Grid and Cloud Computing Technology: Interoperability and Standardization for the Telecommunications Industry

Overview

The European Telecommunications Standards Institute formed the Grid Technical Committee (TC GRID) in 2006 with participation from over 20 organizations spanning private sector, government, and academia. It had a specific mandate to address the convergence of the IT and Telecommunications industries in particular in the domain of interoperability. Grid computing, while having originated in the public sector, was experiencing a high degree of interest from commercial organizations. TC GRID intended to study the state of the art in grid computing standards, technology, and systems to evaluate their suitability for commercial adoption. In 2007 a Specialist Task Force (STF) was formed with a 2 year programme to provide a series of technical reports studying stakeholders, standards, technology, interoperability, and gaps. This white paper provides a selection of highlights from the technical reports, which are available free of charge from the ETSI website.

Grid vs Cloud

During 2008 and 2009 commercial interest shifted from "grids" to "clouds", with the availability of several on-demand compute and storage resources. What is the relationship between a "grid" and a "cloud"? The concepts are independent and complimentary. The similarities are that both aim to provide access to a large compute (CPU) or storage (disk) resource. Beyond that, a cloud utilizes virtualization to provide a uniform interface to a dynamically scalable underlying resource, with the intention that the virtualization layer conceals physical heterogeneity, geographical distribution, and faults. The cloud environment only provides direct support for single user or single organization access, and current models typically have a high cost to integrate computing, data, or network transfers from outside of the cloud. This model suits environments where compute and data resource needs can be isolated to a single location and rapid scaling (up or down) of compute, network, and data availability are important. Pricing models are variations on normalized CPU-hours, GB/day storage, and MB network I/O, or are based on a "cloud" product that can be licensed and used with local physical resources. Existing cloud systems use proprietary command line or browser (web) based interfaces. In contrast, grid computing aims to provide a standard set of services and software that enable the collaborative sharing of federated and geographically distributed compute and storage resources. It provides a security framework for identifying inter-organizational parties (both human and electronic), managing data access and movement, and utilization of remote compute resources.

Grid computing can benefit from the development of cloud computing by harnessing new commercially available compute and storage resources, and by deploying cloud technology on grid-enabled resources to improve the management and reliability of those resources via the virtualization layer. Cloud computing can benefit from grid concepts by integrating standard interfaces, federated access control, and distributed resource sharing. The current state of the art favours cloud computing for single organization commercial applications that can be deployed in their entirety onto a cloud environment. The dynamic provisioning of storage, compute power, and network bandwidth allows rapid scaling for intensive utilization either directly by the organization or by the public via Internet-based interfaces. Grid technology continues to dominate public sector scientific computing environments due to the collaborative nature of this work and the need to manage existing data sets and computing resources across organizational boundaries. The more advanced state of interface standardization within grid technology allows some degree of choice between various software and hardware systems. Deploying data and applications into a cloud environment, however, limits an organization to a single cloud provider or requires duplicated effort to repeat the deployment process for additional cloud environments.

Grid Requirements

In the same way the Web has federated and liberated unstructured information exchange between individuals, companies, and the academic community, there is a desire to do the same with data, software systems, services, compute resources, and storage resources. Progress has been made over the last decade to achieve this in an interoperable and extensible way, led primarily by grid computing initiatives. In many cases, users and providers of these more "advanced" services are forced to use proprietary systems, which do not inter-operate, do not scale, cannot easily be replaced, and cannot easily be reused by others. To address this, grid computing aims to provide "coordinated resource sharing and problem solving in a dynamic, multi-institutional virtual organization." A virtual organization (VO) is a set of individuals, organizations, and resources along with their mutually agreed sharing and access policy.

