



e-Infrastructure for the 21st Century



EIROforum partners are intergovernmental research organisations – CERN, ESA, EMBL, EFDA, ESO, European XFEL, ILL and ESRF – covering disciplines ranging from particle physics, space science and biology to fusion research, astronomy, and neutron and photon sciences. The partner organisations have a truly European governance, funding and remit, and in many cases share a global engagement. They are world leaders in basic research, as well as in managing and operating large research infrastructures and facilities. The EIROforum collaboration is helping European science reach its full potential through exploiting its unparalleled resources, facilities and expertise. By combining international facilities and human resources, EIROforum exceeds the research potential of the individual organisations, achieving world-class scientific and technological excellence in interdisciplinary fields. EIROforum works closely with industry to foster innovation and to stimulate the transfer of technology.



Publication prepared by CERN on behalf of EIROforum IT Working Group.
Front image credit: Shutterstock.com

e-Infrastructure for the 21st Century

Executive summary

Over the past decade Europe has developed world-leading expertise in building and operating very large scale federated and distributed e-Infrastructures, supporting unprecedented scales of international collaboration in science, both within and across disciplines. We have the opportunity now to capitalize on that investment and experience, to build the next-generation infrastructure to enable innovation and opportunities for European science, education, industry and entrepreneurs.

We are now in a period of explosive data growth. The foundations for handling the “Data Tsunami” or “Big Data” have been laid in the last 20 years as we have moved from simple commodity computing (“Farms”), to commodity distributed computing (“Grid”) and then commodity computing services (“Cloud”). These have prepared the ground for handling the large amounts of data being produced today. The era of “Data-Intensive Science” has begun.

To address these challenges for the diverse, emerging “*long tail of science*” conducted by researchers that do not have access to significant in-house computing resources and skills, we propose creating a common platform for the future that builds on the experience of the last decade and is flexible enough to adapt to technological and service innovations. Such a platform must provide the underlying layers of common services, but must be adaptable to the very different and evolving needs of the research communities. A key feature should be that commodity services be operated by European industry, while development of new services may be funded via public-private partnerships. The proposal has three distinct layers of services:

1. European and international networks; services for identity management and federation across all European research and education institutions and integrated with other regions of the world;
2. A small number of facilities to provide cloud and data services of general and widespread usage.
3. Software services and tools to provide value-added abilities to the research communities, in a managed repository:
 - a. The tools to enable those research communities that already have access to large sets of resources to federate and integrate those resources and to operate them for their community, potentially sharing with other communities;
 - b. Tools to help build applications: e.g. tools to manage data, storage, workflows, visualisation and analysis libraries, etc.
 - c. Tools and services to allow researchers to integrate everyday activities with the e-Infrastructure: collaborative tools and services; office automation, negotiated licensing agreements etc.;
 - d. Tools to help research communities engage the general public as citizen scientists.

These layers would be supplemented by investment in application software in order to build and share expertise in ensuring that applications are capable of exploiting evolving computing architectures.

The expectation is that a continuum of financial models is appropriate, ranging from sponsored resources for peer-reviewed scientific cases to communities who would pay for the services they receive. Thus, the services they receive must be appropriate and provide a clear value. The governance of the platform would be created by representation from the user communities to ensure the services provided remain relevant and to build a trustful relationship with service providers. The majority of existing distributed computing infrastructures are supported by public funding and provide services that are free at the point-of-use. The financial support provided by the funding agencies is normally based on a fee linked to the cost of setting-up and operating a service rather than its level of usage. By introducing a pay-per-usage scheme as part of the funding model, the funding agencies will have the information to be able to measure the level of usage of a service and whether it justifies their investments.

The objective of the implementation plan is to put in place the *e-infrastructure commons* that will enable *digital science* by introducing *IT as a service* to the public research sector in Europe. The rationale calls for a hybrid model that brings together public and commercial service suppliers to build a network of *Research Accelerator Hubs* offering a range of services to a wide user base. The exploitation platform will make use of and cooperate with existing European e-infrastructures, including volunteer computing structures, by jointly offering integrated services to the end-user. This hybrid model represents a significant change from the status-quo and will bring benefits for the stakeholders: end-users, research organisations, service providers (public and commercial) and funding agencies. *Research Accelerator Hubs* can be owned and operated by a mixture of commercial companies and public organisations. Their portfolio of services - using as a starting point those listed by e-IRG and the High Level Expert Group on Scientific Data - will be made available under a set of terms and conditions that are compliant with European jurisdiction and legislation, with service definitions implementing recognised policies for trust, security and privacy, notably for data protection. A funding model engaging all stakeholder groups is described.

The ability to fully exploit the potential for knowledge and job creation that is locked-up in the datasets and algorithms to be hosted by the *Research Accelerator Hubs* will require the nurturing of a new generation of *data scientists* with a core set of ICT skills. A management board, where all the *Research Accelerator Hubs* operating organisations are represented, will provide strategic and financial oversight and a user forum, through which the end-users themselves, in a cross - disciplinary body collaborate to define requirements and policies for the services.

A pilot service is proposed that can be rapidly established by building on the existing investments. The pilot service will demonstrate the feasibility of the network of *Research Accelerator Hubs* model for a range of scientific disciplines and evaluate the suitability for the ESFRI Research Infrastructures that are currently under-development and represent Europe's future "*big data factories*". Implementation will start in 2014, initially offering a limited set of services at several prototype *Research Accelerator Hubs*.

Table of Contents

Executive summary.....	2
Vision	6
Proposal for an e-Infrastructure	7
A set of basic infrastructure services:	8
Cloud services.....	8
Data facilities	8
Distributed infrastructure.....	9
Software services and tools.....	9
Investment in software.....	9
Relationship with the HPC community.....	10
Building the data continuum.....	10
Providing leadership.....	10
Funding models	11
Governance.....	11
Implementation	12
Rationale.....	12
Benefits of the hybrid model	15
Architecture of the e-Infrastructure commons	16
Pilot service	17
Portfolio of services	18
Training and education.....	18
Integration of public and commercial service suppliers into a hybrid model .	19
Governance model	19
Relationship to existing European e-Infrastructures.....	20
GEANT.....	20
PRACE	20
EGI.....	21
Volunteer computing.....	21

EUDAT data services.....	21
Engagement of funding agencies	22
Prototype Research Accelerator Hubs.....	24
EMBL-EBI as a Research Accelerator Hub.....	24
CERN and Wigner as a Research Accelerator Hub	25
Funding models for the prototype Research Accelerator Hubs	28
Timeline.....	29
Governance.....	30
The problem.....	30
e-Infrastructure activities.....	31
Membership	32
Objectives.....	33
Deliverables.....	34
Recent examples of user coordination.....	35
FIM4R.....	35
Helix Nebula	35
Joint development with industry.....	35
Relationship with other activities.....	36
Research Data Alliance	36
e-Infrastructure Reflection group (e-IRG).....	36
Process	37

Vision

Looking forward over the coming 10-15 years, there are exciting challenges ahead to capture, manage, and process the vast amounts of data likely to be generated, not only by the established fundamental research domains, but in a growing range of scientific disciplines, large and small. This new frontier in computation will be pushed forward by the needs of data-driven science, simulation, modelling and statistical analysis in areas from climate change to life sciences, art and linguistics. All will see incredible growth and accelerated breakthroughs due to unprecedented access to data and the computational ability to process it.

Europe must preserve its intellectual capital and provide the opportunity for it to be nurtured, developed and to grow. Major advances in technology that have taken the world by storm, from Linux to the World Wide Web, have often been conceived in Europe but ultimately exploited elsewhere. This is a loss to Europe, in terms of skills, employment and business.

We must take the step to make unprecedented scales of IT resources available to the next generation of emerging scientists, researchers and entrepreneurs, nurturing them from education to start-up activities and then sustainable businesses or research communities. As well as serving the direct needs of computing, the opportunity to innovate and explore new technologies is essential. The correct environment for innovation will allow many ideas to be tested, and explored, which will lead to the truly unexpected and ground-breaking discoveries and inventions that will shape this century.

Over the last decade, driven with sustained funding from the EC, the e-Infrastructure landscape across Europe has grown from regional prototypes to a set of pan-European production resources, including EGI, GEANT and PRACE. But to go forward much more coordinated effort is needed. Today's efforts leave gaps in the overall strategy, and suffer in part from inadequate funding by the stakeholders.

