.

The user groups include projects or communities requiring access to e-Infrastructure resources, members of the academic and research community already familiar with the use of the academic network, and existing collaborations who are extending their work.

Two approaches were used to identify and liaise with the user groups and projects:

- Meetings with projects on the ESFRI (European Strategy Forum on Research Infrastructures) roadmap [ESFRI Roadmap] via workshops organised by DG Infso and DG Research; these meetings were attended by projects within an ESFRI Working Group or with a common discipline orientation
- One-to-one meetings between GÉANT and specific projects

Meeting the projects on the ESFRI roadmap via the ESFRI workshops provided a good opportunity to understand the individual projects and the common e-Infrastructure requirements of their field. Consulting with ESFRI projects from the same field or Working Group (brought together in the discipline-orientated meetings) made it easier to understand their network requirements. In many cases more detailed discussions are required to fully understand their networking requirements. These requirements will develop as the computing models become more fully defined in each project.

In dealing with specific projects we have developed the concept of a “lead” NREN, who, together with DANTE, will take on a coordinating role when interacting with projects and communities distributed over several countries. This approach has worked well and allowed the specific NREN knowledge of the local environment or user community to be fully utilised.

NA4 Task 3 continues to work with specific projects to facilitate collaborative performance testing, proof of concept demonstrations and assistance with the formulation of the future networking strategy. It seems likely that several case studies can be made with these emerging communities.

1 Introduction

The globalisation of research and technology has become an increasing factor in the last two decades. In 2000 the concept of the European Research Area (ERA) was developed. ERA has as an objective the European organisation and funding of research co-operatively with European Union (EU) member states. The EU Research Framework Programme is explicitly designed to support the creation of ERA. Co-operative, distributed research and development relies on advanced ICT and therefore it is a goal of the GN3 project (GÉANT) to better understand the potential telecommunications and networking requirements of researchers. It is an objective of GN3 to liaise with user projects and communities, in order to understand their networking needs. At the same time, GN3 can inform projects of the connectivity and network services available in the GÉANT portfolio, and provide technical support to enable projects to use the GÉANT network to the best advantage.

Existing projects, as well as potentially new projects, are brought together by ESFRI, the European Strategy Forum on Research Infrastructures [ESFRI]. ESFRI provides a forum for facilitating multilateral initiatives, leading to the better use and development of research infrastructures, at EU and international level. ESFRI organises ad-hoc working groups to analyse topical issues and to report to the forum. The ESFRI roadmap identifies discipline areas and European projects within these areas, as well as the Research Infrastructures to support them [1].

1.1 Meetings with ESFRI projects via workshops

GÉANT made use of the framework of ESFRI to engage with the pan-European projects. While this has the advantage of providing a single forum in which to meet with project representatives, one limitation of this approach is that the disciplines covered by ESFRI (ranging from Social Sciences and Humanities through to Biological and medical sciences) are a heterogeneous group with different networking requirements.

A group approach to understanding the networking needs of pan-European projects (by getting them all together in a single workshop) is likely to be less effective, due to the different mixture of needs, requirements and experience, as well as their different expectations. Although all projects have requirements for both

[1] The ESFRI roadmap is a strategic roadmap for Europe, approved by the Council of the European Union, which describes the scientific needs for Research Infrastructures for the next 10-20 years. The roadmap identifies vital new European research infrastructures of different size and scope, covering all scientific areas.

computing and networking, some have an understanding of what they need, while others are unclear about the benefits, capabilities and limitations of the technology.

Bearing these facts in mind, rather than trying to address the ESFRI projects as a homogenous group, via a generic workshop, a more targeted approach has been taken, based on meetings with each of the following fields or discipline orientations within ESFRI:

- Social Sciences and Humanities
- Environmental Sciences
- Energy
- Biological and Medical Sciences
- Materials and Analytical Facilities
- Physical Sciences and Engineering

Involvement of the Directorate General for Research

The EC Directorate General for Research (DG-Research), together with the member states, is responsible for developing the European research area and for funding research infrastructures. In order to encourage convergence on planning among the different areas of research, DG-Research, with support from DG-Infso, has organised a number of discipline-oriented workshops, which GÉANT has attended, as described below.

Involvement of EEF members

It was decided to address the project's needs co-operatively with computing and grids initiatives so that a broader picture of their e-infrastructure needs could be developed. The European e-Infrastructures Forum (EEF) is a loose alliance of networks, supercomputing and grids activities that is jointly addressing the e-infrastructure needs of ESFRI projects. Members of EEF with expertise in these e-infrastructure areas were invited to attend the workshops being arranged with the ESFRI disciplines. For more information on the involvement with EEF, see Appendix C.

To meet the objectives of NA4 Task 3, GÉANT, as a member of EEF, has used the ESFRI workshops to gain a better understanding of the network needs of groups of projects. This level of segmentation, addressing collections of projects within their field or discipline, has been a more effective approach to identifying user requirements.

1.2 Meetings with individual projects

In addition to the workshops mentioned above, technical discussions between GÉANT and individual projects were held to gather their specific networking requirements.

These two approaches — meetings with projects in the discipline areas organised under the banner of ESFRI and DG-Research, as well as one-to-one meetings with individual projects — represent the efforts that GÉANT has used to gather user requirements and provide support to address specific technical requirements of

advanced user groups. The findings are presented in Section 3 of this document. An analysis of the specific networking needs of these projects is provided in Section 4.

Organisation of this document

Section 2 of this document describes the process and methodology adopted for the requirements gathering work. Section 3 describes the findings of the requirements gathering. Section 4 discusses the implications of the findings and conclusions/next steps. Section 5 outlines planned activities for Year 2. Appendix A provides a list of ESFRI projects with GÉANT requirements. Appendix B provides a copy of the questionnaire used for the requirements gathering.

Appendix A provides a list of the ESFRI and other projects. Appendix B provides the questionnaire used to collect information. Appendix C describes the project coordination with EEF.

2 The process and methodology followed

An Initial requirements gathering workshop with the ESFRI projects was held in September 2009. Following this meeting, two approaches were used to liaise with European projects:

- Meetings with ESFRI projects via discipline-oriented workshops (organised under the banner of ESFRI and DG-Research/DG Infso)
- One-to-one meetings by GÉANT with individual projects and communities.

These two approaches are described below.

2.1 Meetings with ESFRI Projects via Workshops

The first ESFRI requirements workshop was co-located at the EGEE [2] meeting in Barcelona in September 2009 **[EGEE09]** and organised by EEF. Projects attending the first ESFRI workshop included:

ESFRI Area	Project	Description
Social Sciences & Humanities	CLARIN	Research infrastructure to make language resources and technology available and useful to scholars of all disciplines
Social Sciences & Humanities	DARIAH and CESSDA	Digital infrastructure to study source materials in cultural heritage institutions Facility to provide and facilitate access of researchers to high quality data for social sciences

[2] EGEE is Europe's leading grid computing project, providing a computing support infrastructure for over 13,000 researchers world-wide, from fields as diverse as high energy physics, earth and life sciences.

ESFRI Area	Project	Description
Biological and Life Sciences	ELIXIR	Upgrade of the European Life-science infrastructure for biological information
Physical Sciences and Engineering	CTA	Cherenkov Telescope Array for Gamma-ray astronomy
Physical Sciences and Engineering	FAIR	Facility for Antiproton and Ion Research
Physical Sciences and Engineering	SKA	Square Kilometre Array telescope for radio-astronomy
Environmental Sciences	EISCAT3D	Upgrade of the EISCAT facility for ionospheric and space weather research
Environmental Sciences	EPOS	Infrastructure for the study of tectonics and Earth surface dynamics
Environmental Sciences	LIFEWATCH	Infrastructure for research on the protection, management and sustainable use of biodiversity
Materials and Analytical Facilities	XFEL	Hard X-ray Free Electron Laser in Hamburg

The requirements gathering work has been supported by the ESFRI meetings in Brussels, jointly organised by DG INFSO [DGINFSO] and DG Research [DG Research].

