31-03-2014

Milestone MS101 (MJ1.2.1): Network Architectures for Cloud Services White Paper

Milestone MS101 (MJ1.2.1)

Contractual Date:	31-03-2014
Actual Date:	31-03-2014
Grant Agreement No.:	605243
Work Package/Activity:	12/JRA1
Task Item:	2
Nature of Deliverable:	O (Other)
Dissemination Level:	PU (Public)
Lead Partner:	CARNet
Document Code:	GN3PLUS14-976-21
Authors:	D. Regvart (CARNET), Y. Demchenko (UvA), S. Filiposka (UKiM), M. de Vos (SURFnet), T. Karaliotas (GRNET), K. Baumann (SWITCH), D. Arbel (Technion)

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 2007–2013) under Grant Agreement No. 605243 (GN3plus).

Abstract

This document proposes an initial GÉANT Open Cloud eXchange (gOCX) architecture, taking into account the background environment in which it will operate. It presents a brief overview of related projects, as well as possible extensions.

Table of Contents

Table of Figures

Executive Summary

The rapid increase in the use of cloud based services and cloud computing technologies has caused an urgent need to research the general architecture and network technologies that are best suited for cloud-based services.

A brief review of cloud services in the Research and Education environment demonstrates the need to support not only typical users (with email, cloud storage and CloudApps services) but also those Data Intensive Science applications (i.e. “Big Data” science) with access to High Performance Computing, VMs and Virtual Computing Labs, in order to provide practically unlimited resources on demand. NRENs are striving to modify and improve their networks so that they can offer scalable cloud-based services with the high Quality of Service (QoS) users have come to expect.

This document proposes a solution, the GÉANT Open Cloud eXchange (gOCX), that considers these trends in cloud services and addresses the related problems. The main purpose of gOCX is to provide a framework and facilities for higher QoS cloud services delivery from the Cloud Service Providers (CSPs) to the NREN’s customers (universities, research institutes and other organisations) and to end-users. It also aims to simplify the integration of cloud-based applications between universities. Concerned only with service delivery, the gOCX remains neutral with respect to actual cloud services provisioning, limiting its services to the Layers from 0 to 2 of the transport network in order to remain transparent to the current cloud services model.

Section 2 gives an overview of the typical cloud services that have been embraced by universities and research institutes as well as detailed analyses of a few selected use cases in order to obtain a clearer picture on the range of cloud service usage by these institutions. The results of this analysis show a need for a dedicated cloud services delivery infrastructure that will support the emerging advanced research and collaboration at universities and other research centres.

Section 3 of this paper lists the requirements, the architecture and components, and the design elements of the GÉANT Open Cloud eXchange.

Finally, section 4 looks at projects that are related to gOCX and that may provide helpful lessons: Prototype gOCX services implementation by SURFnet using Open Lightpath Exchanges (OLE); SURFConext; automated GLIF Open Lightpath Exchange (GOLE), using eduGAIN as a model for the gOCX TTP trust management and services federation; and the GRNET Okeanos project.

1 Introduction

The use of cloud-based services and Cloud Computing technologies [NISTdef], [NIST-CCRA], [NISTrec], [CIRefFW] is set to increase rapidly among universities and NRENs in the near future. This causes an imperative need to research the general architecture and network technologies that are best suited for cloud-based services. The results of these research activities will help NRENs modify and improve their networks so that they can offer scalable cloud-based services with high Quality of Service (QoS) to their users. The network infrastructure plays an important role in delivering cloud services and can be viewed as being comprised of three interdependent structures:

- The user access network, which connects users to applications.
- Extreme high-speed networks, which interconnect physical servers and offer fast migration of their virtual machines (VMs).
- Megapipe networks, interconnecting storage tiers.

Thus, one of the key challenges of supporting the high demands of cloud computing is to investigate and propose a suitable network architecture and its interdependent structures that will satisfactorily respond to user demands.

Among other factors, the need for enabling dedicated high QoS for cloud services is mostly driven by the increasing demand from emerging Data Intensive Science applications (i.e. “Big Data” science) that require both advanced computing and networking infrastructure and infrastructure to support collaborative groups. Big Data science has created the need for an additional consolidation of technologies that will help address both infrastructure performance and manageability for data/computing-centric/driven tasks and for advanced user support for project-oriented collaborations. At present, NRENs are providing network access and advanced infrastructure services for their constituencies, as well as the infrastructure for federated access control and cross-organisational collaborative groups support. All of the external cloud services provided by outside Cloud Service Providers (CSPs) are being delivered over the common Internet connections together with the rest of the commodity traffic. Taking into account future trends, the work presented in this document focuses on the network elements and orchestration that will help NRENs (and GÉANT as a whole) to support high-demand cloud-based services with guaranteed QoS.

The higher QoS required by cloud services can be delivered in different ways. In many cases, large CSPs can create/establish a Point of Presence (POP) for large customers. A recent example of this approach is the established Microsoft POP at the UK education network [MSinUKedu]. On the other hand, customers with distributed campuses are prepared to consider extending their network to a CSP’s POP. The latter approach is

becoming popular among NRENs at national level. This approach can be also used for Europe-wide projects and communities. At the same time, CSPs are starting to offer dedicated links to their cloud facilities, as is the case with Amazon's Direct Connect service [AmazonDC].

This document proposes a solution, the GÉANT Open Cloud eXchange (gOCX), that considers these trends in cloud services and addresses the related problems. Its main purpose is to outline an initial gOCX architecture and design within the background environment in which it will operate, including standardisation activities related to the inter-cloud technologies.

The main purpose of gOCX is to provide a framework and facilities for higher QoS cloud services delivery from the CSPs to NREN customers (universities, research institutes and other organisations), as well as to end-users. It also aims to simplify the integration of cloud-based applications between universities. Concerned only with service delivery, the gOCX is not involved in actual cloud services provisioning and limits its services (taking into consideration the transport network) to the Layers from 0 to 2, in order to remain transparent to the current cloud services model.