CERN, in collaboration with the EC, national funding agencies and the High Energy Physics community, has successfully built - and today operates - the world's largest scientific e-Infrastructure. This worldwide infrastructure is in daily operation and has been used to produce results from the huge volumes of data delivered by the LHC and its detectors. The development of this distributed grid federating resources around the world took close to 10 years from conception through to production use at the necessary scale, and required novel developments in terms of physical infrastructure, middleware, application software, and policy development. This experience also highlights a key aspect of the future e-Infrastructure model if it is to become *the* infrastructure of choice for the European Research Area: there must a long-term commitment by all the stakeholders to make the e-Infrastructure the means by which they will provide/use production IT services. The future research infrastructures currently in construction, such as FAIR, XFEL, ELIXIR, SKA, ITER and upgrades to ILL and ESRF, need to be convinced that the e-Infrastructure will exist and continue to

evolve throughout their construction and operation phases if they are to take the risk and invest in its creation and exploitation.

In considering the way forward, it is important that we foresee an infrastructure that supports all of the scientific and academic needs of the European community, including the “*long tail of science*” conducted by researchers that do not have access to significant in-house computing resources and skills. Consequently, this should not be thought of as a one-size-fits-all solution; rather as a broad but coherent set of services and tools which must be available to allow the specific needs of each community to be met. This common platform should also be able to act as an incubator for new businesses and scientific activities. It is essential that European industry engage with the scientific community in building and providing such services, but it is also important that the user community has a strong voice in the governance. This view has been documented in a recent Response to EC (DG CNECT) Paper “Research Data e-Infrastructures: Framework for Action in H2020” produced by the EIROforum IT Working Group.

Proposal for an e-Infrastructure

While the grid model has been extremely successful for High Energy Physics and similar high-throughput computing applications (such as astrophysics), it is not suited for many other sciences, which have very different requirements. Technological advances are continuous. Thus, during the time of the development of the grid, distributed computing and the underlying technologies have advanced significantly in academia and have been adopted by many business sectors. Cloud computing technologies, and the huge increase in available networking capabilities are leading examples. The growing computing needs of sciences, in particular those that have never before needed large-scale computing, will benefit from many of these advances. Thus it is vital for the future needs of scientific e-Infrastructures, that a model be adopted encompassing a wide spectrum of facilities and tools that can be of direct benefit to a range of science and research use cases. Rather than build such a structure as a single integrated e-Infrastructure (like the grid) it will be far more advantageous to provide a set of collaborating core infrastructure services, and a variety of facilities, together with a broad set of easily adaptable tools. This will allow the research communities to select the services or tools that they require, and only those, without additional complexity. One of the lessons from the grid experience is that unnecessary complexity should be avoided in the infrastructure layers.

There are many existing efforts within Europe that can be drawn on to fulfil a vision for the future. These “pathfinder” initiatives have prototyped many aspects of what will be needed in the future. This includes much of the work in the grid projects, but also projects such as EUDAT, CRISP, Helix Nebula and OpenAIRE, as well as thematic data projects, such as Transplant and many others.

In order that such an effort be sustainable and permit maximum flexibility across domains, as well as being able to fulfil the goals of working together with industry and key global players, it is essential that the future service infrastructure and tools be fully based on open standards, open software, and promote open access to the data.

This integrated model representing a common platform for the future must have the following key components:

A set of basic infrastructure services:

- The core network; building on today's GEANT/NREN and commercially operated networks to provide excellent connectivity and operational services to all scientific institutions, and fully extending to countries on the edges of Europe, as well as ensuring the international and global connectivity required by today's sciences;
- Federated identity management services, allowing existing identities of researchers to be used across the full set of e-Infrastructure services. Bearing in mind the substantial achievements of the ORCID project, it must support persistent digital identifiers that uniquely distinguish every researcher throughout their entire career, providing integration in key research workflows such as manuscript and grant submission, supporting automated linkages between professional activities and ensuring the researcher's work is recognised. The FIM4R document¹ produced by representatives from a range of research disciplines provides detailed requirements for such services.

Cloud services

A (small number of) publicly provided research community cloud facilities that can be used for applications that require the instantiation of a few long-lived services, or access to compute or storage resources for a relatively short time. Such a cloud resource could be outsourced to commercial providers, or be the product of a public-private partnership. The Helix Nebula project provides an example for what may be involved in putting this in place.

Data facilities

A general data storage facility (for example, public science archives). Such a resource would be of immense value to data producers, providing a sustainable, dependable and accessible archive. It would also provide an unrivalled opportunity for data sharing between sciences with the integration of different types and sources of data. The EUDAT project is demonstrating some of the candidate technologies in this area. These facilities would offer open access to the data, and would be a focus for data preservation activities to ensure the long-term guardianship of the data. The facilities would also provide persistent identifiers for data objects at an appropriate granularity, as well as metadata services. They would allow secure data sharing between sciences, so that a range of data-access policies can be supported, including certification requirements of Data Access Committees. As such, the e-infrastructure will offer a knowledgebase consisting of a collection of data, organizational methods, standards, analysis tools, and interfaces representing a body of knowledge.

¹ Federated Identity Management for Research Collaborations, <http://cds.cern.ch/record/1442597>

Distributed infrastructure

A set of high-level software services that allow research communities to implement a federated and distributed computing infrastructure in order to integrate resources often explicitly provided for those applications. These are typically useful where the computing and storage requirements are large, where there is a need to collaborate, and where the occupation of the resources is very high. These services would be a generalisation of today's grid services, but should focus on the move to more open and standard implementations, and may benefit directly from cloud implementations. By democratizing access to data and computational resources, the services will enable any laboratory or project, regardless of size, to participate in a transformative community-wide effort for advancing science and accelerating the pace toward its exploitation. Thus, the e-Infrastructure services will facilitate building a broader scientific community that will contribute to fundamental science within the European Research Area.

Software services and tools

1. A set of software services that allow researchers to integrate e-Infrastructures with their everyday activities and personal devices, for example a "dropbox" functionality, office automation and collaborative tools and services. Many of these would be hosted on the infrastructures described above. Today there are many tools available, but most are not widely known or used. This action would also provide a means by which software licenses (which today represent a significant and rapidly growing cost) could be potentially negotiated and managed on behalf of the entire scientific community.
2. A set of tools of wide and general utility that can be used by the applications. This would include a set of tools to manage data transfer, storage, and other data related activities, as well as coordinating a repository for useful software tools. The ideas outline by SciencePAD² could play a role here.
3. A set of tools to allow scientific communities to build a citizen-cyberscience facility where that is appropriate and useful.

Investment in software

In addition to the above, an organisation put in place to coordinate such a set of coherent services and activities would also be the natural way to broker new collaborations. One area that will be of strategic importance in the coming years will be a significant investment in software capability that will be absolutely essential to obtain the best performance from current and future computer and storage architectures. Many sciences today benefit from commodity CPU and storage, and this is likely to change as the consumer market shifts from PC's to tablets and smartphones. This investment in software is essential to maintain European competitiveness in this area, and should include coordination of existing expertise to the benefit of diverse communities.

There may also be traditional software and tools at the application layer that would benefit from a European-wide collaboration. Examples here may include a mechanism

² <http://www.sciencepad.org>

to obtain better licensing conditions, or collaborations to build specific application software of general benefit to a broad community.

Relationship with the HPC community

The relationship with the supercomputing community (HPC applications) should also be re-defined. One aspect to consider is the frontier-science challenges that need the most significant HPC resources. In this case, the HPC facilities should be viewed as scientific instruments in their own right that produce science data for their application communities. Today, such large-scale simulations produce huge volumes of data. Those application communities are then naturally users of an e-Infrastructure on which to distribute and analyse their (supercomputer-produced) data, and those applications would also be scientific stakeholders in a general scientific e-Infrastructure.

There are other aspects of HPC facilities that are complementary to the cloud and high-throughput resources. Some applications that require modest levels of an HPC resource may well be deployable on suitably configured cloud- or data-intensive computing resources. There are also workflows that cross HPC and data-intensive computing resources and would benefit from an integrated service environment. The HPC facilities and their scientific communities would also benefit from the underlying technologies mentioned above (the networks, federated identities, policy work, etc.).

Building the data continuum

A data continuum, that is to say a system capable of navigating the data evolution by linking the different stages of the data lifecycle, from raw data to publication is necessary to accelerate the rate of scientific discovery and increase the impact of research on society. Elements of the data continuum exist and a range of projects, including openAIRE³ where CERN provides the Invenio⁴ software technology that supports this open access repository and many more around the world, have created repositories for initially, publications, and now extending to data that can give good examples of what can be achieved. But these remain independent projects and have not been integrated into the overall e-Infrastructure landscape.