So far there have been meetings of the Bio-Medical projects in December 2009, the Social Sciences and Humanities in January 2010 and the Environmental Sciences in March 2010. Members of GÉANT who are working on NA4 Task 3 have participated in these meetings and held discussions with the projects. The common requirements gathered from these meetings are described in this document. For a list of ESFRI and other projects, see *Appendix A*.

It is understood that the Commission intends to hold similar meetings with the other ESFRI disciplines, such as Energy, Physical Sciences and Engineering and Materials and Analytical Facilities.

2.2 Meetings with Individual Projects

In addition to discipline-based workshops, the NA4 Task 3 project team has held one-to-one meetings with individual projects and communities. This included:

- ELIXIR and 1000 Genome (Bio Informatics)
- European Maritime Safety Agency (EMSA)
- EUMETSAT
- World Meteorological Organization (WMO)
- newGRID and outGRID (Neural Science)
- EXPReS and NEXPReS (Radio Astronomy).

For a complete list of all projects that were contacted, see *Appendix A*.

The NA4 Task 3 project team was also invited to the Bio-Informatics workshop held in Beijing in October 2009 to present and discuss networking options with the Chinese and European Bio-Informatics communities together with CSTNET and CERNET, the Chinese NRENs.

These meetings focused on understanding the networking needs of these projects and initiating discussions on how to best match their needs with the GÉANT portfolio. In several cases, this has resulted in collaborative proof of concept testing over the GÉANT-NREN network.

Meeting the networking requirements of any project involves a subset of the GÉANT NRENs; the precise partners involved depends on the geographic spread of the project. In any project there is, in general, a lead organisation, responsible for the overall management of the project. In order to organise appropriate involvement with the projects, it is been agreed that this will primarily be led by the NREN who has regional headquarters in their geographic area. The lead NREN, together with the GÉANT project coordinator, will act as the primary point of contact, involving other interested NRENs, as required. This allows the specific NREN knowledge of the local environment or user community to be fully utilised.

3 Findings of the requirements gathering

This section reports on the findings from:

- Meetings with ESFRI projects via discipline-oriented workshops (organised under the banner of ESFRI and DG-Research/DG Infso)
- One to one meetings by GÉANT with individual projects

3.1 Findings from meetings via the Project workshops

This sections details the work performed in liaising with the Trans-European infrastructures on the ESFRI roadmap [ESFRI Roadmap] to determine their e-Infrastructure requirements, in particular their network needs.

3.1.1 Social Sciences and Humanities

The Social Sciences and Humanities contribute actively to and are necessary instruments for understanding the cultural, social, political and economic life in Europe as well as for the process of European cohesion and bringing about changes. In practice these disciplines make significant contributions to important areas such as strengthening employment, modernising social welfare and education systems, and securing economic reform and social cohesion as part of a knowledge- based economy.

Meetings

The meeting of the Social Sciences and Humanities (SSH) ESFRI projects, jointly organised by DG INFSO and DG Research, was held in January 2010.

Projects

The following projects attended:

- **CLARIN** is a large-scale pan-European collaborative effort to create, coordinate and make language resources and technology available for the whole European Humanities (and Social Sciences) community. See [CLARIN].

- **ESS** [ESS] (The European Social Survey) is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations.
- **DARIAH** [DARIAH] (Digital Research Infrastructure for the Arts and Humanities) aims to enhance and support digitally-enabled research across the humanities and arts, as well as to develop and maintain an infrastructure in support of ICT-based research practices and to share expertise and tools for the creation, curation, preservation, access and dissemination of data.
- **SHARE** [Share] (Survey of Health, Aging and Retirement in Europe) is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks. **CESSDA** [CESSDA] is an umbrella organisation for social science data archives across Europe. The CESSDA Catalogue enables users to locate datasets, as well as questions or variables within datasets, stored at CESSDA archives throughout Europe.

Requirements

Within the Social Sciences and Humanities community several common requirements were identified.

- *Data archiving and curation*

Data archiving and curation is a common need for several of the ESFRI projects. To enable this there is a requirement for a flexible repository system and a system to provide Persistent Identifiers (PIDs) which together will provide the basis for data storage/archiving and management.

Research by the Social Sciences and Humanities has the need for both long-term storage as well as short term usage of their databases and archives; however the detailed requirements may well be different. Typical sizes of the database for the SHARE project are a few hundred Gigabytes. For the CLARIN project, 50 to 100 TBytes would be the typical size of the data for language material. With the exception of the linguistic computations of CLARIN, this area had light-weight, computing storage and networking requirements.

- *Metadata standards*

Metadata standards used to form an extensible interoperable framework are essential. There are two aspects: the description of the data and the cataloguing of the data. Common tools supporting data archiving, data discovery and mining seem attractive.

- *Web-based access*

There was a strong need for good access to the data archives by the academic and research user community, usually via web-based transactions.

- *Authentication and Authorization*

The sensitive nature of the data to be stored leads to a need for a fine grained, distributed Authentication and Authorization (AA) system, it also is virtually imperative that such a system provides Single Sign On (SSO) functionality.

There has been some involvement with the related work on AAI in the GÉANT project. The scale is thousands of data sets in many countries. In some projects, such as the surveys of SHARE, the security and confidentiality of the data is vital— for example, when it contains blood samples, DNA, wealth and social security data

- *Cloud/Grid computing*

The ability to use grid or cloud computing facilities for the processing of the stored data is also foreseen in some projects.

- *Education and training*

Education and training covering the e-infrastructures and associated technologies was clearly requested by SSH the community.

- *IT Experts*

There was recognition of a need for contact with real IT experts, but that it would be best to find experts internal to the SSH area.

3.1.2 Environmental Sciences

The ESFRI Environmental Science projects represent a diverse set of demands in respect of e-infrastructures, including measurement and monitoring facilities, access to analytical facilities such as synchrotrons, as well as large-scale access to unique global facilities and distributed facilities on a pan-European basis.

The overall objectives are to support the sustainable management of the environment by monitoring and measuring major environmental systems.

Meetings

A meeting in Brussels for the Environmental Sciences on the ESFRI roadmap was jointly organised by DG INFSO and DG Research during March 2010.

Projects

The following projects attended:

- **EISCAT-3D** [EISCAT-3D] is planned as a distributed network of incoherent scatter radar, capable of making measurements of the upper atmosphere.
- **EMSO** [EMSO] is a European Multi Disciplinary Network of seafloor observations, providing permanent monitoring of the deep sea.
- **EPOS** [EPOS] is an integration of existing Plate Observation Systems into a coherent distributed research infrastructure.

- **EUFAR-COPAL** [EUFAR] is a proposal for a heavy pay-load, long-endurance aircraft to provide a platform for airborne measurements across a range of disciplines.
- **EURO-ARGO** [EURO-ARGO] is a proposal to develop the European component of a global ocean observation system.
- **EUSAAR-I3** [EUSAAR] is a project to provide for the integration of atmospheric aerosol properties measured at a distributed network of European ground stations.
- **EARLINET-ASOS** [EARLINET] is a cooperative activity among operations of Aerosol LIDAR systems across Europe.
- **IAGOS** [IAGOS] is a project exploiting the routine measurement of atmospheric composition by installing instruments on commercial aircraft.
- **ICOS** [ICOS] is a project which plans to integrate terrestrial and atmospheric observations of greenhouse gases into a single dataset.
- **LifeWatch** [LifeWatch] is a network of observations and biological collections brought together in a virtual laboratory to measure biodiversity.

Requirements

The disciplines that make up the Environmental Sciences use many different techniques, and their laboratories are also quite different including fixed sites, sensor networks, ships, planes and satellites, which may be located all over Europe or the world. The data is complex and sometimes fuzzy and both primary and processed data are important.