The motivations behind the development of gOCX are discussed by giving examples of general use cases, based on which the gOCX requirements are defined. A brief overview of other related projects is also presented in Section 4. Finally, a summary of future activities and possible extensions is provided.

2 Cloud Services in the R&E Environment

The growth in cloud services is mostly due to the extensive number of use cases and usage scenarios that offer great cost benefits based on the pay-as-you-go model along with its great scope for scalability. General use cases and usage scenarios have been defined by the industry and documented by NIST [NISTusecases], Open Data Center Alliance [ODCA], and the Global InterCloud Technology Forum [GICTF].

While universities and research institutes have embraced the typical cloud services offered (e.g. email), there are also a number of specific cloud service usage scenarios for research and education (R&E) institutions that go beyond general use cases. Thus, in order to obtain a clearer picture of the full range of cloud service usages by these institutions, it was important to gain an overview of the typical cloud use by universities and the research community, as well as to carry out a deeper detailed analysis of a few selected use cases. The results of this analysis show what motivates the need for a dedicated cloud services delivery infrastructure that will support emerging advanced research and collaboration at universities and other research centres.

2.1 Use Cases and Scenarios

The cloud services provided by GÉANT and NRENs are targeted at universities and research institutes/centres as service users, with academic staff and students as end-users. Thus, the potential number of users may be measured in tens of millions. Cloud services can make academic network usability, administrative processes, research and scholarly collaboration more efficient and effective on a local, national, regional and global scale. Because of the vast number of end users, it is essential that all supported cloud-based (as well as non-cloud-based) systems can be accessed via a Central Authentication Service (CAS) using single sign-on (SSO). SSO can also be combined with Federated authentication, allowing members of one organisation to use their authentication credentials to access a cloud service of another institution. Hence, trust relationships that allow different entities to accept each other's assertions are essential.

2.1.1 Typical Uses of Cloud Services

Cloud computing services are seen as a solution for the near insatiable demand from researchers for bandwidth and computing power and from students for sound- and video-intensive applications. An important option is the development of collaborative service offerings among universities. The following are the typical/popular uses of cloud services by universities on a departmental/institutional level as well as at the individual level for staff and students.

2.1.1.1 *Email*

In any organisation, email is usually among the first services to be moved to a public cloud environment from the on-site data centre. The cloud solution provides one consolidated place for all distributed entities inside the organisation, while applying less limitations (e.g. inbox size) using scalable and well-distributed resources. This is especially the case for universities that do not have the resources to provide for their large and rapidly changing user base. Migrating email to the cloud also has the benefit of global accessibility, which is vitally important for the mobile research community.

2.1.1.2 *Storage services*

Storage services are very popular both for backup purposes and for sharing documents and data. Documents and data sharing are important services to support inter- and intra-organisational collaboration. Cloud storage options allow organisations to choose one or more cloud locations to store files. Once saved, the files are easily accessible via the web, reducing the need for multiple copies or additional disk space on the computer or mobile device. Cloud storage providers include Box, Dropbox, GoogleDrive and SkyDrive. All of the supported locations can be accessed using a single point of entry Central Authentication Service (CAS). Despite the existing security concerns these services are quite popular for regular, non-security-critical cases. The TERENA community is developing secure cloud storage sponsored by TF-Storage [TFstorage].

2.1.1.3 *CloudApps services*

The advantages of cloud applications (i.e. Software as a service – SaaS) to researchers and students are making them quite popular. SaaS users can access the application “anytime, anywhere”, share data and collaborate more easily, as well as keep the data stored safely in the infrastructure.

Universities can use CloudApps to effectively implement collaborative learning approaches where the students are able to work alongside students from other locations in order to achieve a common goal. CloudApps and collaboration technologies can improve educational services, giving students access to low-cost content, online instructors, and communities of fellow learners, thereby greatly enhancing e-learning capability and making distant learning more effective and more efficient.

From the researchers' point of view, SaaS applications can provide seamless collaboration on joint international projects and a way of making e-Infrastructure resources available. CloudApps allow researchers to easily define a computational task and provide fast access to results.

2.1.2 Scientific data access – Big Data

Many general research and specialised scientific software and applications are provided in the form of SaaS. Although SaaS are usually mostly interactive applications, there are also opportunities for batch-processing for jobs that can analyse terabytes of data and take hours to complete. The cloud computing model is a perfect match for Big Data-related research since it provides unlimited resources on demand. Currently, increasing amounts of scientific data are available to research communities and collaborative groups in the form of Grid and cloud storage resources (e.g. LHC experiment data and genome data) provided worldwide.

Programming abstractions such as Google's MapReduce [MapReduce] and its open-source counterpart Hadoop [Hadoop] allow programmers to deal with Big Data tasks while hiding the operational complexity of the parallel execution across hundreds of servers. Integrating data from a widespread network of sensors is another example of applications spanning across many different fields. Moreover, the latest versions of different software packages are capable of using cloud computing to perform intensive evaluations. Another interesting model is based on keeping the data in the cloud and relying on having sufficient bandwidth to support visualisation and a responsive GUI back to the end-user.

2.1.3 Computational Power (e.g. HPC access, VMs, Virtual Computing Labs)

Members of the academic community rent/use on-demand computational and storage services offered by:

- Public vendors, or
- Other organisations inside NRENs' national networks, or
- Academic organisations reachable via GÉANT.

Cloud services can be provided using a high-performance supercomputer cluster, or virtual machines, depending on the needs of the researchers. National and organisational data centres are increasingly providing access to High Performance Computing (HPC) as IaaS cloud services.

Using cloud Infrastructure as a Service (IaaS) services is popular for deploying multiple VMs that can be used for running user-designed services and experimenting with new infrastructure services and protocols. Different types of IaaS cloud offerings can be set up (e.g. Nimbus, OpenStack, or Eucalyptus clouds). The virtual appliances provided can also be used for training and education or for the creation of virtual computing labs. Moreover, science experiment setups can contain references to cloud applications, making such experiments easier to replicate.