Providing leadership

In order to build such a long-term and broad-scope e-Infrastructure to benefit the entire European community, we must leverage the tremendous assets that have been built up during the last decade: in particular the knowledge and skills, as well as the working prototypes of each of the core services noted above.

What we envisage is a continuum addressing the needs of education and speculative innovation through to growing and established entities. There must, therefore, be a range of infrastructure and services to satisfy the needs of such a broad range of maturity of activities. This continuum must cover the different axes of financial models (user-pays, provider-pays) as well as infrastructure (industry supplied and in-house).

³ <https://www.openaire.eu/>

⁴ <http://invenio-software.org/>

User representation in governance is paramount and we envisage a user activity for e-Infrastructures as a well-defined activity.

The development of new and novel services and software may be funded from any combination of public and private sources, but as these become production ready and commoditised they should move into the industrial service operation. Open source and open standards are essential to ensure maximal adoption, and to allow new entrants to be able to leverage the innovation made with public funding. The platform should also enable business innovation for new services and new users to create wealth and employment.

Funding models

Past experience has shown that as the number of communities and activities that could benefit from European e-Infrastructure continues to grow and evolve, there is no “one-size-fits-all” solution that is appropriate.

Within the lifecycle of a given activity, it may be appropriate to have supplier-financed resources for a time and then some subsidised resources as the activity evolves. Fully paid resources by the activity may come at a later stage. In terms of e-Infrastructures, it is important that activities can rely on the resources and services so it is critically important that timeframes for the e-Infrastructures be long and the funding stable.

Where it is expected that users should pay for the services, those services must be relevant and attractive in order for that to happen. The added value of the proposed services must be made absolutely clear. In order to remain attractive to the user communities, the offering must evolve and adapt to the changing needs. A trustful relationship must be established between the users and service providers. This evolution can begin with a set of managed federated services that are recognised as a common need, with new services being added as commonalities are explored and understood. The proposed platform must provide solutions in a timely and relevant way. This is another reason why the user community stakeholders must be directly involved in the overall governance.

A model for moving from innovative development (where a funded activity may develop a service) to industrial operation is essential to avoid unproductive competition between publicly funded activities and commercial offerings. Organisations such as CERN have specialist knowledge of large-scale tendering and coordination of tendering and brokering that would enable cost efficiencies to be gained through the application of scale.

Governance

It is important that a future European e-Infrastructure be driven by the scientific stakeholders. Some key strategic research communities could be selected that would drive the frontiers of the technologies in several different but complementary aspects.

Implementation

The objective of the implementation plan is to put in place the *e-Infrastructure commons* that will enable *digital science* by introducing *IT as a service* to the public research sector in Europe.

This implementation plan establishes a pilot service aimed at addressing the challenges ahead to capture, manage, and process the vast amounts of data to be generated, not only by the established fundamental research domains, but in a growing range of scientific disciplines, large and small. This pilot service is seen by the research organisations as an exploratory activity to be operated and funded in addition to their baseline plans for the production services they offer to their user communities.

Rationale

E-infrastructures have made remarkable progress over the last few years, moving from bespoke installations for single applications to distributed structures serving many user communities. In order to fulfil the vision for the future these e-infrastructures must be brought together to provide integrated solutions to a wider range of users during all stages of the research life-cycle. The justification, or rationale, to push for the implementation of this vision is that the impact of combining the e-infrastructures to form an *e-Infrastructure commons* will generate far greater added value for all the stakeholders. This rationale has been inspired by the Helix Nebula initiative's publication "*Helix Nebula – The Science Cloud: A catalyst for change in Europe*"⁵ from which extracts have been taken.

Helix Nebula was conceived as a way of bringing coherence to a highly fragmented IT services industry through the vision of a federated 'science cloud' integrated with publicly-funded scientific e-Infrastructures. Two fundamental steps remain if the vision of Helix Nebula is to become reality. One is to bring together a critical mass of supply-side interests and the other is to do the same for the demand-side.

The work of Helix Nebula on an architecture model⁶ has produced a number of scenarios exploring the means by which publicly funded infrastructures can interoperate with commercial cloud services⁷. Such hybrid systems are in the interest of the users of publicly funded infrastructures and funding agencies because they will provide '*freedom and choice*' over the source of resources to be consumed and the manner in which they can be obtained, facilitating the goal of '*making every researcher digital*'. Hybrid systems serve the interests of suppliers by encouraging the adoption of cloud services by more researchers and thus creating a larger and more vibrant marketplace in which they can offer their services.

⁵ <http://cds.cern.ch/record/1537032/files/HelixNebula-NOTE-2013-003.pdf>

⁶ <http://cdsweb.cern.ch/record/1478364/files/HelixNebula-NOTE-2012-001.pdf>

⁷ http://cds.cern.ch/record/1548323/files/HelixNebula-D6_1.pdf

This integration will allow the public infrastructure users to strike their own balance between publicly funded resources and commercial cloud services while taking into account aspects of policy and cost. Private sector service providers can supply resources/services and additional capacity not available in the publicly funded infrastructures. Research and innovation activities that have the potential for commercial exploitation can work with the private sector to unlock that potential, demonstrate feasibility and market acceptance. It is expected that services will migrate from publicly funded infrastructures to commercial suppliers as suitable offers, addressing cost and policy aspects, become available on the market. In this manner, Helix Nebula is implementing what the e-Infrastructure Reflection Group (e-IRG) refers to as the *e-Infrastructure Commons* in its 2012 Roadmap paper⁸.

This hybrid cloud has the potential to bring the public and private sectors together, so that the public sector has something to offer the private sector beyond a simple demand for cloud services. The creation of a continuum of services across the public and private sector, instead of a simple supply-demand model will alleviate the current situation which is inhibiting the establishment of hybrid cloud infrastructures. It is also inhibiting the establishment of big-data services, since there are a number of sensitivities (explained below) that need to be addressed before a model can be envisaged where all datasets are hosted on commercial cloud services.

Business models need to be developed that ensure the Data Stewards⁹, which are frequently publicly funded research centres, retain control over the curation and usage of the datasets and the associated intellectual property. Researchers also need guarantees about the long-term availability of the cloud services and the data. Such a hybrid model would address these points and allow the progressive migration of services between public and private providers. Furthermore, it is of utmost importance that data catalogue services, such as those being developed by the EUDAT project¹⁰, are available to all research communities' datasets hosted by both public and private cloud services.

⁸ http://www.e-irg.eu/images/stories/publ/e-irg_roadmap_2012-final.pdf

⁹ http://en.wikipedia.org/wiki/Data_governance

¹⁰ <http://www.eudat.eu/>

There exist several institutes in the biomedical sciences with well-established data infrastructures, such as the European Bioinformatics Institute and the Swiss Institute of Bioinformatics. The new, distributed infrastructure ELIXIR will take data management, curation and interoperability to the next, Europe-wide level. ELIXIR, in turn, is coordinating the FP7-funded BioMedBridges project, which aims to integrate data and thus construct computational 'data and service' bridges between the new biomedical sciences infrastructures on the ESFRI roadmap, including BBMRI, INSTRUMENT and ECRIN. Together these span a wide range of disciplines from the biological, medical, translational and clinical domains. Their diverse data management needs can be best catered for by a combination of domain-specific services that are integrated into a wider ecosystem of domain-agnostic services. Other community-specific research data infrastructures have been established in recent years, such as The Catalogue of Life¹¹, iMARINE¹² and GENESI-DEC¹³, which have produced valuable data-curation tools and expertise, along with data-sharing policies. Being able to interface these new data e-infrastructures into a hybrid cloud model, as proposed here, will allow a larger user base to exploit the data, bigger opportunities to contribute scientific data to multi-disciplinary research, and provide sustainability models for their continued existence.

Initial investigations on the potential impact of cloud services in the research community suggest that the commercial public cloud services are likely to be adopted initially for the 'long tail of science' conducted by researchers that do not have access to significant in-house computing resources and skills. Efforts must be made to simplify access to commercial cloud services for such groups that generally have straightforward requirements and frequently do not have sufficient in-house IT expertise to manage and operate their own computing resources.

Conversely, it will require further reduction in the costs of commercial services for large research users with important in-house computing capacity to find commercial cloud services as financially attractive as is the case for small scale-users. As one industry representative put it 'why would I hire a car on a daily basis if I know in advance I will use it every day for 3 years – it will be cheaper to buy my own'¹⁴.