- *Data archiving and curation*

There was considerable discussion about access to data as well as the long-term sustainability, permanence and curation of the data – for example, who selects what data to keep, who pays, how much it costs who curates for the future – maybe for generations to come.

There was concern regarding the need for permanent identifiers and authorisation where they saw grids as being a potential, partial solution.

- *Metadata framework*

There is a clear need to be able to exchange data across the disciplinary areas, implying a need for a metadata framework with semantic interoperability, ontology and provenance. In this context a common database infrastructure and management was thought to be useful.

- *Networking requirements*

This diverse group of activities had quite varied networking requirements with data moving needs varying from bulk data transfer, through to collecting data from intelligent networks of autonomous sensors and observatories to quasi real-time requirements such as the operation of remote and/or mobile instruments and observatories.

- *Cloud/Grid/HPC computing*

The environmental sciences are increasingly data rich and, as a data-driven science, will benefit from data services and computation in data analysis. The analytical and modelling platforms need both high-performance computing (HPC) and distributed grid or cloud computing, implying an underlying requirement for efficient high-performance reliable networking.

- *Fixed and mobile data collection*

Significant investment is proposed in both fixed and mobile server systems, collecting data from land, sea and air measurements, using fixed and mobile data collection. The ICT challenges associated with the sector include data capture, particularly from sensor networks, and the combining, processing and storage of large and complex data sets. There are also some significant real-time requirements in terms of collecting and processing data.

3.1.3 Energy

The ESFRI roadmap stresses the importance of economically competitive, environmentally friendly and sustainable energy resources for European development. A coherent policy for Research Infrastructures is needed to maintain Europe's world leadership in efficient use of energy, in promoting new and renewable forms and in the development of low carbon emission technologies.

The areas covered are Carbon dioxide capture and storage, nuclear fission and fusion, wind energy, solar energy, biofuels, ocean/marine energy, hydrogen, and smart energy grids.

Meetings

Meetings will be scheduled in Year 2.

3.1.4 Biological and Medical Sciences

The biological and medical sciences (BMS) projects within ESFRI cover a range of disciplines with a general focus on health and drug development. There is also some work in the area of Marine biology. Developments in the field and the application of ICT are leading to huge increases in the amounts of data available. These in turn require access to well structured databases, which should be broadly accessible. Recognizing that the value of the data being collected far exceeds the costs of storing and accessing it, the development of distributed infrastructure to store, curate and provide access globally is a key part of the planning.

Meetings

The meeting of the ESFRI Biomedical projects was jointly organised by DG INFSO and DG Research in December 2009. Dante was invited to a follow-up meeting at EBI in Hinxton in January 2010, where the common actions of this ESFRI area were agreed.

Projects

The following BMS projects attended:

- **ELIXIR** [ELIXIR] is a secure, evolving platform for biological data collection, storage and management, consisting of an interlinked set of core and specialist resources.
- **BBMRI** [BMRI] is a distributed pan-European infrastructure of bio-banks, incorporating biomolecular research tools and biocomputational tools.
- **ECRIN** [ECRIN] supports multi-national clinical research trials in Europe by connecting together nationally coordinated networks of clinical research networks and clinical trials units.
- **EMBRC** [EMBRC] will provide an infrastructure connecting the main coastal marine laboratories in Europe, to facilitate common research and training.
- **ERINHA** involves the development and cooperation of European bio-safety level 4 laboratories in Europe.
- **Euro-BioImaging** [Euro-BioImaging] is planning the construction and operation of connected facilities, providing access to imaging technologies, covering both biological and medical applications.
- **Infrafrontier** [Infrafrontier] is organising infrastructure among 15 European laboratories to provide large-scale phenotyping and archiving of mouse models.
- **Instruct** [Instruct] is organising a distributed infrastructure of core and associated centres for integrated structural biology.
- **eNMR** [eNMR] aims to provide the European biomolecular nuclear magnetic resonance community with a platform for access to appropriate computational methods.
- **neuGRID** [neuGRID] is planned as a GRID-based facility for the neuroscience community to assist in research on degenerative brain disease.

Requirements

In order to organize user-friendly data access, a major investment in computer infrastructure and storage is envisaged, along with the development of appropriate standards and ontology's. Developments in imaging are likely to give rise to significant increases in data volumes, which will place new demands on computing, networking and storage. The main demands on e-infrastructures are seen in terms of storage, grids, networks and general computing, with high performance computing seen as being of lesser importance.

- *Networking requirements*

The current data scales are petabytes moving to exabytes over ESFRI period, the data growth has been exponential for many years, and has exceeded improvements in CPU and storage. There is a need to ensure that the major computer centres have network connections with sufficient bandwidth, or point-to-point circuits where appropriate, to allow the required exchange of data to occur in a timely manner.

Data from these archives is made available to a worldwide community of millions via the web, implying the need for a very good routed IP service. Similar database information resources are located in the US and Japan and there is a requirement to exchange data updates on a daily basis. It is planned to explore the use of point-to-point circuits with EBI for these cases.

- *Interoperability and Metadata*

There is a common need for data interoperability. The heterogeneous data that needs to be stored and exchanged has a vast span including genomes, nucleotide sequence, genes, proteomes, protein families' interactions and structures, cells, and bio-systems.

3.1.5 Materials and Analytical Facilities

The development of new materials contributes to all areas of EU activity, from energy generation and storage through to medical implants and computer components. Advances in material sciences has been fuelled by the capability of observing materials at an ever increasing level of microscopic detail — down to the nanometre scale.

Examples of European research projects in this area include the European Synchrotron Radiation Facility (ESRF), European Spallation Source (ESS) and European Magnetic Field Laboratory (EMFL). Research sites range from small to large facilities dispersed throughout Europe. Planned site upgrades, new site installations and development of the infrastructure connection between sites are the major requirements in this area over the next few years.

Meetings

Meetings will be scheduled in Year 2.

3.1.6 Physical Sciences and Engineering

Physical sciences deal with phenomena at all scales and complexities, from astronomy and astrophysics to particle and nuclear physics. Research facilities have become larger, technically more complex and expensive, requiring a joint pooling of intellectual and financial resources from EU members. Such facilities drive the development of new technologies and new methods of working, for instance in ICT and in the applications of nanotechnology and superconductivity.

Examples of European research projects in this area include the European Extremely Large Telescope (E-ELT), the Square Kilometre Array (SKA) and Cubic Kilometre Neutrino Telescope (KM3NeT).

Meetings

Meetings will be scheduled in Year 2.

3.2 Findings from meetings with individual projects

These are communities or individual projects. Requirements from these projects were gathered by GÉANT on an individual basis rather than through the ESFRI and EEF forums.

3.2.1 European Bioinformatics Institute

EBI [EBI] collect, organise and distribute biological information with a worldwide base. This data includes genome, molecular and protein information types, as well as an archive of biomedical publications. There are 20 member states as well as Israel and Australia. They have five locations in Europe with the headquarters in Heidelberg. Typically there are about half a million users worldwide; there are two major collaboration sites—one in the US and the other in Japan. There are also numerous interactions and data exchanges with the Chinese Bioinformatics community.

Meetings

A first meeting was held with this community at Hinxton in the UK in May 2009. A technical follow-up visit was made to EBI in July 2009 for an in-depth discussion on their network topology and to prepare for joint tests to explore network performance issues. Detailed UDP and TCP performance tests were then made between July and September from Hinxton and London to locations in Europe and China using the ORIENT link. An initial report has been made and discussed with EBI.

NA4 Task 3 was also invited to the Bioinformatics workshop held in Beijing in October 2009 to present and discuss networking options with the Chinese and European Bioinformatics communities together with CSTNET and CERNET (the Chinese NRENs). The meeting provided a good view of the current and emerging requirements of the Chinese and European Bioinformatics communities.