2.2 Reference Use Cases

In most cases, cloud services are fulfilled over the Internet and do not require any specific network services. However, advanced research activities (i.e. Big Data science) require access to very large datasets, scientific instruments and HPC. Combining all these components into one scientific infrastructure requires a dedicated network infrastructure and related services to support researcher collaborations.

In terms of generic Cloud Services Models [NISTarch], [CIRefFW], [ICAII], [Demchenko], such a functionality can be defined as an InterCloud Access and Delivery Infrastructure (ICADI) with the main purpose of delivering cloud-based services to organisational customers and end users.

The following reference use cases that require a dedicated infrastructure for delivering cloud services to campus users can be currently identified:

- Streaming high-speed, high-volume experimental or visualisation data to (and from) labs in campus locations that may require dedicated links;
- Scientific data processing with Massively Parallel Processing (MPP) tools available in facilities that are distributed among universities and research organisations;
- CSP and campus network peering over dedicated L0–L2 fibre links.

These reference use cases are collected for the purpose of identifying the required functionality for the ICADI that structurally includes all infrastructure components between the CSP, the final consumer and other entities involved in cloud services delivery and operation.

2.3 Benefits of Dedicated CSP Connectivity

An overview and analysis of the different R&E use cases highlights where a dedicated connection to the CSP is needed to achieve a guaranteed QoS to improve the performances of scientific applications. As previously discussed, these are first of all the Big Data cloud-based applications as well as other use cases that demand an unusual bandwidth and involve moving a vast amount of data between the CSP and the end users. Since all of the cloud services offered by providers outside the NRENs are provided via the Internet, the current network architecture faces difficulties trying to accommodate these specific use case scenarios. One possible solution to this problem is to establish dedicated links between the GÉANT communities and CSPs and to utilise these links for special-purpose cloud services, in order to guarantee the required QoS and satisfy the end-user demand. As mentioned earlier, this actually represents one of the main goals of gOCX, and its introduction would result in a number of benefits in this respect for the academic community.

Using gOCX, a consortium-wide sourcing model can be implemented. This model is based on a not-for-profit aggregation of demand, so that organisations can then either provide cloud services for themselves, or contract to have them provided by a commercial or institutional cloud provider. Aggregating demand would help keep providers accountable and focused on addressing client needs and would allow the universities to preserve

negotiating leverage. Cloud providers on their part would benefit from not having to negotiate agreements with each organisation separately, but only with the gOCX as a single customer.

Since gOCX enables the sharing of private dedicated links to CSPs, this means that other NRENs can also benefit from an already established private link at a given NREN by enabling other universities to reuse the separate fast, stable and secure GÉANT network.

The GÉANT shared cloud services environment also guarantees a certain degree of standardisation, and this best practices approach can ensure that providers meet “reasonable standards” expectations.

3 Open Cloud Exchange

The GÉANT's Open Cloud eXchange (gOCX) is a design solution that addresses the current difficulties in delivering cloud services to organisational/enterprise customers and end users who demand guaranteed QoS while generating large quantities of network traffic. gOCX intends to bridge the gap between two major components of the cloud services provisioning infrastructure:

- The CSP infrastructure, which has a global footprint or is intended to serve the global customer community, including customers who have global/international presence or deliver services to the global/worldwide community; and
- The cloud services delivery infrastructure, which, despite the fact that the cloud services concept is based on ubiquitous Internet connectivity, in many case requires dedicated local network infrastructure.

As discussed while analysing the different use case scenarios, there is an emerging need for joining/combining the CSP infrastructure and the local access network infrastructure. This is especially the case when facing the "last mile" problem in delivering guaranteed QoS cloud services to customer locations and end users.

3.1 Definition

The gOCX concept is based on and extends the Internet eXchange (IX) and Optical eXchange model with additional facilities/functionalities to enable the establishment of *ad hoc* dynamic InterCloud federation and unrestricted cloud provider and customer peering. Peering of local infrastructure providers can also be provided in cases where the delivery of cloud services requires their involvement.

In addition to providing the physical location for interconnecting networks of the gOCX members, the gOCX adopts two basic principles in order to simplify and facilitate service delivery:

- No third-party services (e.g. service brokering, integration or operation) – this means that gOCX will not be part of or involved in business relations related to the actual cloud services delivery;
- Trusted Third Party (TTP) services to facilitate the establishment of *ad hoc*/dynamic federations – gOCX may provide services such as a trusted repository of PKI certificates and the services directory operating under community (and representative's) supervision. gOCX will act as a policy authority for security and operational practices and may provide a clearing house service for SLA and PKI Certificate policies.

The gOCX's role as a TTP will facilitate the creation of dynamic federations and the establishment of dynamic trust relations. Using the membership service in gOCX, member CSPs will be able to establish a trust relation via cross-certification or simply by providing trusted certificates in a repository similar to TACAR (TERENA Academic CA Repository) [TFstorage].

The envisioned gOCX functionality is part of the general ICADI as defined by the Intercloud Architecture Framework (ICAF) [ICAII], [Demchenko]. However, gOCX may limit its services to Layers from 0 to 2 in order to remain transparent to the current cloud services model.

3.2 Requirements

The general requirements for the gOCX functionality and its design are presented below. The rationale for the requirements is provided where necessary.

- Generally speaking, gOCX should follow and leverage the Internet eXchange (IX) design and operational principle, adopted in such a way that it supports the specifics of cloud services provisioning. From this perspective, similarly to IX, gOCX can be defined as a place for inter-connection and peering between cloud service providers and customers.

Big cloud providers are increasingly becoming global service and infrastructure providers with their own international infrastructure, which does not have to be TCP/IP based, and can therefore handle a significant amount of their own and customers' traffic internally.

gOCX may also benefit from being co-located with a colocation service provider, NREN exchange points, or a regional data centre servicing the regional or national research community.

- Primarily, gOCX should provide Layer 0 to Layer 2 network services to interconnect CSP Points of Presence (PoP). This will make it fully transparent to current cloud services models, which generally use Layer 3 network infrastructure virtualisation for deploying VMs and interconnections.