It is recognised that not all publicly funded research centres are in a position to make accurate estimations of the total cost of ownership of in-house IT services since some contributing costs are borne by different departments. But in order for the demand-side users to be encouraged to purchase cloud computing services, the services offered must be economically advantageous compared to other means of procuring IT services.

These alternatives include purchasing and operating IT equipment internally which requires capital investment and IT expertise but remains economically attractive for IT-intensive applications with a sustained and predictable usage.

¹¹ <http://www.catalogueoflife.org/>

¹² <http://www.i-marine.eu/Pages/Home.aspx>

¹³ <http://www.genesi-dec.eu/>

¹⁴ Similar cost arguments are made in this article: http://www.theregister.co.uk/2013/04/26/cloud_magic_number/

Benefits of the hybrid model

The hybrid model outlined in the rationale above represents a change from the status-quo but is necessary to kick-start new and emerging sciences use of e-infrastructures. The identified benefits of this model for the stakeholders include:

Users

- A wide range of integrated services and data resources providing a more complete working environment supporting the full lifecycle of research activities
- Easy means to publish their own research results and data
- Freedom and choice over which services to use according to preferences for functionality, policy and budget
- An environment into which their own bespoke services can be integrated
- Readily available training material and support
- Time to focus on their research objectives rather than having to learn about and dedicate time to operate IT infrastructure
- Keep control over their results and data with the assurance that they will be available over the long-term
- Have a voice on how the services they use are governed

Research organisations

- A means to expose their research products to a wider community
- An exploitation platform through which research organisations can be more closely connected to their users and react more quickly to their changing needs
- Ability to procure additional resources on-demand
- Grouped negotiation of service procurement with other research organisations
- A ready-made channel for improving knowledge and technology transfer with industry

Service providers (public and commercial)

- Creation of a market with an extensive user base in to which providers can offer their services
- Privileged communication channel with end-users and purchasing organisations
- A means of estimating the scale and sophistication of services that would be suitable for the public research sector
- Reduced risk and up-front investment during the development and introduction of innovative services

Funding agencies

- A means of encouraging the uptake of cloud services by the public sector
- Creation of a public-private innovation platform for ICT
- A simple means for implementing open access policies of publications and data
- Democratisation of access to data and compute resources, so that they are available to any users, laboratory or project, regardless of size and location
- A channel for increasing the impact and exploitation of publicly funded research
- A means to monitor the impact of public investments

Architecture of the e-Infrastructure commons

The e-infrastructure vision proposes a common platform for the future that builds on the experience of the last decade and is flexible enough to adapt to technological and service innovations. Such a platform must provide the underlying layers of common services, but must be adaptable to the very different and evolving needs of the research communities. The proposal has three distinct layers of services:

1. European and international networks; services for identity management and federation across all European research and education institutions and integrated with other regions of the world;
2. A small number of facilities to provide cloud and data services of general and widespread usage.
3. Software services and tools to provide value-added abilities to the research communities, in a managed repository:
 - a. The tools to enable those research communities that already have access to large sets of resources to federate and integrate those resources and to operate them for their community, potentially sharing with other communities;
 - b. Tools to help build applications: e.g. tools to manage data, storage, workflows, visualisation and analysis libraries, etc.
 - c. Tools and services to allow researchers to integrate everyday activities with the e-infrastructure;
 - d. Tools to help research communities engage the general public as citizen scientists.

Figure 1 shows how the layers of services are brought together into an overall architecture. A key element is the *Research Accelerator Hub (ReACH)* which is a centre that offers a set of managed IT services tailored to the needs of their users. Together the *Research Accelerator Hubs* form a network of interoperating centres that can support workflows addressing the full research life-cycle. The underlying high-performance network brings together the *Research Accelerator Hubs* with the other elements of the *e-infrastructure commons*. The high-performance network extends to scientific instruments which act as big-data factories and additional data providers hosting important datasets around the world. Supercomputing sites connected to the network have a dual role: like scientific instruments they can be seen as big-data factories but they also offer capability-style data analysis facilities for datasets held at the *Research Accelerator Hubs* or provided by external data providers. Volunteer computing structures are an integral part of the *e-infrastructure commons* providing significant computing capacity, as well as channels for engaging citizen scientists and the general public. All the elements of the *e-infrastructure commons* share a federated identity management system enabling single-sign-on access for users to any resource for which they have authorisation. A common digital object identifier service promotes the citation and re-use of any publication or dataset in the *e-infrastructure commons*.

E-infrastructure commons

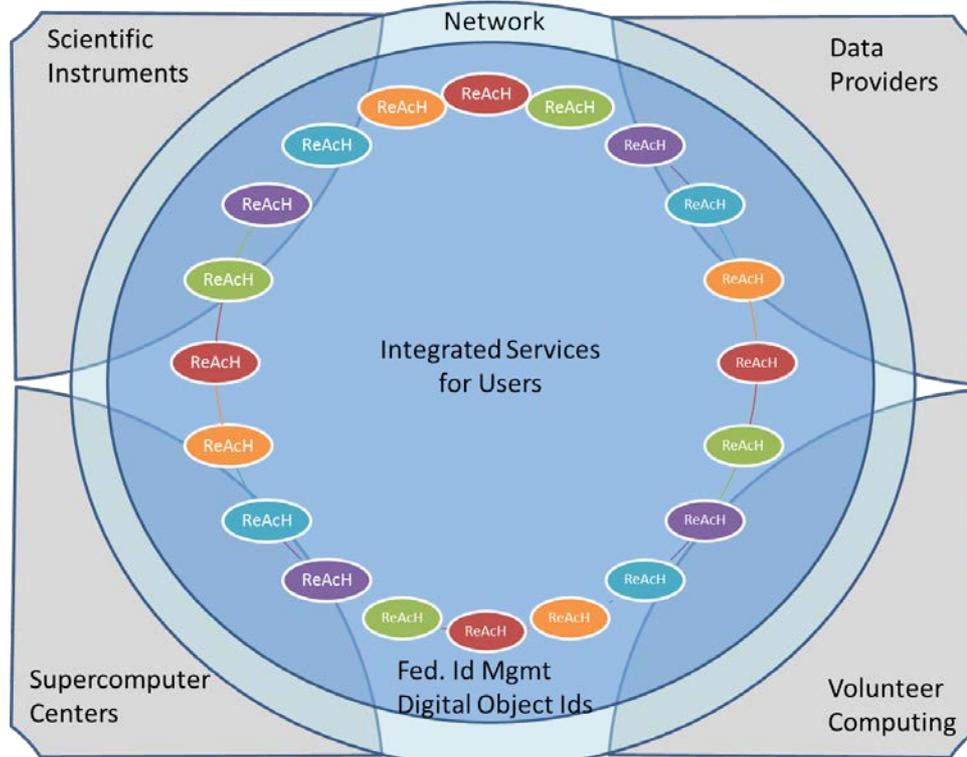


Figure 1 e-Infrastructure commons architecture

Pilot service

In order to verify the feasibility of this vision and rationale, a pilot service is proposed that can be rapidly established by building on the existing investments to demonstrate added value to the research community and the sustainability of the approach.

Implementation will start in 2014, initially offering a limited set of services at prototype *Research Accelerator Hubs*. This ramp-up phase will include a roadmap for how to extend the e-infrastructure to a group of independently operated and cooperating *Research Accelerator Hubs* offering a richer set of layered services and higher total capacity to serve the European Research Area. Such facilities will be *Research Accelerator Hubs* providing capacity style High Performance Computing (HPC)¹⁵ services for data centric applications and will be federated into a network of such centres across Europe. The *Research Accelerator Hubs* will not be duplicates, but rather offer a set of services tailored to their users' specific needs and business models.

¹⁵ For a definition of HPC capacity computing refer to the ETP4HPC Strategic Research Agenda document http://www.etp4hpc.eu/wp-content/uploads/2013/06/ETP4HPC_book_singlePage.pdf

Portfolio of services

The *Research Accelerator Hubs* will offer an initial portfolio of services taken from the list documented by e-IRG in its blue paper of 2010¹⁶ with the technical characteristics identified by the High Level Expert Group on Scientific Data in their “Riding the Wave” report from the same year¹⁷.

While each *Research Accelerator Hub* will offer their own set of services, all such services will incorporate state-of-the-art security measures, including host-based protections, system logging and access traceability. All services will make use of a federated identity management system addressing authentication and authorization to ensure seamless and secure access to the services and data. While the *Research Accelerator Hubs* will take the role and responsibilities of data custodians¹⁸, the stewards of the data sets hosted by the *Research Accelerator Hubs* will be responsible for defining and managing the access policy for their data. To ensure a trustful relationship is established with the users, it will always be possible for stewards to extract or remove any material that has been deposited with the *Research Accelerator Hubs*.