Requirements

Organising the biological data involves:

- Entering biological data into databases – incoming bulk data transfers at almost constant rate.
- Sharing the data in these databases between the US and Japanese partners so that all three sites have replicas of all the data.
- Serving the data to many users all over the world. Often they use “local” web or web services portals to ensure good performance to the end users.

There have been issues with achieving the required throughput performance when using usinf TCP/IP for inter-continental bulk data transfers, e.g. to NCBI in the US and to Japan. They use “aspera” a reliable UDP/IP based transfer protocol from a commercial supplier.

There are three main user-groups in China; in 2009 each needed to move about 1 TByte/day to and from Europe; however it is estimated that during 2010 this will increase to an average of 3 TBytes peaking at 5 TBytes.

Next Steps

As well as continuing to work with EBI, in collaboration with the relevant NRENs, on performance matters and current international issues, we will also work with them to help develop a strategic plan as to how the expanding future communications needs of their community can be best met.

3.2.2 DC-NET

The main objective of the DC-NET project [DC-NET] is to generate a comprehensive plan of joint activities for the implementation of a new data and service e-Infrastructure to form a multidisciplinary virtual research community for Digital Cultural Heritage. DC-NET aim to develop and strengthen the co-ordination among the European countries and their public research programmes in the sector of the digital cultural heritage. The project will integrate the research capacities of the participant member states and foster the cooperation between the cultural sector and technology and infrastructure providers

Meetings

A member of NA4 Task 3 was invited to attend the kickoff meeting in Rome, where the project objectives and work plans were presented.

Requirements

DC-NET will start by establishing a Network of Common Interest, between the cultural heritage ministries, agencies and research centres, and the working groups which focus on specific project areas. They will then explore the e-Infrastructures available in each member state, the policies and regulations which govern access to the e-Infrastructures, and the mechanisms and technical standards that need to be complied with. Armed with this knowledge about e-Infrastructures, DC-NET will match what can reasonably be achieved with the priorities of program of the cultural ministries. This will allow them to define a Joint Activities Plan for digital cultural heritage research to effectively use e-Infrastructures.

DC-NET recognises the importance of working with GÉANT and the NRENs. They see potential use of the network, not only for transporting images but also to provide users with a three-dimensional (3D) view of an artifice or an archaeological site. This would require both bandwidth and a real-time response from the network and e-Infrastructure, and is a more demanding application than those foreseen by the projects in the ESFRI Social Science and Humanities area.

Next Steps

NA4 Task 3, together with members from the NRENs, will attend the DC-NET meetings that discuss the network requirements and work with them on selecting the most appropriate connectivity and network services from the GÉANT-NREN portfolios.

3.2.3 European Maritime Safety Agency

The European Maritime Safety Agency (EMSA) [EMSA] developed the CleanSeaNet service [CleanSeaNet] a satellite-based monitoring system for marine oil spill surveillance and detection in European waters. The service provides rapid delivery of oil spill alert information and images, using radar satellite imagery acquired by SAR satellites [3], to Member States. Identification of potential spills in near real-time is essential and EMSA has contracted a network of ground stations throughout Europe able to downlink, process and analyse satellite data within a maximum of 30 minutes after satellite overpass. These images of the sea and coastline round

[3] SAR (Synthetic Aperture Radar) is Germany's first reconnaissance satellite system, operated jointly with other EU members.

Europe will be sent in real-time from the ground stations to EMSA premises in Lisbon, Portugal using the GÉANT-NREN network.

Meetings

NA4 Task 3 has held discussions with EMSA to fully understand their work and networking requirements. The connectivity options were presented and discussed and EMSA were advised on the suitability of the routed IP service.

Requirements

The transmission of one image (~500 Mbytes) from a ground station to EMSA should not last longer than one minute. This indicates a required bandwidth of 66 Mbit/s. EMSA will have a 1 Gigabit connection to FCCN the NREN in Portugal.

The protocols required include VPN, sFTP and a two-way secure SSL - HTTP protocol.

Next Steps

NA4 Task 3 will continue to work with EMSA to ensure that they are able to move their data from the ground station over the GÉANT service area network to their computer centre in Portugal within the required minimum delay period.

EMSA have agreed to present their satellite surveillance activities in an oral presentation at the ICT exhibition in Brussels on 27–29 Sep 2010 and are happy to demonstrate how the GÉANT network will support these applications.

3.2.4 EUMETSAT

EUMETSAT [EUMETSAT] operate a number of satellites collecting earth observation data and images. The data from the satellites is received by a number of ground stations and sent to EUMETSAT for processing; the resulting data is then disseminated as various data products to users and the research community using commercial telecommunication satellite links, in a similar way to satellite television. There are a wide range of applications for the data products including numerical weather predictions, climate modelling, land surface analysis, Ozone and atmospheric chemistry monitoring, hydrology and water management, and active fire monitoring.

Meetings

Contact was first established with EUMETSAT at the GEO World Plenary in Bucharest in November 2008, and a follow-up meeting was held at their offices in Darmstadt in January 2009 to understand their requirements in detail and discuss what options of the academic networks would be most suitable.

A further meeting during July 2009 discussed the advantages of four different connection scenarios and explored operational aspects of the GÉANT service area.

Since then there have been regular telephone meetings to discuss progress and plan the proof of concept tests.

Requirements

There is a continuous flow of data from acquisition, processing to dissemination. The timing is critical as the information must be with the consumers within five minutes of the images being taken.

The current satellite dissemination system is based on standard Digital Video Broadcast (DVB) technology with a total bandwidth of about 15 Mbit/s. However with new image systems and more detailed products, they expect to require between 100 and 400 Mbit/s. Hence they are interested in complementary use of satellite and the academic network.

Both unicast and multicast transmission methods need to be tested.

Next steps

Following the visit from EUMETSAT in July 2009, members of NA4 and SA1 have worked out the traffic patterns for data flows from Germany to the countries of interest and to EUMETSAT, and demonstrated the possible reduction in traffic if multicast were used to move their data.

Test programs to investigate the detailed behaviour of multicast packets in a multi-domain environment have been developed, tested, and made available to EUMETSAT together with some documentation. Detailed tests of high-bandwidth multicast flows have been made over the GÉANT backbone and further multi-site, multi-flow tests are in progress. A technical note describing the work and results is in preparation. The multicast test programs and scripts will also be made publically available.

An exercise to obtain budgetary costs for EUMETSAT is almost complete for a sample of countries. Although some variation in the cost of a connection was anticipated due to the different funding models of the NRENs, the actual range of costs was considerable. Unfortunately, this exercise took much longer than expected and did not provide a clear budget figure. In parallel to this exercise, checks were made to determine if the facilities requiring access to EUMETSAT data were already connected to their respective NRENs. It turned out that most of the sites were already connected with sufficient bandwidth to the NRENs.

A collaborative proof of concept set of tests over the GÉANT-NREN network using unicast and multicast will be performed in the near future using EUMETSAT's dissemination software.

EUMETSAT have agreed to this collaborative work being reported as a poster at the coming TNC [TERNA CONF].

3.2.5 NeuGRID

Description

neuGRID [neuGRID] aims to build a new user-friendly Grid-based research e-Infrastructure, where the collection and archiving of large amounts of imaging data is paired with computationally intensive data analyses of the images and the ability to visualise the results. They aim to provide large sets of brain images that will provide neuroscientists with the capability to identify neurodegenerative disease markers, for example, for Alzheimer's disease, through the analysis of 3D magnetic resonance brain images. The Grid infrastructure is

based on EGEE's GLite middleware. Currently there are eight partner research sites distributed across Europe, each with 100 processing cores and terabytes of disk storage.

The outGRID project [outGRID] is a Coordination and Support Action which aims to help the three e-infrastructures—neuGRID in Europe, CBRAIN in Canada, and LONI at UCLA in the U.S—to converge into one unique worldwide facility.

outGRID will organize workshops to promote the exchange of technical information, direct the development of the infrastructures towards interoperability, and promote specific international calls aiming to achieve full interoperability.