However, further performance optimisation for the cloud infrastructure may require Layer 2 network virtualisation. Consequently, this may require gOCX services at the lower layers.

- gOCX should support secure topology information exchange between its peering members, as an important component/requirement for effective services interconnection at different networking layers.
- gOCX's interconnection network infrastructure must guarantee high QoS parameters (e.g. bandwidth, latency, and jitter) in accordance with the cloud services provided by the connected CSPs and the agreed SLAs.
- gOCX should provide smooth service delivery and integration between CSPs and customers. Besides network connectivity, this implies including support for federated services integration and operation.

These additional services can be generally defined as Trusted Third Party Services (TTP) and may include but are not limited to:

- Trusted introducer service that can be supported by the Trusted Certificates Repository (similarly to the TACAR service offered by TERENA [TACAR].
- CSP and Cloud Services Directory and Discovery Service.
- SLA repository and clearinghouse.

The gOCX architecture should allow for a flexible operational scenario. In order to increase flexibility it may have hierarchical architecture and can be operated by both, NRENs and GÉANT.

When implemented with modern optical network technologies (e.g. Lambda and DWDM), gOCX can be easily implemented using different distributed topological models: extended, collapsed or hierarchical.

3.3 Architecture and Components

Figure 3.1 illustrates the general case of implementing an enterprise or scientific workflow on a heterogeneous multi-provider infrastructure. A gOCX that can be placed between the customers/campuses and cloud providers will provide facilities for interconnecting all members and entities of the federated cloud infrastructure.

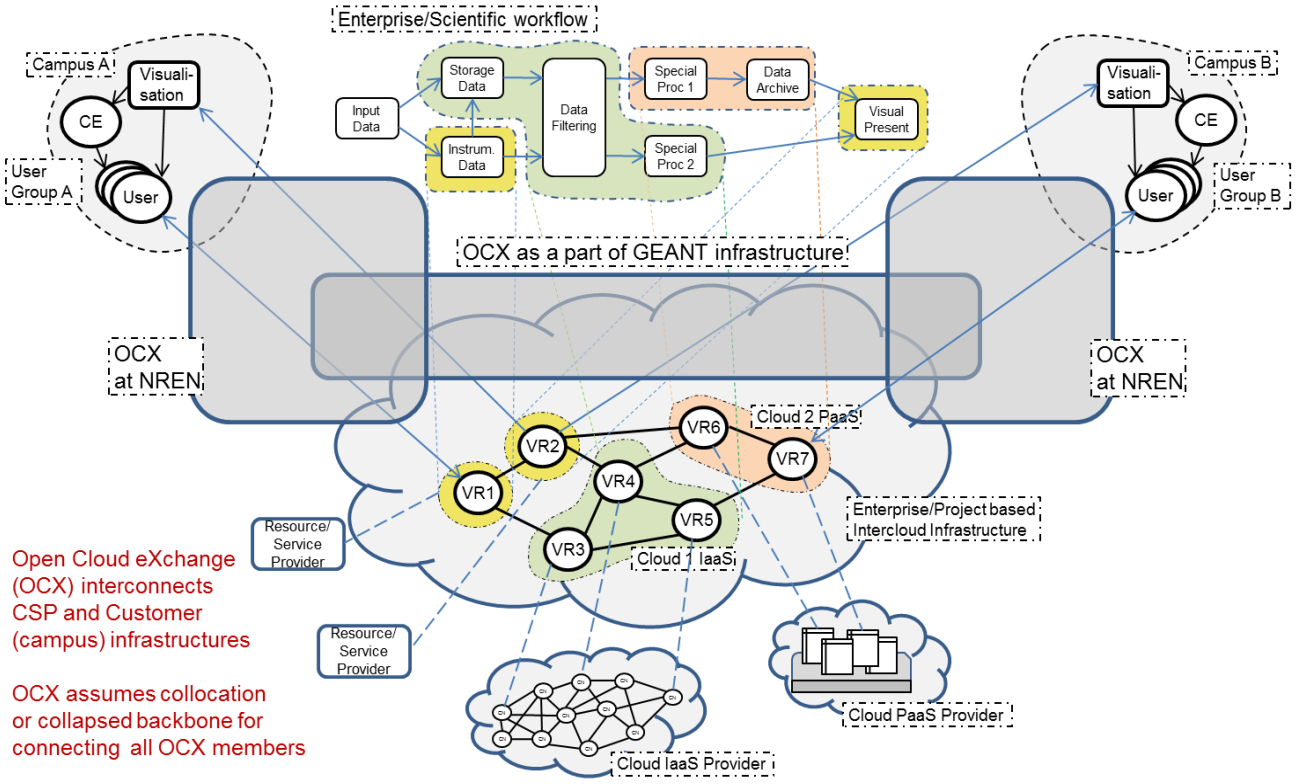

Figure 3.1: Enterprise or scientific workflow implemented on a heterogeneous multi-provider infrastructure

Figure 3.2 shows another view where gOCX services can be provided by the Cloud Carrier or on the Network Provider level, in this case by the NREN or GÉANT.

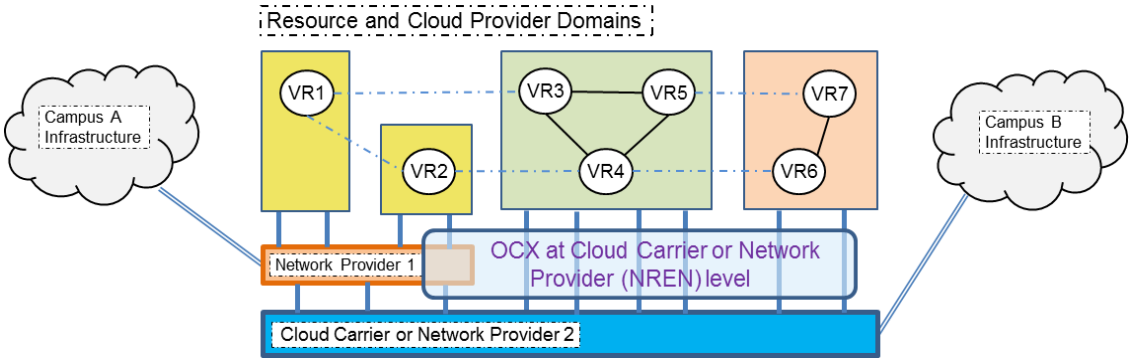

Figure 3.2: gOCX on the Cloud Carrier or Network Provider level

Architecturally and functionally the gOCX includes the following services and functional components (see Figure 3.3):