Brokering services will be deployed so that the full portfolio of services from all *Research Accelerator Hubs* is visible to all users and can be compared and accessed.

Training and education

The design, creation and operation of e-infrastructure services are essential tools in the development of skills and competencies for the European market. The services offered by the network of *Research Accelerator Hubs* will provide a focus for the creation of public-private teams of skilled personnel across the cooperating research organisations, service suppliers and users, enabling the sharing of knowledge to develop a co-design approach¹⁹ to services and applications. The ability to fully exploit the potential for knowledge and job creation that is locked-up in the datasets and algorithms to be hosted by the centre will require the nurturing of a new generation of *data scientists* with a core set of ICT skills. To ensure the training and education services impact as large an audience as possible while controlling the costs of offering such services, massive open online course (MOOC²⁰) techniques will be employed, which will be supported by the *Research Accelerator Hubs* model itself via the digital library services. The EIROforum organisations have core competences in training and education²¹ which can form the basis of this activity. Since its formation in 2002, EIROforum joint initiatives have already resulted in an important impact in the areas of outreach and education. High-profile activities have involved thousands of European science teachers and students, thus emphatically supporting Europe-wide efforts to raise interest in science and technology and to secure a sound recruitment base for European R&D efforts in the

¹⁶ http://www.e-irg.eu/images/stories/eirg_bluepaper2010_final.pdf

¹⁷ <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

¹⁸ http://en.wikipedia.org/wiki/Data_governance

¹⁹ http://echallenges.org/e2010/outbox/eChallenges_e2007_ref_195_doc_3562.pdf

²⁰ http://en.wikipedia.org/wiki/Massive_open_online_course

²¹ <http://www.scienceinschool.org/>

future. Such collaboration in the area of human resources also has a bearing on the European Research Area: by attracting and retaining world-class researchers in Europe, through technological exchange and common development, and by organising multi-disciplinary scientific conferences.

Integration of public and commercial service suppliers into a hybrid model

The investigations within Helix Nebula have revealed that with current pricing of commercial services many public organisations will continue to find it financially attractive to provide the services themselves in-house. Hence, this implementation plan foresees that *Research Accelerator Hubs* can be owned and operated by a mixture of commercial companies and public organisations. Each *Research Accelerator Hub* will include a pay-per-usage element in its funding model, but the operating organisation is free to define its own pricing strategy and implementation. The *Research Accelerator Hubs* are unlikely to have the same funding models because legislation and regulation for e-infrastructure, including aspects such as state-aid, procurement, data protection, and software licensing, vary significantly across Europe. The portfolio of services will be made available under a set of terms and conditions that are compliant with European jurisdiction and legislation with service definitions implementing recognised policies for trust, security and privacy notably for data protection.

By integrating commercial and publicly operated *Research Accelerator Hubs* into a hybrid model it will be possible to migrate workloads to commercial suppliers, once acceptable pricing and terms and conditions have been negotiated. In this way, the *Research Accelerator Hub* model will bolster the creation of a market to which the commercial suppliers can offer their services.

Governance model

The governance model foreseen for the pilot service will be based on two bodies:

- A management board where all the *Research Accelerator Hubs* operators are represented and the strategic and financial aspects of the *Research Accelerator Hub* model are addressed. Policy aspects impacting the operation of the network of *Research Accelerator Hubs* will be addressed by the management board. The introduction of a new *Research Accelerator Hub* into the network will be approved by the management board subject to verification of the policies (see Integration of public and commercial service suppliers) under which it operates and the quality of the services provided.
- A user forum²², through which the end-users themselves, in a cross-disciplinary body collaborate on requirements and policies concerning the e-Infrastructure services. The User Forum will also act as a consultation body for the management board.

²² “Science Strategy and Sustainable Solutions; A Collaboration on the Directions of e-Infrastructure for Science”, CERN-OPEN-2013-017, <http://cds.cern.ch/record/1545615/files/CERN-OPEN-2013-017.pdf>

The pilot service will demonstrate the feasibility of the e-infrastructure *Research Accelerator Hubs* model for a range of scientific disciplines. The pilot service will also provide a means for the ESFRI Research Infrastructures that are currently under-development and represent Europe's future "*big data factories*", to evaluate the suitability of the services provided by *Research Accelerator Hubs*. The integrated service portfolio offered by the network of *Research Accelerator Hubs* will enable multi-disciplinary research across the Research Infrastructures and contribute to ensuring excellence in science.

Relationship to existing European e-Infrastructures

The *Research Accelerator Hubs* will make use of and cooperate with existing European e-infrastructures to build the e-infrastructure commons by jointly offering integrated services to the end-user. The relationship with these existing e-infrastructures will be defined via written agreements and is described below.

GEANT

The *Research Accelerator Hubs* will be connected to the GEANT network in order to provide high performance access for the whole European Research Area and ensure integration with other regions of the world, so that Europe can be a leading participant in global research challenges. Building on the experience gathered through the Helix Nebula initiative, which is extending GEANT connectivity to commercial data centres, will greatly simplify the establishment of a hybrid model engaging both public and commercial service providers.

PRACE

The capacity-style HPC services for data centric applications offered by the *Research Accelerator Hubs* will be complimentary to the capability style HPC services provided by PRACE. The expertise developed by PRACE and related projects in efficient parallel programming paradigms and optimising software for a range of architectures is directly relevant to the e-Infrastructure *Research Accelerator Hubs* and application/service developers. The European Technology Platform for High Performance Computing project²³ recently published a Strategic Research Agenda for achieving HPC leadership in Europe²⁴, which specifically highlights the upcoming big-data challenges for leading research activities and the relevance of cloud services:

*"Europe is in a unique position to excel in the area of **HPC Usage and Big Data** owing to the experience level of current and potential users (and the recognition of the importance of data by such users as CERN, ESA, and biological data banks) and the presence of leading ISVs for large-scale business applications. Europe should exploit that knowledge to create competitive solutions for big-data business*

²³ <http://www.etp4hpc.eu/>

²⁴ http://www.etp4hpc.eu/wp-content/uploads/2013/06/ETP4HPC_book_singlePage.pdf

applications, by providing easier access to data and to leading-edge HPC platforms, by broaden the user base (e.g., through Cloud Computing and Software as a Service (SaaS), and by responding to new and challenging technologies.”

So, as well as sharing expertise, the services offered by the *Research Accelerator Hubs* and the PRACE Tier-0 and Tier-1 centres should be federated to form part of the overall e-infrastructure ecosystem. This will require the PRACE HPC centres to participate in the federated identity management scheme and data sharing services if the PRACE centres are to be fully federated as service providers within this model.

EGI

The experience gathered by EGI in managing a federated grid infrastructure will be directly relevant to the network of *Research Accelerator Hubs* model. EGI has also been evaluating cloud technologies via the EGI federated cloud²⁵. It is proposed that a consolidated set of EGI sites become *Research Accelerator Hubs*. This will give the EGI distributed computing infrastructure a clear direction for how to contribute its experience and make a larger portfolio of services accessible to its existing user-base, while introducing the innovation potential created by the uptake of cloud computing in research and business sectors.

Volunteer computing

Volunteer computing initiatives across Europe have established production structures which serve a range of research communities. Such structures allow research and education organisations, as well as individuals and citizen scientists, to contribute and participate in research activities. Significant computing resources are assembled by structures such as the International Desktop Grid Federation²⁶ that can support a growing range of application types with very modest operational and coordination overheads. It is important that such structures become an integral part of the e-infrastructure commons.

EUDAT data services

It is expected that data services currently under development by various projects, notably EUDAT, will provide candidates for future services and potentially additional e-infrastructure *Research Accelerator Hubs*. A goal will be to introduce services that can profit from the co-location of data and compute services to support multi-disciplinary research. Metadata and indexing facilities across the set of services in all the e-infrastructure *Research Accelerator Hubs* are seen as being particularly relevant. It is essential that new services are fully integrated with existing services to preserve the data and compute continuum of the exploitation platform and support the *e-infrastructure commons*.