Meetings

After an initial exchange of email with members of newGRID, which occurred during the Open Grid Forum (OGF) [OGF] meeting in Alberta Canada in October, NA4 Task 3 had a positive meeting with US and Canadian Neural Biologists to discuss how they might use the GÉANT links between Europe and the US and Canada to work with the neuGRID and outGRID projects. This matter was also raised with members of CANARY who were joint hosts of the OGF meeting.

A meeting to discuss the neuGRID and outGRIG projects and fully understand their requirements was arranged with senior members of both projects at the GÉANT Launch event in Stockholm at the beginning of December 2009.

Requirements

neuGRID regularly transfers brain scan images from databases in the US for analysis and visualisation with the newGRID algorithms. Once transferred, selected data is distributed to the Grid sites at the European partners.

As part of outGRID, the Neural Science community, who already operate neuGRID in Europe using GÉANT, wish to extend their work to collaborate with researchers in Canada and the US via the outGRID project. They also wish to exchange data on brain images with the CBrain and GBrain projects.

Next Steps

GÉANT will follow up with the project to determine if the performance concerns raised at the meeting have been resolved, and discuss requirements for links to partners in Canada and the US.

3.2.6 EXPReS and NEXPReS

The objective of EXPReS [EXPReS] is to connect up to 16 of the world's most sensitive radio telescopes on six continents to the central data processor of the European VLBI Network at the Joint Institute for VLBI in Europe (JIVE), and also pioneer the movement of data between telescope and correlator at 4 Gigabits/s.

NEXPReS builds on the work of the EXPReS project and will offer enhanced scientific performance for all users of the European VLBI Network and its partners. By providing high performance, transparent data buffering mechanisms at the Telescopes and the correlator it will be possible to address the current bottlenecks in e-VLBI and avoid limitations in connectivity, bandwidth, and processing capacity while eliminating the need for

the transport of physical media. Real-time grid computing and high bandwidth on demand will also be investigated to improve performance and prepare EVN for the use of higher bandwidths. The project has just been accepted for funding by the commission.

Meetings

DANTE was a partner in the EXPReS project, which started in March 2006, with responsibility for organising the EVN-NREN forum that brings together experts from the radio astronomy community and the NRENs to discuss EXPReS and the network requirements. From January 2008 DANTE also led the performance measurements on the EXPReS 4 Gigabit data flows using FPGA[4] systems.

Two substantial papers: "Commissioning and Using the 4 Gigabit Light path from Onsala to Jodrell Bank" and "iNetTest a 10 Gigabit Ethernet Test Unit" describing the work done by DANTE and GÉANT in the EXPReS project were presented at the 8th International e-VLBI Workshop [e-VLBI workshop] held in Madrid in June 2009, and were accepted for publication by the Proceedings of Science journal [POS1] [POS2].

EXPReS received a positive final review in Brussels with the reviewers liking the way in which EXPReS collaborated with GÉANT and the NRENs both directly and through the EVN-NREN meetings organised by DANTE.

Next steps

DANTE is a partner in NEXPReS with responsibility for organising the EVN-NREN forum that brings together experts from the radio astronomy community and the NRENs to discuss NEXPReS and the network requirements. Recent developments that are relevant to NEXPReS and related work from other activities, e.g. Network performance and Bandwidth on Demand, will also be discussed.

3.2.7 World Meteorological Organization (WMO)

To predict the weather, modern meteorology depends upon near instantaneous exchange of weather information across the entire globe; this is coordinated by the World Meteorological Organization (WMO) [WMO]. The Meteorological offices and associated institutes in each country collect local data and run simulation models to predict the weather patterns on various timescales. Some run climate models. The output, data, together with the corresponding metadata, is made available to the local regional Global Information System Centre (GISC), which is linked via the WMO Weather Information System (WIS) to all the GISCs worldwide. Metadata (and data) updates are synchronised between all the GISCs so the weather information is available throughout the world. High priority messages such as typhoon warnings are also exchanged between the GISCs. Many of the computations must be performed in a timely manner, such as the hourly weather synopses.

Meetings

NA4 Task 3 has had discussions with Deutscher Wetterdienst (DWD), The World Meteorological Organization (WMO) and The European Centre for Medium-Range Weather Forecasts (ECMWF) on the possible inclusion

[4] A field-programmable gate array (FPGA) is an integrated circuit designed to be configured by the customer or designer after manufacturing.

of the academic Internet for use in their WIS, which would be a replacement for their current Global Telecommunications System (GTS).

These discussions led to an invitation from ECMWF to make a presentation, in 25th February 2010 at a "vendor's information event", on the suitability of the GÉANT-NREN network for use in the WIS. A joint presentation was given by members of DANTE and the UK NREN JANET. While ECMF accepted that the academic network had the required coverage, was designed to be reliable and would deliver high bandwidth to the user, some of those present had experienced a poor level of 24/7 support for the service. They were disappointed that, unlike their commercial provider (Orange Business systems), we could not provide a simple total cost for the service, and did not normally have SLAs; they also preferred a single point of contact, rather than having to contact multiple areas for their support.

NA4 Task 3 was also invited to a meeting of the Inter Commission Coordination Group on WIS in Korea to give a presentation on the potential use of research networks. The main concern about using academic networks as a complete WIS solution was the need for 24/7 support. Several examples were given where this did not occur, including US, Korea and France. Also due to the way the academic networks were procured in the US they were not allowed to carry "operational traffic". However, there was interest in complementary use of commercial and academic networks in linking the GISCs for high bandwidth exchanges. DWD and the Chinese Met office had already been making performance tests of exchanging metadata and were interested in trying the academic network.

Requirements

The current bandwidths used on the commercial networks are a few Mbit/s, which is sufficient for the current small text-based data /metadata exchanges. With new and more detailed data and images about 100Mbit/s or up to 1 Gbit/s might be required.

Next Steps

A Memorandum of Understanding (MoU) between DANTE on behalf of the GÉANT project and WMO has been drafted and sent for comment to DWD and WMO. This will be progressed.

We will follow up the establishment of a route over GÉANT and the ORIENT link between Europe and China for collaborative tests between DWD and the Chinese Met office.

3.2.8 European Institute of Innovation and Technology (EIT)

EIT is a new independent community which was set up to address Europe's innovation gap. See [EIT] and [EIT brochure].

The mission of the EIT is to grow and capitalise on the innovation capacity and capability of stakeholders from higher education, research, business and entrepreneurship from the EU and beyond through the creation of highly integrated Knowledge and Innovation Communities (KICs).

Meetings

At meeting was arranged at the Information Event for the 7th Call for Proposals e-Infrastructures on 18 June 2009 in Brussels. This was attended by EU project officers with responsibility for EIT together with representatives of GÉANT and EGEE. GÉANT gained a clear idea of the thinking and concepts behind EIT and the KICs and discussed how these new joint academic-industrial communities could best be made aware of the potential of the European funded e-Infrastructures such as the GÉANT and NREN networks, GRID computing from EGEE, HPC computing from DEISA, as well as the work on e-repositories and collaborative tools.

Requirements

To introduce the services available from the European integrated e-Infrastructure to the European Institute of Innovation and Technology (EIT), GÉANT drafted the document "The GÉANT Network and EGEE Grid Services for e-Infrastructure Users" [EIT board doc].

This document makes EIT and the KICs aware of how the layered e-Infrastructure could support researchers, educators and innovators over the multi- domain environment of Europe from the Network connectivity services provided by GÉANT through the Grid computing from EGEE to the digital repositories and the knowledge infrastructure was well received by the commission.

Next Steps

During year 2 GÉANT will make contact with the three KICs that have been formed to discuss and understand their networking requirements.