In order to achieve this, a grid environment must be:

- *Large scale*: Incorporating resources on a scale larger than what is typically provided by a single organization or accessible to an average user.
- *Dynamic*: Resource state and availability must be able to change during the course of system operation. This may be planned or unplanned. New resources may be added or removed without significant impact to the system behaviour.
- *Scalable*: Ability to increase in size by orders of magnitude.
- *Inter-operable*: Services provided by the components of the system can be replicated or replaced with alternate implementations, and dynamic resource interaction patterns are possible based on common resource interfaces.
- *Extensible*: New services and functionality can be incorporated or realized by integrating existing services in new ways.
- *Secure*: Providing a high level of trustworthiness between resources and users with configurable access control policies.
- *Heterogeneous*: Resources with different underlying properties integrated into the overall system via common interfaces. Ability to access these resources at abstract (common) interface layer, or via resource-specific interfaces.
- *Enabling*: Services which provide capabilities beyond what is already available to the user or a collaboration group must be achieved.
- *Usable*: The provided functionality must be easier to access and utilize than a comparable custom solution.
- *Manageable*: Providing facilities to track and control user and resource activity.
- *Federated*: Multiple wholly independent administrative domains with policies for resource sharing, user authorization, charging.
- *Geographically distributed*: Resources located at numerous independent sites.
- *Reliable*: Providing fault tolerance and a suitable level of service for long term continuous use.

Standards for Grid and NGN

The ETSI Technical Report 102 659-1 provides a details of major grid computing initiatives, commercial providers, users, key standards, and research projects. The following summarizes the key points of this report:

Open Grid Forum: The OGF has been leading body for grid-specific standards. Over the past few years these have primarily focused on SOA-based standards that fit the Web Services Resource Framework (WSRF) model, although there has been slow adoption of some of these standards and of the WSRF model in general.

Distributed Management Task Force: This group is responsible for the Common Information Model (CIM) standard, and several other standards related to enterprise system modelling, monitoring, and control. Recently DMTF has begun work on the Open Virtualization Format for managing virtualized environments.

European Telecommunications Standards Institute: ETSI is a leader in telecommunications standards and through TISPAN (Telecommunications and Internet converged Services and Protocols for Advanced Networking) has been developing both the NGN (Next Generation Network) architecture, and IMS (IP Multimedia Subsystem). These aim to provide advanced networking services and an environment supportive of rapid application deployment, possibly from third parties.

Internet Engineering Task Force: The IETF standardizes Internet-related protocols and interfaces. The primary IETF standards that impact grid computing are those around security: X.509, PKI, TLS/SSL, SASL.

Organization for the Advancement of Structured Information: OASIS produces several standards related to XML and Web Services. The two dominant standards are XML Access Control Language (XACML) and SAML (Security Assertion Mark-up Language).

World Wide Web Consortium: The W3C is responsible for the XML standards, the original XML/web-based RPC mechanism SOAP, and various Web Services standards.

Storage Network Industry Association: SNIA produces several standards related to storage network monitoring and control.

World Wide Web Consortium: The W3C is responsible for the XML standards, the original XML/web-based RPC mechanism SOAP, and various Web Services standards.

In addition to the various standards bodies, there have also been numerous significant grid projects and the deployment of national, international, or domain-specific grid infrastructures.

NGN and Grid Gap Analysis

The ETSI Technical Report 102 659-2 considers barriers to interoperability of grid technology from the perspective of gaps in existing standards. Five areas are covered in detail: Architecture, Service Level Agreements, Charging, Security, and Service Discovery.

In the presence of a multiplicity of network technologies and the resulting integration problems, the vertical integration of network layers is increasingly gaining in importance. These architectural elements are critical to the Next Generation Network (NGN) and to the formation and operation of dynamic large-scale grid infrastructures.

The architecture must support end to end services, where the quality of network services requested by the grid layer need to be independent of the underlying networking technologies. It is the responsibility of the network service provider to map and enforce the required quality of service. Currently neither NGN nor grid domains provide suitable interfaces or models to manage this relationship.

Horizontal integration of grid and NGN architectures needs to support the co-existence of multiple network service providers for widely distributed grid applications. These providers need to allow collaborative mechanisms for end-to-end service establishment. Current cross-network standards focus primarily on network provider interfaces and relatively static topologies. To realize the full potential of an integrated NGN and grid environment it is necessary to expose the cross-network routing and QoS interfaces to third-party applications for real-time dynamic service provisioning.