²⁵ <https://wiki.egi.eu/wiki/Fedcloud-tf:UserCommunities>

²⁶ <http://desktopgridfederation.org>

Engagement of funding agencies

The implementation plan and associated funding model have been designed so that the e-infrastructure *Research Accelerator Hubs* can be sustained by their operating organisations according to a continuum of financial models, ranging from sponsored resources for peer-reviewed scientific cases to communities who would pay for the services they receive. Additional resources will be required in order for these services to be expanded and to serve a wider range of users. The European Commission and national funding agencies will be invited to become stakeholders and contribute to the expansion of the e-infrastructure *Research Accelerator Hub* model. The guiding principle is that funding from such stakeholders will be focused on innovation of services and uptake by new user communities and business actors, while the operational costs will be borne by the operating organisations and the users themselves. Below is a non-exhaustive list of areas where funding agencies may contribute to the expansion of the *Research Accelerator Hub* model:

- Development of new services to be deployed on the e-infrastructure. Significant effort will be required to co-develop scalable services that can operate in a distributed environment and serve a wide range of users.
- Financial incentive scheme to increase adoption of services (both public and private) by users including 'long tail of science' research groups and SMEs.
- Engaging the use of the services by new research communities (e.g. curation of data-sets, connection of identity federations, deployment of community specific services, training for new users, etc.)
- Develop training and educational activities building on the e-infrastructure services to maximise their impact. This can also include expansion of services to engage citizen-cyberscience communities and the general public in science.
- Organisation of user forum events as well as outreach and dissemination to a range of audiences and production of material for policy related activities.
- International collaboration (beyond Europe) through interoperation with equivalent structures in other regions of the world.
- Expansion of the network of the *Research Accelerator Hubs* across the European member states to address national and thematic needs.
- Many research organisations that operate research infrastructures do not have the mandate to provide e-infrastructure services to their users for the management and processing of their experimental data. This represents a gap in the scientific lifecycle and a missed opportunity to highlight the results and impact of public funded research. These research organisations will require assistance to bridge this gap by supporting their users so they can make use of e-infrastructure services to manage and process their experimental data.

The recently completed eInfraNet project²⁷ brought together national funding agencies²⁸ from across Europe to explore the economics of e-Infrastructures and align

²⁷ <http://e-infranet.eu/>

²⁸ <http://e-infranet.eu/partners/>

European policies with the needs of the member states. In its final report²⁹ the project recommended:

“The use of standards-based private cloud infrastructures may also provide the needed flexibility in case of jobs that exceed available resources through cloud bursting (the use of external public clouds for the exceeding capacity) or by pooling individual institutional private clouds into a federated cloud that can provide the necessary resources in a federated way, following the extremely successful approach already in use for GRIDs.”

The model for a network of *Research Accelerator Hubs* described in this document can serve as a blueprint to be expanded across Europe. The establishment of *Research Accelerator Hubs* could be supported by national funding agencies to address local or thematic needs and provide a focus for innovation in e-Infrastructures. Such national or thematic *Research Accelerator Hubs* should be integrated with existing European e-Infrastructures (see Relationship to existing European e-Infrastructures) and will be in a position to host data, information and services that for policy reasons need to remain within a national context. As these *Research Accelerator Hubs* come online they must be federated together to provide a European dimension to the *e-infrastructure commons*.

²⁹ Cloud Computing Economics: An evidence-based approach for Research Applications <http://e-infranet.eu/wp-content/uploads/2013/02/Cloud-Computing-Economics-An-evidence-based-approach-for-Research-Applications1.pdf>

Prototype Research Accelerator Hubs

A network of *Research Accelerator Hubs* will be required to satisfy the needs of the public research sector. Here we outline two EIROforum prototypes that can act as “pathfinders” for the overall model and pilot service. It is proposed to use the resources installed by CERN at the Wigner Research Centre for Physics in Budapest, Hungary³⁰, and by EMBL-EBI at Teletcity’s Powergate site in London, UK, to develop prototype *Research Accelerator Hubs* for the pilot service.

EMBL-EBI as a Research Accelerator Hub

EMBL-EBI will operate a pilot *Research Accelerator Hub* to serve the broad life science community based on its successful Embassy cloud, which it has been piloting since 2011. EMBL-EBI is Europe’s principal provider of life science services, and provides many well-used resources, including UniProtKB, Ensembl, PDBe and ENA. The largest datasets (such as the ENA’s DNA sequence archive which requires around 2 Petabytes of storage) are doubling every year, and many other datasets are growing rapidly. The Embassy cloud offers Infrastructure as a Service (IaaS) to other organisations (described as “tenant” organisations) using the proven VMWare technology. The service is operated from EMBL-EBI’s resources within its tier-3 London data centres.

The principal benefit offered by the Embassy cloud is the potential of having virtual machines running on the same high performance network as the substantial public biology data resources held by EMBL-EBI. This allows high-bandwidth, low-latency operations to successfully occur, such as mounting file systems directly to the virtual machines. This eliminates the requirement to replicate such data within the tenant’s own remote infrastructure. Tenant organisations build and administer their own virtualised infrastructure (both machines and networks) with access to EMBL-EBI’s portfolio of datasets. There is no fundamental change of data access rights by being a tenant organization; the benefit is removing both the need to duplicate data and the complex data management of keeping a remote copy in sync with a EBI dataset.

As of June 2013 there are 8 tenant organisations participating in the Embassy Cloud trial, from both the academic and private sectors. The private sector will be paying “at cost” for the machine usage; currently, academic users are selected by their use case. Use cases have included testing direct access to the 1000 Genomes data collection, as well as live public services operated by third parties such as a European mirror for OMIM.

EMBL-EBI plans to continue to evolve the Embassy model from 2014. This will involve a significant scale-up in capacity to include high-performance compute, providing a more efficient combined pool for internal and tenant organisations. Planned use cases include large-scale analysis of genomic data via partnerships with the International Cancer Genome Consortium and the metagenomics dataset of the high profile Tara Oceans

³⁰ <http://press.web.cern.ch/press-releases/2013/06/cern-and-wigner-research-centre-physics-inaugurate-cern-data-centres>

project. Further developments will target other life science use cases in medicine, agriculture and ecosystems, integration into a wider ecosystem of federated cloud infrastructures, and further adoption of open technologies as they become available. Many of these activities will be based upon the outcomes of EIROforum Helix Nebula initiative, as well as related ELIXIR Pilot Actions.

CERN and Wigner as a Research Accelerator Hub

The Wigner centre was contracted by CERN following a competitive call to tender, to which approximately 20 sites around Europe responded. Wigner is operated as an extension to the CERN Data Centre (hosted on the CERN site in Switzerland) and consequently CERN can assure the business continuity for the critical systems in case of major problems. The Wigner data centre is connected to CERN via multiple 100Gbs network links. CERN is integrated with many of the world leaders in international networking, including hosting a GEANT Point of Presence (PoP), as well as being a major centre for commercial networking for both telecoms suppliers and Internet Service Providers (ISPs). Additional network options provided by public and private bodies are welcome, according to the needs expressed by the users. These range from commodity IP services to dedicated circuits directly to Wigner, or via CERN.

The following initial portfolio of services will be provided by this prototype *Research Accelerator Hub*:

- A virtual multi-tenant compute environment to provision and manage networks of virtual machines on-demand.
- A 'dropbox' style service for secure file sharing over the internet.
- A point-to-point reliable, automated file transfer service for bulk data transfers³¹.
- An openAIRE-compliant open-access repository³² for publications and supporting data, allowing users to create and control their own digital libraries. Persistent digital object identifiers will be assigned to all publically available uploads so as to make them citable and permit the creation of a digital data continuum spanning from experimental data through to publications including links to commercial publishers.
- A long-term archiving service³³.
- Integrated digital conferencing³⁴ tools allowing users to manage their conferences, workshops and meetings.
- Training services: online training material³⁵ will be made available and re-enforced with advanced training via brief residential programmes and secondments, such as CERN's openlab summer student programme³⁶, which has

³¹ <http://information-technology.web.cern.ch/services/file-transfer>

³² <http://www.zenodo.org/features>

³³ <http://castor.web.cern.ch/>

³⁴ <http://indico-software.org/>

³⁵ <http://indico.cern.ch/categoryDisplay.py?categId=88>

³⁶ <http://openlab.web.cern.ch/news/cern-openlab-summer-student-programme-invites-2013-applications>

recently been expanded with the engagement of EIROforum organisations and more companies.

The services will be accessible in a single sign-on (SSO) manner supported by a federated identity management system. These services will be implemented using the OpenStack³⁷ open source software suite accessible via the most popular interfaces (including EC2, S3 and potentially OCCI). OpenStack has been selected as the framework for the e-infrastructure services offered by the prototype *Research Accelerator Hub* because it has emerged as a clear leader in the vibrant cloud management market. CERN has extensively tested the software and has adopted it for the management of its production resources at its data centre in Geneva. OpenStack is fully open source and distributed under the Apache 2 license, supported by a growing consortium of more than 180 private companies and public organisations with a well-defined and transparent governance structure representing a global community of more than 9,000 people across 87 countries. More than 800 developers are contributing to the software. All of these attributes means OpenStack does not require public funds to be maintained and offers one of the best environments for innovation.