- Physical Point of Presence (PoP) for providers and customers.
- L0-L2 network interconnection facility (optionally also connectivity using dedicated optical links). The associated service should allow topology information exchange between providers and customers in a secure and consistent way.
- Trusted Third Party (TTP) services in order to support dynamic peering, business/service and trust relation establishment between gOCX members. The specific services may include:
    - Trusted Certificates repository and associated Trusted Introducer service in order to allow establishment of dynamic trust associations and/or federations.
    - Trust Broker service supported by either or both the Trusted Introducer and privacy/data security policy Registry or clearinghouse.
- Publish/subscribe Services Directory and Discovery, in order to provide a list of all CSPs and their cloud services offered so that customers can subscribe to chosen services. An SLA clearing-house service can also be provided.
- Cloud Service Broker (this is an optional additional component) that will provide service advice and integration for the contracted community.

Figure 3.3 provides an overview of the core gOCX components as a part of the federated Intercloud infrastructure that involves multiple CSPs that use gOCX for transparent exchange of traffic/communications between cloud services.

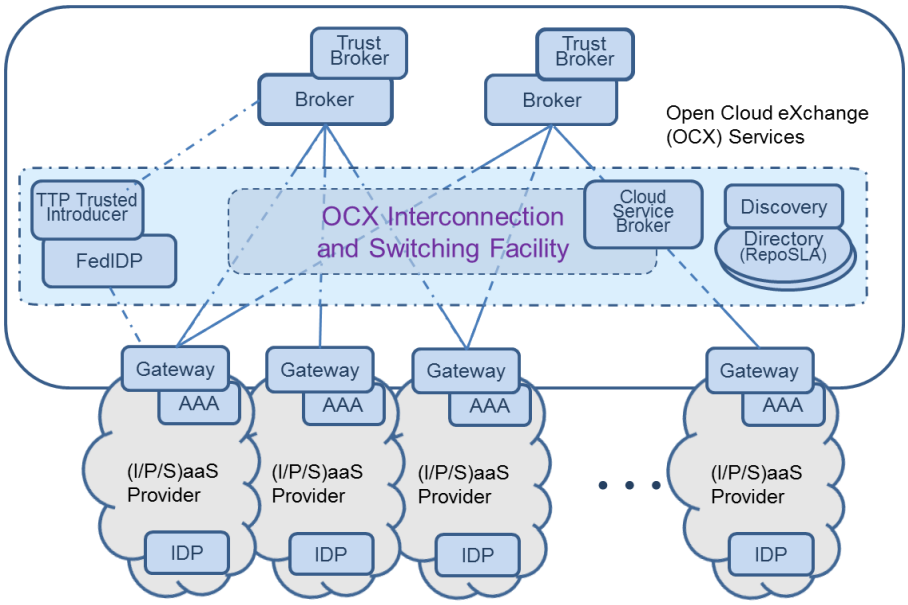

Figure 3.3: gOCX functional components (core and optional)

3.4 Design

The following design proposals are derived in order to implement the functionalities described in the Requirements and Architecture and Components sections.

Topologically, gOCX should allow any-to-any interconnection at Layer 0, Layer 1 and Layer 2. This can be implemented by using corresponding L0–L2 optical switches. Figure 3.4 illustrates the switching topology of gOCX. Please note that the final design implementation must provide for a secure topology information exchange, as well as guaranteed QoS parameters.

Figure 3.4 also illustrates how gOCX can operate as a Trusted Third Party for establishing direct/dynamic trust relations between the gOCX members. These trust relationships can be used for establishing identity management federations among the gOCX members. The research work presented in [Chadwick] and [Ngo] can provide a solution and mechanisms for the establishment of the trust relation in the federated cloud environment. It is recommended that members have trust policies to define the trust criteria. As for defining and implementing a Trusted Introduction Protocol for the establishment of Dynamic Secure Federations using gOCX TTP services, the design should be based on the solutions developed by the IETF Application Bridging for Federated Access Beyond web working group [IETF-ABFAB], [TrustRouter], as well as on the research work carried out by the University of Kent [Chadwick] and the University of Amsterdam [Ngo].

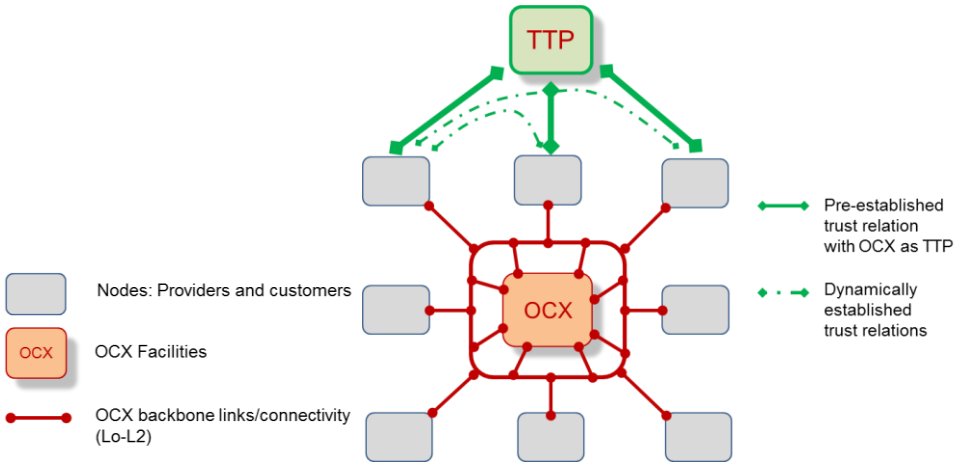

Figure 3.4: gOCX topological model and TTP role for establishing dynamic trust relations

The characteristics of the gOCX require fast decision-making and policy-enforcement mechanisms to enable the transparent coordination of cloud service transactions traffic. gOCX can benefit from adopting the SDN's main design principle, namely the separation of the control and data planes. In this way, the data plane can be optimised by simply applying forwarding rules efficiently at any layer (L0–L2), while the SDN controller will implement the control using data forwarding, data filtering, policy enforcement, TTP services, etc. A modular implementation of the SDN controller such as the one offered by Floodlight [Floodlight] or an implementation using the Network as a Service (NaaS) concepts developed by frameworks such as OpenNaaS [OpenNaaS], will provide the flexibility and extensibility that allow easy adaptation of interconnectivity requirements.