4 Implications of the requirements gathering

The requirements that projects have described are broad-ranging. These requirements can all be accommodated using current and predicted technology. The requirements themselves range from:

- Multiple access of databases from a diverse population of users
- Concentrated flows between key project locations
- Real-time performance requirements.

The global requirements that have been articulated to date are within the activity footprint of the GÉANT global reach, and relationships are in place to assist in organising network service beyond Europe. The portfolio of services that is available across the GÉANT service area (including high-performance IP, configurable Point to Point connections and, where appropriate, dedicated wavelength capacity) is capable of meeting the needs that have been articulated.

Their specific issues are described in more detail below.

4.1 Main issues identified

A number of points have emerged from the discussions with individual projects. The projects:

- Have quite specific (often not especially well-informed) views about their networking requirements.
- Have limited, if any, understanding of the benefits of networking or of the technology or the organisation of networks.
- Look to have a single point of contact and / or a single offer and a single point of billing.
- Need a lot or organisational support and request a more organisational and functional focus to such assistance.
- Have concerns about the quality of service, service level agreement and user support available from the research networking community (in contrast to the level of service provider by a commercial provider).

The following comments are made about the more general implications for GÉANT and the NRENs in supporting the current and developing needs of projects.

Geographic Scope

Availability of services should not be a problem but for specific locations access capacity to those locations, including capacity available for elements of the service portfolio, would need to be confirmed and, if appropriate, addressed.

Identity Management (with single sign-on)

All the ESFRI projects consulted identified consistent identity management and single sign-on as a basic requirement. This is something that is available as part of overall network capability. Its precise role in supporting project needs requires further development.

External Access

There is a need to allow some non Research and Education access to information. This is in general not a significant problem, but performance may be an issue. It needs to be better scoped and understood.

Complex Requirements

For more complex requirements, the design and the implementation of the network service will require cooperative effort between the research network community and the project participants. Providing support from the NREN community to the project during their requirements analysis, implementation and ongoing service operations is the main issue.

Common Proposal development

There is currently no process in place for coordinating proposals for connecting new users to the NREN. When questions arise from the new users (such as whether their particular site is already connected to the NREN and what is the budgetary estimate for a connection) finding answers to these questions takes far too long. (In contrast, a commercial provider can usually respond in a much more rapid and coherent manner.)

One of the difficulties is that each NREN's proposal is different in format, content and presentation. The NRENs use different funding and charging models as well as service provision conditions. This makes it difficult to collate the separate elements into a single, coherent proposal that summarises both the various national and GÉANT elements. An example of this was the difficulty in putting together a proposal for the European Space Administration's (ESA), who were interested in a NREN and GÉANT supported network to link a number of their sites.

Support and problem ownership

Some user communities already have considerable experience of operating private network connections, leased from commercial providers and used to link their sites. The advantage of a commercial service is that it provides a single point of contact and problem ownership. Although the user communities see the benefits of using the academic network for research and non-commercial traffic, they are concerned about ownership of the NREN-GÉANT network service, its operational availability and the lack of Service Level Agreements (SLAs).

GÉANT, together with lead NRENs, have addressed these concerns by emphasising the role of the GÉANT service area in the integrated European e-Infrastructure, showing the approach for designing the network for reliability, and demonstrating the high availability.

Performance

Performance, particularly where demanding applications are being supported, will need monitoring and fine-tuning. This is a well-understood issue and the techniques of addressing it are established. However, overall performance in terms of complex systems can be challenging, as it involves interactions between different systems under separate management control. Network tools to help debug such problems are available. As part of customer support, GÉANT is prepared to analyse and diagnose performance problems.

The discussions with both existing and the new user communities have indicated that they are encountering similar performance issues in moving bulk data or accessing remote data to that which other users (such as the Particle Physics and Radio Astronomy community) have already dealt with. There is scope for sharing of information, to enable new users to benefit from the experiences of existing users.

Training and education

All ESFRI projects have expressed the need for training, education or external expertise in their use of e-infrastructures.

GÉANT provides training on network performance analysis and improvements. GÉANT has an E-Learning portal [e-Portal] which hosts self-paced training modules about perfSONAR (multi-domain network monitoring tool) [perfSONAR] and has other self-paced education/training courses. See [courses].

GÉANT advertises the services it has via a portfolio. See [Portfolio]. The presentation of services via a portfolio could be extended to all e-infrastructure services.

4.2 Conclusions and next steps

The meetings with the discipline areas highlighted their common requirements and allowed contacts to be made with the individual projects, within their discipline area. The meetings with individual ESFRI roadmap projects provided good insight into the aims and network requirements of the projects.

Feedback from the ESFRI projects has indicated that involvement with the core e-Infrastructure providers via the EEF has been well received and has provided timely input of the services available to the projects.

Collaborative work is ongoing to complete the EEF questionnaire (see Appendix B) to obtain input from projects representative of each of the ESFRI discipline areas. In several cases, the ESFRI roadmap projects have limited awareness of their network requirements, although they are better informed about their data processing, security and storage needs. It should be possible in further dialogue to improve our understanding of the network requirements. We plan to have meetings with individual ESFRI projects to more fully determine their needs and requirements.

To address performance issues, during Year 2 of the project we plan to create a user forum, where it would be possible to help both the new and existing users by passing on tips and techniques to improve performance. This would also create another opportunity to interact with the user communities, to inform them of the services that are available as well as gather their ongoing requirements.

We need to strengthen our ability to work with the projects. As a project, GÉANT is still struggling to present proposals for design and implementation of service and ongoing service support in a consistent and timely manner. The policy of working with a lead NREN when interacting with projects and communities distributed over several countries has worked well and allowed local knowledge of the environment or user community to be fully utilised. This is an area that needs further development.

5 Plans for Year 2

Looking at the requirements of projects and the experience gained from supporting projects in Year 1, the following ideas are suggested in respect of gathering user requirements and developing project support in Year 2.

5.1 Further liaison with Projects

We will continue with the strategy of meeting with individual projects, as well as attending workshops organised under the ESFRI framework, using the lead NREN policy.

5.2 Common Proposal Development

There is an urgent need to develop a common proposal framework so that the GÉANT can provide a single, coherent proposal when required. The existence of different funding and charging models as well as service provision conditions among the European NRENs is a complex challenge, but harmonisation or best practice needs to be put in place.

In future, new proposals could be developed by the coordinating partner together with some of the NRENs who have contributed to such proposals in GN2. An obvious starting point would be the group who organised the proposal to the European Space Administration.

5.3 Ongoing support and problem ownership

The question of providing ongoing operational support to an implementation also needs to be considered. In the context of GN2 the end-to-end coordination unit was established and perfSonar was rolled out to assist the LHC OPN but these services were never formalised as part of a portfolio nor was there a clear definition of roles. Further work is required here.

5.4 Performance improvement

To address performance issues, during Year 2 of the project we plan to create a user forum, where existing users can pass on tips and techniques to improve performance.

5.5 Training and Education

Many projects, particularly but not exclusively from non-science-based disciplines, have little, if any, knowledge of what research networking can offer them, let alone what the technology is. It is proposed to develop an educational workshop to promote the general capabilities of research networking, to provide some very limited introduction to technology and to get some insight into the sort of performance that is possible.

Appendix A List of Projects Requiring GÉANT Services

These are the projects that have been liaised with on an individual basis.