Recent developments lead CERN to conclude that the open source software technology will shortly be available to make this federation of a network of *Research Accelerator Hubs* technically feasible. Federation is stated as an objective of the OpenStack Foundation³⁸:

“The vision of federation across deployed OpenStack clouds is critical to the OpenStack vision. At the core OpenStack is a framework for managing and provisioning compute, network and storage. From that core many things are possible. But, we hope and will push for the OpenStack core to remain robust, consistent and stable. Yes, every deployment will have its own configuration and technical specifications, but as long as the core concepts and native APIs are robust and consistent, the possibility to drive real interoperability will exist and the value from this is a huge advantage of the open cloud concept”.

Long-term sustainability, governance and collaboration with industry are achieved by building the prototype *Research Accelerator Hubs* on the existing open source community and means national and European funding can be focused on innovation.

The services provided by the prototype *Research Accelerator Hub* will not be made available on a commercial basis. CERN rents space in the Wigner data centre and commercial suppliers can discuss directly with the local authorities if they are interested in locating their own services in Wigner. This co-location model is already used in a number of centres around Europe offering services to the business and public research sectors, including an internet exchange point at CERN’s data centre in Geneva³⁹

³⁷ <http://www.openstack.org/>

³⁸ <http://www.rackspace.com/blog/an-open-letter-to-the-openstack-community/>

³⁹ The CERN Internet eXchange Point, <http://cixp.web.cern.ch/>

and the EMBL-EBI Embassy cloud that enables external organisations to perform secure data analyses on large datasets using virtual machines.

The proposed funding model for the pilot service is that CERN will enter into formal agreements with partner research organisations that wish to jointly develop and use the services offered by the prototype *Research Accelerator Hub*. CERN will take responsibility for operating the services at the *Research Accelerator Hub* and cover the costs during the first year of operation. Beyond this first year, each partner research organisation engages to fund the cost of the services their registered users consume according to a pay-per-usage model, the details of which will be jointly-developed during the first year. Should a partner research organisation withdraw from the pilot service then access to the services of the prototype *Research Accelerator Hub* for their registered users will be disabled following a grace period, after which the host organisation will no longer guarantee the integrity of any material under its custody. The costs of operating the *Research Accelerator Hub* at the capacity foreseen for CERN's own usage will continue to be paid by CERN itself. CERN will charge to the partner research organisations the additional costs incurred when increasing the site's capacity for the purpose of the *Research Accelerator Hub*. These additional costs will include processors, storage, network switches, cabling, energy, software licensing costs and operational staff. CERN will operate the services on the same hardware configurations⁴⁰ and at the same level of quality as those offered to its high energy physics user community. As the data custodian, CERN will not exert any ownership or intellectual property rights over the material that has been deposited. Tools for monitoring the resources consumed and the quality of services provided will be made available. The services will be integrated in a bottom-up manner, starting with basic IaaS services and allowing each partner research organisation to integrate and manage their own higher-level services and portals. Any new proposed service will only be introduced if there is at least one identified partner research organisation that is prepared to contribute to the costs of operating the service. Support for the cloud services will be integrated into CERN existing support structures (e.g. the standard service portal⁴¹). Effort will be required from each partner research organisation to curate their data-sets, connect their identity federations, deploy their community specific services and portals as well as manage the interaction with their registered users and associated support activities.

⁴⁰ Technical Description Servers and Storage for Physics Data Processing, Acquisition and Control, CERN, MS-3903/IT, http://cds.cern.ch/record/1472671/files/MS-3903_Technical_Description.pdf

⁴¹ <https://cern.service-now.com/service-portal/>

Funding models for the prototype Research Accelerator Hubs

To ensure the sustainability of the prototype *Research Accelerator Hubs*, funding models are envisaged where all the stakeholders participate. Sustainability is seen as a process rather than a fixed solution and the funding models will, therefore, evolve over time.

The rationalisation offered by the prototype *Research Accelerator Hubs* and their integration into a hybrid model will provide a number of advantages:

- Provide a clear example of the scale and nature of the services that would be required from commercial suppliers.
- Offer “user aggregation” of procurement, since the prototype *Research Accelerator Hubs* will be in a position to negotiate with suppliers on behalf of all the partner research organisations. This will ensure that the most cost-effective means of providing a service will be available at any moment in time. This will also simplify the contractual and administrative aspects of procuring services from commercial suppliers for the cooperating research organisations.
- The procurement and deployment cycle for the *Research Accelerator Hubs* will take several months and so the integration with commercial cloud service providers will be used to provide additional capacity. This will bring added value to the cooperating research organisations and their users by making it simpler for them to use commercial cloud services in ‘bursts’ to access additional resources to fulfil any short-term requirements.
- The technical aspects of migrating user workloads to commercial services will be simplified and accelerated.
- The prototype *Research Accelerator Hubs* will make it possible to host data-sets on a cost-recovery basis while offering open access so they can be exploited via commercial services. This will remove the need for commercial suppliers to make investments to host data-sets without a clear estimation of how much business they will generate.
- The prototype *Research Accelerator Hubs* will, in relation to Europe’s HPC strategy, address the need for more flexible, easier-to-use, more productive and more cost-effective capacity-style HPC systems.

Timeline

The timeline below shows the major milestones foreseen from conception through to the establishment of the pilot service.

Fourth Quarter 2013:

- Discuss implementation plan with relevant EC projects
- Refine implementation plans for the prototype *Research Accelerator Hubs* with further details of each service offered, SLA, terms and conditions of service
- First meeting of the user forum

First quarter 2014:

- First meeting of management board
- Initial portfolio of services (see Portfolio of services) made available for testing by a small number of research partner organisation(s) via the prototype *Research Accelerator Hubs*

Second Quarter 2014:

- Management board meets to approve opening of the pilot service

Third Quarter 2014:

- Start of pilot service

Governance

As the importance of European e-infrastructures grows and matures it becomes increasingly important that user communities are able to voice requirements and help drive the direction of their evolution. However, the diversity of communities and the relative maturity in their international collaborations makes this difficult to implement in an efficient and cost effective manner.

Pan-European e-infrastructure providers need to understand the requirements of a wide variety of possible communities and have contact with them. In general, this is not the end-scientist or -engineer, but the institute or project that supports them.

What is proposed here is a pan-European forum for organisations and projects that operate at an international level, in order to present to the policy makers and the infrastructure providers where there are common needs and opinions and where there is divergence. This will enable both policy makers and e-infrastructures providers to have a view across many research domains and be able to take strategic decisions that will reflect the commonalities, and differences, that exist. This forum should be independent of any supplier and engage across research domains so that shared information and strong representation can be established.

The problem

E-infrastructure investments on behalf of the publicly funded research community represent an important and growing budget item. As identified in the GEANT Expert Group report⁴², the user communities will increasingly be called upon to pay for the services they receive if e-infrastructures on which users can depend are to continue to survive. E-infrastructure costs will be an integral part of the cost of doing science and, consequently, e-infrastructure investments must make a substantial and sustainable impact in order to be justified. To evaluate the impact, it is essential that the market of end-users is well understood by funding agencies and e-infrastructure service providers. An unpopular way of describing this is that there must be a business case for an e-infrastructure, but indeed it is precisely the considerations of who will pay what for the services that needs to be addressed. This issue was highlighted by the e-Infrastructure Reflection Group (e-IRG) in its 2012 roadmap report where it stated there is a “lack of business models based on secure and sustainable funding streams for the use and innovation of e-Infrastructures”.

This is a difficult topic as more and more IT services are commoditized and available on the open market and the research communities will adopt the most cost effective options for them.

E-Infrastructure services provided to the research community then must be innovative and either address needs that are clear or create opportunities for evolving the

⁴² <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/geg-report.pdf>

“business” of science. In either case they must be relevant enough to attract the investment of the user communities that they target.

This body intends to drive this process by providing information that helps e-infrastructure providers to create solutions that are of value to the user communities and develop the business case.

e-Infrastructure activities

There are currently a number of infrastructure activities that engage with communities in a variety of ways. Recently, the GEANT project (GN3+) has reviewed its governance model and has concluded that it needs better representation from user communities. This awareness has been a growing trend across all providers.

e-IRG	Engages with country representatives concerned with e-infrastructure policy.
EGI	Engages with National Grid Initiatives (NGI's)
GEANT	Engages with the National Research Networks (NREN's)
TERENA	Engages with National Research Networks, large users and commercial providers
PRACE	Engages with the Supercomputer users.