In addition to the SDN-based design, the gOCX design needs to include and define the additional major API's needed to establish and access the gOCX services. These will be typically activated in the process of subscribing to or changing the CSP and during customer setup and establishment of trust relations. In this context, it should be stressed again that gOCX is intended to be transparent to higher-level cloud services protocols and communications.

The following is a non-exhaustive list of groups of APIs that need to be defined:

- gOCX interconnection API that should include the following functions:
  - Interconnection topology, presumably "many-to-many".
  - Links bandwidth and other QoS parameters.
  - VPN/virtual circuits setup, etc.
- Access (publish-subscribe) to gOCX for CSPs, Customers together with a Service Directory listing.
  - Presumably Web Services SOAP or REST interfaces.
- Access SLA and security policy repository and clearinghouse.
  - Presumably Web Services SOAP or REST interfaces.
- Access gOCX Trusted Anchors Repository that will collect and store certificates of gOCX members.

- Presumably Web Services SOAP or REST interfaces.

- (Optional) Secure Token Service (STS) and Federated Identity Provider (FIDP) that can be provided by the GÉANT network or the local NREN.

3.5 Implementation Models

There are different possible locations for gOCX placement: at NRENs, at GÉANT premises and combined versions based on a hierarchical gOCX infrastructure and extended gOCX backplane/backbone.

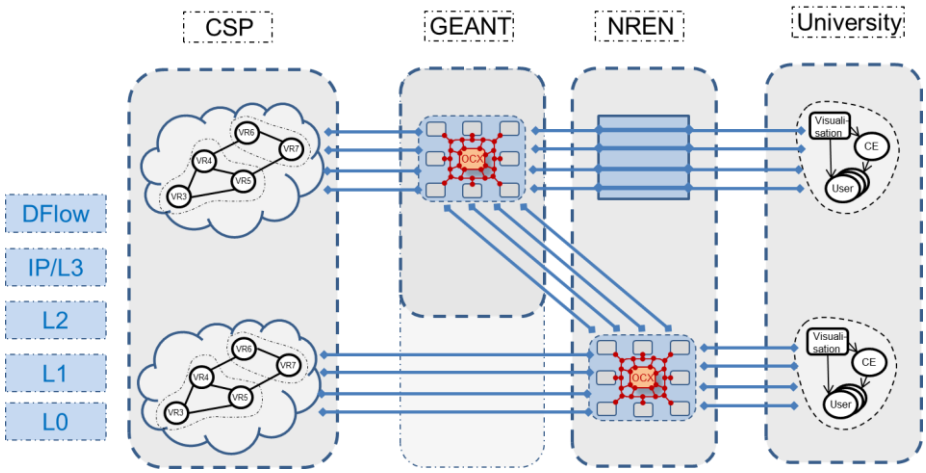

Figure 3.5: Single gOCX located at GÉANT or NREN premises: hierarchical

In order to avoid any problems with a single point of failure inherent in the solution that includes a single gOCX on L0–L3 (IP) and DFlow located on GÉANT's or the NREN's premises, a hierarchical gOCX architecture is favoured. Figure 3.5 illustrates the hierarchical approach with a gOCX at the Trans-European/GÉANT level interconnected with the national gOCXs run by NRENs. Together, they create a cross-border cooperative access infrastructure to cloud services. Each gOCX operates independently, while dedicated (virtual) links are used to interconnect them. Another possibility is a distributed gOCX topology that uses an extended backplane approach where the gOCX switching backplane is extended to a remote location.

In order to establish proof-of-concept, a gOCX demo has been created [TNC] according to the hierarchical implementation model involving a gOCX on the GÉANT level with three gOCXs at different NRENs. The demonstration enables cross-domain VM creation and migration as well as cloud storage usage. The user activities during the demo were also used as a basis for network monitoring and establishing a baseline for the QoS guarantees.

4 Related Projects

4.1 Prototype gOCX Services Implementation by SURFnet

Open Lightpath Exchanges (OLE) allow for any party connections and do not impose any limitations on possible cross connects. These cross connects are always transparently made at L2 or below. SURFnet and NetherLight have been performing pilots with cloud providers who want to offer their services with lightpath characteristics to SURFnet-connected institutions. NetherLight's OLE concept is similar, but with a possible marketplace function. Via this marketplace, cloud providers can offer their services to SURFnet-connected institutions, as well as to other R&E institutions and to other cloud providers via connected OLEs.

Most of the research community would like to use dedicated paths to the cloud provider in order to ensure secured access with no sensitive traffic traversing the Internet and make it possible to incorporate cloud machines into their own network. Connecting the services of GreenQloud (a company from Iceland that provides storage and computation in the cloud) with lightpaths via NetherLight gives SURFnet-connected institutions some advantages over regular IP services, such as guaranteed bandwidth and latency, traffic separated from the regular Internet, possibility of domain extension (VMs appear as if inside the campus network) and cost reduction from offloading traffic. In addition, GreenQloud is also connected to the SURFconext platform, thus providing all their researchers, students, teachers and ICT staff with safe and easy access.