Project Name	Potential Lead NREN	Project Scope	GÉANT Services / Issues
HEP	SWITCH	Moving Particle Physics Data from CERN over LHCOPN to Country Tier1 sites.	Provision of dedicated 10Gigabit point-to-point links Provision of network monitoring Assistance with resolving performance issues
VLBI	SURFnet	Very Long Baseline Interferometry. Moving data between telescopes and correlator	Provision of dedicated point-to-point links Resolution of multi-gigabit performance issues
NEXPReS	SURFnet	Novel EXplorations Pushing Robust e-VLBI Services. Will test deployment of bandwidth on demand and a transparent data caching system that allows all scientific VLBI experiments to benefit of the increased sensitivity, flexibility & robustness of real-time VLBI.	Provision of dedicated point-to-point links Provision of monitoring services Provision of dynamic services Assistance with resolving performance issues
DEISA	DFN	Distributed European Infrastructure for Supercomputing Applications	Provision of dedicated 10Gigabit point-to-point links Assistance with resolving performance issues
GEO	SWITCH	Group on Earth Observations. A member of the Architecture and Data Committee	Analysis of user requirements Synthesis of world network architecture

Project Name	Potential Lead NREN	Project Scope	GÉANT Services / Issues
EUMETSAT	DFN	Monitoring weather and climate from space. Integration of use of Satellites and the Academic Internet	Analysis of user requirements Collaboration on proof of concept data transfer tests
DWD	DFN	Weather and Climate studies	Analysis of user requirements
WMO-WIS	SWITCH	The World Meteorological Organization. The role of academic networking for weather & climate research and dissemination.	Analysis of user requirements Presentation of possible use of Academic Internet Synthesis of network architecture
MPI Meteo	DFN	Max Planck Institute for Meteorology (MPI-M), Hamburg. Climate Research.	Assistance with resolving performance issues
THEOS		Earth Observations by Thailand's state-of-the-art satellite imaging system, operated by the Geo-Informatics and Space Technology Development Agency (GISTDA), Ministry of Science and Technology.	Analysis of user requirements Assistance with resolving performance issues Collaboration on proof of concept data transfer tests
ITER	Renater	ITER is a large-scale scientific experiment that aims to demonstrate that it is possible to produce commercial energy from fusion.	Analysis of user requirements Assistance with resolving performance issues Collaboration on proof of concept data transfer tests
EBI	JANET	European Bioinformatics Institute	Analysis of user requirements Assistance with resolving performance issues Collaboration on proof of concept data transfer tests
1000 Genomes	JANET	A Catalogue of Human Genetic Variation	Analysis of user requirements Assistance with resolving performance issues
ELIXIR	JANET	European life sciences infrastructure for biological information. The mission of ELIXIR is to construct and operate a sustainable infrastructure for biological information in Europe.	Analysis of user requirements Assistance with resolving performance issues

Project Name	Potential Lead NREN	Project Scope	GÉANT Services / Issues
Ensembl	JANET	The Ensembl project produces a genome databases for vertebrates.	Analysis of user requirements Assistance with resolving performance issues
BBMRI		Biobanking and Biomolecular Resources Research Infrastructure	Analysis of user requirements
World Lecture Theatre	SWITCH	Lectures on embodied intelligence, broadcast by videoconference from Jiao Tong University in Shanghai (China-Essex-Zurich India US)	Assistance with resolving performance issues
ESA	GARR	European Space Agency. Using the Academic Internet to link their sites.	Analysis of user requirements Presentation of possible use of Academic Internet Synthesis of network architecture
neuGRID		Grid-based research e-Infrastructure and large sets of brain images enabling study of neurodegenerative diseases by the European neuroscience community	Analysis of user requirements Assistance with resolving performance issues
outGRID		Promote the harmonisation and interoperability of CBRAIN in Canada, LONI in the US, and newGRID in Europe	Analysis of user requirements Presentation of possible use of Academic Internet
EMSA	FNCC	European Maritime Safety Agency. Using satellite imagery to detect oil slicks and tanker accidents.	Analysis of user requirements Help with possible use of Academic Internet
ETI-KITs		European Institute of Innovation and Technology	Generation of documentation showing an overview and advantages of an integrated European e-Infrastructure
PRACE	DFN	The Partnership for Advanced Computing in Europe	Analysis of user requirements Presentation of possible use of Academic Internet
SKA	JANET	Square Kilometre Array	Analysis of user requirements Presentation of possible use of Academic Internet
Karlsruhe	DFN	Interest in Transatlantic dynamic links	Information exchange

Project Name	Potential Lead NREN	Project Scope	GÉANT Services / Issues
Institute of Technology (KIT)			
DIESIS	DFN	Design of an Interoperable European federated Simulation network for critical InfraStructure.	Analysis of user requirements Presentation of possible use of Academic Internet

Table 1.1: Projects requiring GÉANT services

Appendix B The EEF Questionnaire

Survey of e-infrastructure requirements for research facilities

This document is a questionnaire to gather information about the requirements for the use of e-infrastructures by research facilities such as those represented by the ESFRI projects.
This questionnaire has been prepared by the **European e-Infrastructure Forum** (EEF) which is a forum for the discussion of principles and practices to create synergies for distributed Infrastructures. The goal of the EEF is the achievement of seamless interoperation of leading e-Infrastructures serving the European Research Area. The focus of the forum is the needs of the user communities that require services which can only be achieved by collaborating Infrastructures. Its current membership includes GÉANT, Terena, EGEE, EGI, DEISA and PRACE. The forum recognises the importance of data access and management and is seeking to add a member specialising in service provision in this area for multiple research communities.
As the recipient of this questionnaire, we ask to you answer the questions and return the completed document to the email address below. If you already have documents that provide the requested information, please feel free to provide pointers (i.e. URLs) to such documents in response to the questions.
Based on the information received, EEF will produce a summary report which will be distributed to all respondents, relevant policy bodies and used as input to a roadmap intended to address the common e-infrastructure needs. We thank you in advance for your participation and should you have any further questions please do not hesitate to contact us at the email address below.

European E-infrastructure Forum
Europe-einfrastructure-forum@cern.ch
Date: 18th December 2009

Revision	Date	Author	Comment

Question 1: Contact details
Respondents Name:
Respondents Email address:
Respondents Institute:
Research facility represented (e.g. ESFRI project name):

Question 2: ICT vision
Please outline how ICT (Information and Communication Technologies) contributes to the overall research facility vision and goals.

Question 3: Outline the major ICT requirements taking into account aspects such as:
Data sources (from the facility, outside sources etc. – where and how the data originates)

***Storage** needs (scale, longevity, update rate etc.)*
– Average size of the datasets & files
– Expected growth patterns and rates (e.g. 1 TByte/day or /month)

***Compute** needs (e.g. scale and specific computing architecture requirements)*

***Networking** (distribution of data sources, user communities etc.)*
Please consider the following aspects in responding to this point:
Required network usage:
- *Access to the data/information by many users e.g. "web site" data access*
- *Bulk data transfers to some "well known" sites*
- *Real-time or time critical data movement*
- *Other use of the network*

Expected patterns of moving data:
- *Continual replication or movement of data e.g. 1 TByte/day transmitted throughout the day*
- *Periodic rapid transfer of data (e.g. 1 TByte transmitted in 1 hour every 10 hours)*
- *Other patterns of communication over the network*

***Security** (for data access etc.)*

***Access** for user community (estimate size of community, geographical location etc.)*

Question 4: ICT architecture
If you have already defined an ICT architecture then please provide an overview and diagram.

Question 5: Current Status & Timeline
Please state what is the current status of your ICT planning and eexplain what is the timeline for the construction and operation of your research facility. What are the key dates for your computing infrastructure?

Question 6: Challenges and suggestions
Highlight what you see as the major unresolved challenges for the ICT aspects of your research facility. List any suggestions for how European ICT structures (grid, networks, super-computing) can work with/help your research facility.

Any additional information or comments:

Appendix C Project coordination with EEF

It was agreed that the EEF members would establish the requirements matrix which the ESFRI projects had for all e-infrastructure projects in a joint co-operation with the ESFRI projects.

The following process was agreed for gathering the requirements:

- EEF to distribute a basic template of the matrix to the ESFRI projects asking them to modify the template if they thought additional subjects need to be covered.
- Once the template was agreed, EEF would work with the ESFRI projects on a first version
- EEF to analyse the information gathered and then focus on the most common points in more detail with the ESFRI projects

It was agreed that EEF's mandate concerns user communities and hence industrial representation on the membership was not considered appropriate. However, the members were willing to explain what are the relationships of each infrastructure with industry to determine if there is any commonality and hence opportunities for synergy.