The e-IRG 2012 roadmap highlighted the need for *“a single organisation with a central role for user communities with a particular emphasis on involving large, advanced and well-organised user communities at a European level and beyond.”* and that *“on the strategic level user communities will have to organise themselves to drive the long-term strategy.”*

The purpose of this proposed user community body is to provide timely, relevant and coordinated information to all the above, and any future activities, in order to synthesize the trends, discontinuities and emerging needs of the European scientific and technical community.

Membership

It is the intention for this activity to be lightweight, open and have no rigid constraints. In particular it is important to understand the issues facing emerging and future activities in the science and technology domain, such that policy makers can react accordingly and e-infrastructure providers can have a coordinated response.

To have a universal representation would lead to an unmanageably large group being formed with practical consequences on the activities that can then be performed.

The first approach will be to *solicit individual membership* from a limited but representative set of activities:

- The ESFRI cluster projects (BioMedBridges, CRISP, DASISH, ENVRI)
- The EIROforum members
- The new Flagship Projects for Horizon 2020 (Human Brain Project, Graphene)
- The ERF (European Association of National Research Facilities) that will provide representation of national research infrastructures across Europe
- The League of European Research Universities⁴³
- The Association of European Research Libraries⁴⁴

The intention is that the members should be able to represent the needs of the end-users supported by these organisations and projects. In addition to this standing body, a number of Open User Forums could be organized as widely publicized events where the work can be presented and additional input solicited from a wider community. Such events could be co-located with existing e-infrastructure or research community events to maximize engagement.

It is essential that the members are able to represent a strategic view and that a broad range of disciplines is represented.

Each member will be charged with identifying a number of individual researchers that are active in their discipline and not engaged with major research organisations. These individual researchers will represent the views of what is referred to as the long tail of science, namely important research being performed by small groups or establishments. These individuals will be invited to participate in the Open User Forums.

Members are expected to contribute in a written and oral fashion to the activities.

⁴³ <http://www.leru.org/index.php/public/home>

⁴⁴ <http://www.libereurope.eu/>

Objectives

The intention of this body is to coordinate the discussions broadly across organisations that have growing, or emerging, needs in the e-infrastructure space. The initial areas of common interest are expected to be:

- Networking
- Cloud
- Big data

This is expected to change over time as new paradigms emerge and become important. Such strategic topics concerning e-infrastructures include, but are not limited to:

- The growth of demand and expectations from e-infrastructures in terms of infrastructure and services.
- The value of a proposed service or infrastructure to the scientific community.
- Interoperability and Sustainability
- Identifying inhibitors to use, including regulatory, procurement and legal as well as technical issues.
- Sharing of best practices and successful approaches.
- Collaboration and creation of common services.
- Creation of common understanding between service providers and the user communities.
- Several ESFRI projects have highlighted the value of “user aggregation” of needs with respect to working with industry. CRISP ran a workshop on how the research labs can work together with the IT industry more closely⁴⁵.

e-IRG in its blue paper from 2010⁴⁶ includes a list of specific subjects that came out of the survey of the ESFRI projects performed by the European E-infrastructure Forum⁴⁷:

- Single sign-on: consistent access to resources
- Virtual organisations (collaboration)
- Persistent storage: long-term preservation of data and its access
- Data Management services
- Standards – web services
- Workflows – support of access to HPC/grid/network resources (compute and data) across Europe
- Training
- Global scope: beyond Europe

This body is not a technical body, but will provide input to solution providers.

⁴⁵ <http://www.isgtw.org/feature/next-generation-scientific-computing>

⁴⁶ http://www.e-irg.eu/images/stories/eirg_bluepaper2010_final.pdf

⁴⁷ <http://www.einfrastructure-forum.eu/>

Deliverables

This body will provide input to e-infrastructure providers and projects, the European Commission and national funding agencies as well as scientific communities themselves. In particular it may be expected to:

- Prioritise and publish issues facing the scientific communities in the areas of e-infrastructures.
- Maintain a database of contact information.
- Provide an estimation of the impact of the e-infrastructures on the research communities.
- Provide information on the potential for a service in terms of market size and likely adoption.
- Organise representative input from the scientific communities through workshops and polls.
- Participate in strategic discussions with e-infrastructure providers and projects.
- Participate and provide input on strategic directions from the scientific community for the e-IRG, European Commission and national funding agencies.

Recent examples of user coordination

A number of recent examples, described below, show the value of user coordination with respect to e-Infrastructure planning as proposed in this document.

FIM4R

The subject of secured access to e-infrastructure resources has been developed under the FIM4R group as a concrete example of what can be achieved:

A series of workshops started in summer 2011 to investigate Federated Identity Management for Research (FIM4R) collaborations. These workshops were started as an initiative by the EIROforum IT working group. Through these workshops, several research communities have converged on a common vision for FIM, enumerated a set of requirements and proposed a number of recommendations for ensuring a roadmap for the uptake is achieved⁴⁸.

The authors of the paper come from different research communities that span all four ESFRI cluster projects (BioMedBridges, CRISP, DASISH, ENVRI). The GEANT project has taken the contents of this paper and used it form part of its new project, GN3+, funded by the EC that started on 1st April 2013.

So here is a concrete example of how the research communities working together to clarify their requirements can provide strategic direction to the e-infrastructures. The FIM4R workshops have also spurred a series of pilot projects that jointly involve the research communities and e-infrastructures which are being used to explore the requirements and check the suitability of the services being offered by the e-infrastructures.

Helix Nebula

Similarly, support of access to HPC/grid/network resources (compute and data) has also seen some concrete results. ESA, CERN and EMBL have come together to specify their requirements for the use of commercial cloud computing services and consequently formed the Helix Nebula initiative⁴⁹ as a public-private partnership with a growing number of suppliers. GEANT and EGI have joined the initiative and worked to develop a federated cloud architecture. In this hybrid model GEANT is now providing network connectivity to commercial data centres and EGI foresees the potential of the interoperation of its publicly funded sites with the commercial cloud providers.

Joint development with industry

On 1st February 2013 CERN hosted a workshop in the context of the CRISP project on the IT requirements for the next generation research infrastructures. More than 100 participants from the physics research infrastructure and IT industry participated to prepare a roadmap for future joint developments. A summary of the workshop is available online⁵⁰.

⁴⁸ <https://cdsweb.cern.ch/record/1442597>

⁴⁹ www.helix-nebula.eu

⁵⁰ <http://indico.cern.ch/getFile.py/access?resId=0&materialId=0&confId=212402>

Relationship with other activities

Research Data Alliance

The subjects addressed by the User Forum include data management aspects and hence are likely to be relevant to the recently formed Research Data Alliance (RDA)⁵¹. The foreseen membership of the user forum includes leading research organisations that are the source of large and continuously growing data-sets actively used by the global research community. These organisations represent '*big data factories*' that will be primary contributors of datasets to the e-Infrastructure commons. Hence the work and deliverables of the User Forum will provide strategically important input for RDA working groups. The RDA governance model is still under evolution and it currently appears that the Technical Advisory Board would be a suitable interaction point.

e-Infrastructure Reflection group (e-IRG)

The e-Infrastructure Reflection Group (e-IRG) consists of official government delegates from all the EU countries. The e-IRG produces white papers, roadmaps and recommendations, and analyses the future foundations of the European knowledge society. The e-IRG can be considered as orthogonal to the user forum, which represents e-infrastructure users and research disciplines rather than countries. The results of the work of the User Forum will provide useful input to e-IRG for its publications.

⁵¹ <http://rd-alliance.org/>

Process

In addition to the formal membership, there will be a steering group composed of a chairman and 3 other members that will:

- Be responsible for arranging the organization of the events. The intention is that each user community will be invited to host these events to encourage greater engagement from their community and network with representatives from other communities.
- Represent the results of the activities to the stakeholders and various international bodies.
- Fix the agenda and topics/issue to be discussed as a function of the interests of the members at large.
- Maintain the information repository and outreach utilities.

2-4 Meetings per year will be organized where working groups may be formed to address specific topics.

1-2 Open User Forums could be organized per year.

Working groups' results will be hosted and made available on a website.

A mailing list of all members will be available.

Reports on the meetings of international bodies where the results of the forum are presented will be documented and made available through the website.

Europe's Intergovernmental Research Organisations



EMBL