Unified communications, providing a combination of voice, chat, audio- and videoconferencing, can be used via the regular Internet. However, this does not provide enough reliability and QoS and therefore a dedicated path between the institution and the cloud provider is required. In a successful pilot, OneXS provided unified communications for Windesheim University of Applied Sciences. A few months later, another SURFnet-connected institution requested to use the services of OneXS via NetherLight. Using the services of OneXS via lightpaths through NetherLight gave the connected institutions additional advantages like protection from DDOS attacks. The lightpaths between OneXS and our connected institutes are dedicated Ethernet services through NetherLight.

4.2 SURFConext

The SURFconext [SURFconext] infrastructure for online collaboration gives users access to the services of different providers, which can be accessed within a single environment. It connects a number of basic building blocks for online collaboration, such as:

- Federated authentication and authorisation – so that users can securely access all available services via the same account that they use at their own institution.
- Group management – making it possible for access to content and functionalities to be managed centrally.
- Standard data interface – for exchanging activities, reports, and group information (OpenSocial) with cloud applications.
- Various cloud applications (e.g. GoogleApps, EDUgroepen, Sharespace, Liferay Social Office).

SURFconext allows institutions to integrate internal and external online services, thus enabling them to offer users a collaboration environment within which they can access the online services that they require.

4.3 The GLIF "Automated GOLE Pilot" Project

The Global Lambda Integrated Facility (GLIF) is a cooperation between NRENs and research institutes with a twofold goal: to create a forum for engineers and researchers to exchange experiences, and to cooperate to make lambdas available for researchers and projects involved in data-intensive research.

In order to make global optical networking possible, GLIF exchange points have been created where long connections can be linked together to create multi-domain lightpaths. These exchange points typically connect users on layers below L2. In order to clarify the policy applied at exchange points, the concept of a GLIF Open Lightpath Exchange (GOLE) has been defined. This mandates that GOLEs treat everyone the same, and be open about any policy that may be applied [GOLE].

Since the start of GLIF, inter-domain lightpath provisioning has involved much manual processing and actions, since each actor has to play an active role in the process of creating the lightpath, ranging from approving the request to changing the configuration of the network to accommodate the request. Often these are distributed over several time zones, which means that a lightpath request can take about two weeks to be implemented. Based on the Network Services Interface [NSI], the Automated GOLE testbed created by thirteen GOLEs with different implementations of the NSI Connection Service has made it possible for lightpaths to be created and destroyed within seconds. The early experiments with the Automated GOLE testbed have also helped to improve the NSI Connection Service version 2, which is now submitted as a draft standard.

4.4 Federated Access to Web resources: eduGAIN

The GÉANT community has long-term experience in federated network and services access. Federations are typically created at the national level and run by NRENs. To allow federated access at the trans-European level, the community has challenged the level of services federation by enabling inter-federation access, that is, the possibility for users from one federation to access services provided by another federation. This process is made possible through an infrastructure that supports the exchange of information between different countries, by technologies that enable the process to take place in a secure fashion and by legal obligations (such as data protection laws or contractual agreements) that ensure that the user's data are securely handled.

eduGAIN [eduGAIN] is an infrastructure that enables the trustworthy exchange of information about authentication and authorisation between GÉANT partners and beyond. eduGAIN can provide a good model for the gOCX TTP trust management and services federation.

4.5 GRNET Okeanos Project

The GRNET Okeanos project [Okeanos] aims to build a Cloud Infrastructure in order to provide an IaaS service similar to Amazon AWS (Cyclades), a Storage Service (Pithos+), Virtual Networks, and Virtual Firewalls.

Okeanos supports KVM-based VMs. VM storage volumes are physically stored as objects in a distributed, redundant, object-based storage backend. The storage backend is deployed in commodity physical nodes, in a distributed, shared-nothing architecture (i.e. none of the nodes share memory or disk storage) and with no single point of failures (SPOFs). Storage bandwidth and capacity scale with the number of storage nodes, which are added and removed in a live system, with dynamic object replication and automatic rebalancing. This allows for seamless VM migrations between physical nodes.

Okeanos provides virtual Ethernets as a separate resource, giving the user freedom to create arbitrary network topologies of interconnected VMs, e.g., for multi-tiered deployments of enterprise software. Private networks are supported by the API and are exposed all the way to the web UI. Each private network is an isolated Ethernet segment, carrying raw L2 Ethernet frames. This gives an unrestricted choice of IP addressing schemes, allows user set DHCP services to be run, and also supports non-IP traffic. VMs see a separate virtual Ethernet NIC for each private LAN they are part of. The user may protect each public IPv4/IPv6 interface with a virtual firewall, choosing from a number of predefined firewall configurations. Firewalling is provided as a virtual appliance by the infrastructure and works independently from the guest OS running inside a VM.

5 Conclusions

In this white paper, the concept of the GÉANT's Open Cloud eXchange (gOCX) has been proposed. gOCX aims to bridge the gap between the two major components of the cloud services provisioning infrastructure: the Cloud Service Provider (CSP) and the cloud services delivery infrastructure (which in many cases requires dedicated local infrastructure and quality of services that cannot be delivered via the public Internet infrastructure).

The gOCX will provide member peering while at the same time solving the "last mile" problem in delivering cloud services to customer locations. It remains neutral with respect to actual cloud services provisioning and limits its services to L0–L2 in order to remain transparent to the current cloud services model. The initial set of requirements for gOCX has been provided together with a discussion about possible design and implementation models. The proposed gOCX concept will leverage the existing Internet eXchange by the Internet community and the GLIF Open Lightpath Exchange by the GLIF community. Moreover, it proposes additional functionalities that will simplify inter-CSP and customer infrastructure integration while supporting basic cloud-service provisioning models. The most important of these additional features are the Trusted Third Party services that allow for federated infrastructure and access control.

Future efforts will be directed towards a comprehensive definition of the gOCX components, detailing use cases and requirements, after which a pilot implementation can be carried out at a selected number of NRENs and on the GÉANT network. Special attention will be given to enabling secure information exchange between gOCX members.

The project recognises the importance of receiving feedback from the wider community and will propose the gOCX architecture and operational model to standardisation bodies such as the IETF and the Open Grid Forum (OGF) Research Group on Infrastructure Services On-Demand provisioning [ISOD-RG].