Documents created for the requirements gathering

The following documents were produced, to help with the requirements gathering:

- Questionnaire for ESFRI projects. (See *Appendix B.*)
- Letter to ESFRI project leaders to formalise their willingness to work with EEF
- Letter to EEF members to confirm their willingness to contribute to the EEF programme of work
- ELIXIR questionnaire

It was agreed that the project leads would complete the questionnaire for a number of ESFRI projects based on the information already gathered via the EEGE09 session, NEERI09 event and BMS workshop:

- ELIXIR, XFEL, FAIR, CTA
- EISCAT-3d, SKA, EPOS
- Fusion (working with EUFORIA project)

The leads would then distribute the questionnaire to the EEF email list to enable other members to make additions. Once completed the project leads would meet with the ESFRI projects concerned to review the questionnaire and add any other material.

References

[BMRI]	http://www.bbmri.eu/
[CESSDA]	http://www.cessda.org/
[CLARIN]	http://www.clarin.eu/external/
[CleanSeaNet]	http://cleanseanet.emsa.europa.eu/
[courses]	http://www.geant2.net/server/show/ConWebDoc.2757
[DARIAH]	http://www.dariah.eu
[DC-NET]	http://www.dc-net.org
[DEISA/PRACE]	https://www.deisa.eu/news_press/symposium/barcelona2010/deisa-symposium-barcelona-may-10-12-2010
[DGINFSO]	http://ec.europa.eu/dgs/information_society/index_en.htm
[EARLINET]	http://www.earlinet.org/
[EBI]	http://www.ebi.ac.uk/
[ECRIN]	http://www.ecrin.org/
[EEF]	http://www.einfrastructure-forum.eu/
[EGEE]	http://www.eu-egee.org/
[EGEE09]	http://egee09.eu-egee.org/
[eIRG]	http://www.e-irg.eu/
[EISCAT]	http://www.eiscat.se/
[EIT board doc]	Richard Hughes-Jones DANTE, and EGEE, Danielle Venton “The GÉANT Network and EGEE Grid Services for e-Infrastructure Users” for the EIT community “ https://intranet.geant.net/sites/Networking/NA4/T3/Documents/EIT_BoardLevel_final_v2.doc http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=82
[EIT brochure]	http://eit.europa.eu/fileadmin/Content/Downloads/PDF/Background_Information/eit_brochure_en.pdf
[EIT]	http://eit.europa.eu
[ELIXIR]	www.elixir-europe.org
[EMBRC]	http://www.embrc.eu/
[EMSA]	http://www.emsa.europa.eu/
[EMSO]	http://www.emso-eu.org/
[eNMR]	http://www.enmr.eu/
[e-Portal]	http://cbt.geant.net/courses.html
[EPOS]	http://www.epos-eu.org/
[ESFRI Roadmap]	ftp://ftp.cordis.europa.eu/pub/esfri/docs/esfri_roadmap_update_2008.pdf
[ESFRI]	http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=home

[ESnet]	http://www.es.net/
[ESS]	http://www.europeansocialsurvey.org
[EUFAR]	http://www.eufar.net/
[EUMETSAT]	http://www.eumetsat.int/Home/Main/AboutEUMETSAT/index.htm?l=en
[EURO-ARGO]	http://www.euro-argo.eu/
[Euro-BioImaging]	http://www.eurobioimaging.eu/
[EUSAAR]	http://www.eusaar.net/
[e-VLBI workshop]	http://www.oan.es/expres09/
[EXPReS]	http://www.expres-eu.org/
[GÉANT]	http://www.geant.net/
[IAGOS]	http://www.iagos.org/
[ICOS]	http://www.icos-infrastructure.eu/
[Infrafrontier]	http://www.infrafrontier.eu/
[Instruct]	http://www.instruct-fp7.eu/
[LifeWatch]	http://www.lifewatch.eu/
[neuGRID]	http://www.neugrid.eu/pagine/home.php
[OGF]	http://www.ogf.org/gf/group_info/view.php?group=pgi-wg
[outGRID]	http://www.outgrid.eu/site/pagine/home.php
[perfSONAR]	http://cbt.geant2.net/repository/perfsonar_ui/perfsonar_online_training_0.2/player.html
[Portfolio]	http://www.dante.net/upload/pdf/GN3-10-040_DN4-2-1_GEANT_Service_Portfolio.pdf
[POS1]	R.Hughes-Jones, and J. Hargreaves “Commissioning and Using the 4 Gigabit Lightpath from Onsala to Jodrell Bank”, Proceedings of the 8th International e-VLBI Workshop. 22-26 June 2009. Madrid, Spain, http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=82
[POS2]	R.Hughes-Jones, and J. Hargreaves “iNetTest a 10 Gigabit Ethernet Test Unit”, Proceedings of the 8th International e-VLBI Workshop. 22-26 June 2009. Madrid, Spain
[Share]	http://www.share-project.org
[TERNA CONF]	http://tnc2010.terena.org/
[VPH]	http://www.deisa.eu/usersupport/training
[WMO]	http://www.wmo.int/pages/index_en.html

Glossary

CESSDA	Council of European Social Science Data Archives
CLARIN	Common Language Resource and Technology Infrastructure
DIESIS	Design of an Interoperable European federated Simulation network for critical Infrastructure.
DGINFSO	Directorate General Information Society and Media
EARLINET	European Aerosol Research Lidar Network to Establish an Aerosol Climatology
EBI	European Bioinformatics Institute
EC.	European Commission
ECRIN	European Clinical Research Infrastructures Network
EEF	European E-infrastructures Forum
EGEE	Enabling Grids for E-science
EISCAT_3D	Next generation European Incoherent Scatter radar system
ELIXIR	European life sciences infrastructure for biological information
EMBL	European Molecular Biology Laboratory
EMBRC	European Marine Biological Resource Centre
EMSA	European Maritime Safety Agency
EMSO	European Multidisciplinary Seafloor Observatory
eNMR	NMR e-Infrastructure in System Biology
EPOS	European Plate Observing System
ERINHA	European research infrastructure on highly pathogenic agents
eIRG	e-Infrastructure Reflection Group
ESA	European Space Agency
ESFRI	European Strategy Forum on Research Infrastructures
ESS	European Social Survey
EUFAR	European Facility For Airborne Research / Transnational Access
EUMETSAT	Exploitation of Meteorological Satellites
EURO-ARGO	Global Ocean Observing Infrastructure
Euro-BioImaging	European Biomedical Imaging Infrastructure
EUSAAR	European Supersites for Atmospheric Aerosol Research
GÉANT	Pan-European data network dedicated to the research and education community
GEO	Group on Earth Observations
GISTDA	Geo-Informatics and Space Technology Development Agency
IAGOS	In-service Aircraft for a Global Observing System
ICOS	Integrated Carbon Observation System
INFRAFRONTIER	European Infrastructure for Phenotyping and archiving of model mammalian genomes
INSTRUCT	Integrated Structural Biology Infrastructure for Europe

LIFEWATCH	Science and Technology Infrastructure for Biodiversity date and Observatories
MPI Meteo	Max Planck Institute for Meteorology (MPI-M)
neuGRID	Neural Science (newGRID and outGRID)
NEXPReS	Novel Explorations Pushing Robust e-VLBI Services
NRENs	National Research and Education Networks
PARADE	
PRACE	Partnership for Advanced Computing in Europe
NEXPReS	Radio Astronomy
SET Plan	Strategic Energy Technology Plan
SHARE	Survey of Health, Aging and Retirement in Europe
SKA	Square Kilometre Array
TERENA	Trans-European Research and Education Networking Association
VLBI	Very Long Baseline Interferometry
WMO	World Meteorological Organization