References

[AmazonDC]	Amazon Direct Connect service http://aws.amazon.com/directconnect
[Chadwick]	Chadwick, D., M. Hibbert, “Towards Automated Trust Establishment in Federated Identity Management”, Proc. of the 7th IFIP WG 11 International Conference on Trust Management (2013), Malaga, Spain.
[ClRefFW]	Cloud Reference Framework Internet-Draft, version 0.6, January 4, 2014 http://www.ietf.org/id/draft-khasnabish-cloud-reference-framework-06.txt
[Demchenko]	Demchenko, Y., C. Ngo, M. Makkes, R. Strijkers, C. de Laat, “Intercloud Architecture Framework for Heterogeneous Multi-Provider Cloud based Infrastructure Services Provisioning”, IJNGC Journal, July 2013
[eduGAIN]	eduGAIN project http://www.GÉANT.net/service/edugain/pages/home.aspx
[Floodlight]	Floodlight OpenFlow SDN Controller http://www.projectfloodlight.org/floodlight/
[GICTF]	Cloud Computing, GICTF White Paper. August 9, 2010 http://www.gictf.jp/doc/GICTF_Whitepaper_20100809.pdf
[GOLE]	The GLIF “Automated GOLE Pilot” Project. http://staff.science.uva.nl/~delaat/sc/sc10/GLIFAutomatedGOLEPilot.SC.pdf
[Hadoop]	White, T., “Hadoop: The Definitive Guide”, O’Reilly, 2012
[ICAII]	Intercloud Architecture for Interoperability and Integration, Release 2, Draft Version 0.7. SNE Technical Report 2012-03-02, 1 July 2013 http://www.uazone.org/demch/worksinprogress/sne2012-techreport-12-05-intercloud-architecture-draft07.pdf
[IETF-ABFAB]	IETF Application Bridging for Federated Access Beyond web (Active WG) http://tools.ietf.org/wg/abfab/
[ISOD-RG]	Open Grid Forum Research Group on Infrastructure Services On-Demand provisioning http://www.gridforum.org/gf/group_info/view.php?group=ISOD-RG
[MapReduce]	Dean J., S. Ghemawat, “MapReduce: Simplified Data Processing on Large Clusters”, Google http://research.google.com/archive/mapreduce-osdi04.pdf
[MSinUKedu]	Windows Azure to power UK academia’s Janet network http://www.cloudpro.co.uk/cloud-essentials/5614/windows-azure-power-janet-education-cloud
[Ngo]	Ngo, C., Y.Demchenko, C. de Laat, “Toward a Dynamic Trust Establishment Approach for Multi-provider Intercloud Environment”, The 4th IEEE Conf. on Cloud Computing Technologies and Science (CloudCom2012), 3 - 6 December 2012, Taipei, Taiwan

[NIST-CCRA]	NIST SP 500-292, “Cloud Computing Reference Architecture, v1.0.” http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909505
[NISTdef]	NIST SP 800-145, “A NIST definition of cloud computing” http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
[NISTrec]	NIST SP 800-146, “Cloud Computing Synopsis and Recommendations”, May 2012 http://www.thecre.com/fisma/wp-content/uploads/2012/05/sp800-146.pdf
[NISTusecases]	Cloud computing use cases - NIST http://www.nist.gov/itl/cloud/upload/Track-3-Session-5_security_report_out.pdf
[NSI]	Network Service Interface working group in Open Grid Forum https://www.ogf.org/gf/group_info/view.php?group=nsi-wg
[ODCA]	Open Data Center Alliance (ODCA) http://www.opendatacenteralliance.org/
[Okeanos]	Okeanos Infrastructure as a Service, GRNET Cloud service https://okeanos.grnet.gr/home/
[OpenNaaS]	OpenNaaS: Open platform for Network as a Service resources http://www.opennaas.org/
[SURFconext]	SURFconext collaboration infrastructure http://www.surfnet.nl/en/Samenwerkingsomgeving/SURFconext/Pages/ProjectCOIN.aspx
[TACAR]	TERENA Academic Certification Authority Repository https://www.tacar.org/
[TFstorage]	TERENA TF Storage http://www.terena.org/activities/tf-storage/
[TNC]	Demchenko, Y., J. van der Ham, C. de Laat, M. de Vos, S. Filiposka, D. Regvart, K. Baumann, T. Matselyukh, T. Karaliotas, D. Arbel, E. Escalona, T. Breach, “Open Cloud eXchange (OCX): Bringing Cloud Services to NRENs and Universities”, TNC, Ireland, 2014
[TrustRouter]	Trust Router, Internet Draft, March 25, 2012. http://www.ietf.org/id/draft-howlett-abfab-trust-router-ps-02.txt

Glossary

ABFAB	Application Bridging for Federated Access Beyond web
CAS	Central Authentication Service
CSP	Cloud Service Provider
DDOS	Distributed Denial of Service attack
DWDM	Dense Wavelength Division Multiplexing
GICTF	Global InterCloud Technology Forum
GLIF	Global Lambda Integrated Facility
gOCX	GÉANT Open Cloud Exchange
GOLE	GLIF Open Lightpath Exchange
HPC	High Performance Computing
IaaS	Infrastructure as a Service
ICADI	InterCloud Access and Delivery Infrastructure
ISOD-RG	Research Group on Infrastructure Services On-Demand provisioning
IX	Internet Exchange
MPP	Massively Parallel Processing
NaaS	Network as a Service
NIST	National Institute of Standards and Technology
NREN	National Research and Education Network
NSI	Network Services Interface
ODCA	Open Data Center Alliance
OGF	Open Grid Forum
PKI	Public Key Infrastructure
POP	Point of Presence
QoS	Quality of Service
R&E	Research and Education
SaaS	Software as a Service
SDN	Software Defined Networking
SLA	Service Level Agreement
SPOF	Single Point of Failure
SSO	Single Sign-on
TACAR	TERENA Academic CA Repository
TTP	Trusted Third Party
VM	Virtual Machine