A major shortcoming of the grid standards landscape is the lack of a widely agreed upon architectural reference model. While the OGF have produced OGSA and adopted the Enterprise Grid Reference Model from the short lived Enterprise Grid Alliance, neither of these have found much practical use in the development of current grid infrastructures.

In the domain of Service Level Agreement and Quality of Service contracts and control, there is yet to be an established protocol or contract standard. The OGF have produced the WS-Agreement draft specification which is seeing gradual adoption. One challenge in establishing SLAs is the formation of an agreed set of SLA properties. While CIM, GLUE, and JSDL provide starting points for such a set of properties, there has been little success in finding wide-spread adoption of any single ontology, thereby impeding the application of an SLA. A similar lack of agreed SLA properties exists in the NGN domain. SLAs will need to be formed dynamically, implying automated systems for matching resource requests with resource providers and forming contracts with usage, QoS, and charging terms attached. Satisfying the QoS specified by an SLA implies an ability to monitor the various system layers in an integrated way and at an appropriate level of granularity. Current monitoring systems often do not provide suitable end-to-end service monitoring facilities or are unable to differentiate resource utilization by multiple users.

In terms of security, X.509 based Public Key Infrastructure (PKI) has been one of the great successes of grid technology, and the standards surrounding X.509 PKI have been widely adopted. These have also been integrated into some GSM-based SIM card for mobile devices, and there are plans and standards describing the widespread use of PKI in the telecoms domain, with device and end-user PKI identity tokens. The issue of key distribution, binding devices to particular users, and trust of device identity tokens has many similarities with user, host, and service identities in Grids, and both lack standards to guide development.

Furthermore, there are significant standards gaps around the issue of authorization in the grid domain. SAML (Security Assertion Markup Language) and XACML (eXtensible Access Control Markup Language) are the only two broadly applicable authorization standards, however their complexity makes them difficult to use in practice. There is a need for a simplified authorization policy language. Furthermore, there is significant scope for standards concerning authorization policy management: sharing, merging, and updating security policies in an efficient, clear, and secure manner.

A standard model which defines a Virtual Organization, membership, capabilities, and policies would provide an operational framework to improve VO-centric services, in contrast to the current focus on either user- or site-centric services. Finally, there are no standards for data provenance, an important issue in many domains: science, medicine, and financial services, to name a few. Only bespoke solutions for auditing data provenance are available.

Charging is an area where NGN and the commercial nature of network operations has significant experience, while the open collaborative origin of grid computing has resulted in little attention being given to this important topic. The complex and dynamic nature of grid usage patterns makes it a significant challenge to establish pricing models that can be implemented in practice to recover costs from the various providers involved in grid and NGN resource usage. The NGN model allows Billing On Behalf of Others (BOBO), such that a customer of an NGN service provider could elect to enable end users to purchase many products via an NGN account. Some effort has been made in the grid domain to produce Usage Records which could usefully be aligned with NGN Charging Detail Records (CDR) specifications. No standards exist for on line charging for Grid Services. There is scope to develop these jointly with the evolving NGN Standards. It is important that the use of these records for billing customers not limit the flexibility of Service Providers to develop custom pricing models.

Lastly, Service Discovery is a key issue for grid users as the discovery of resources to satisfy a particular need must be done quickly and efficiently, taking into account several characteristics and requirements of the particular usage pattern. Service registry mechanisms can be adapted for this purpose, however these typically do not support a high degree of dynamism (i.e. updates to the state of existing services in the registry, or the addition and removal of entries). Various grid service registries such as Monitoring and Discovery Service (MDS), Grid Resource Information Service (GRIS) and the associated Grid Information Index Service (GIIS), Relational Grid Monitoring Architecture (R-GMA), and Berkley Database Information Index (BDII) and others have been utilized by various grid infrastructures to varying degrees of success, but replication of such services indicates the lack of a single successful standard for resource registration and discovery. Resource and service coordination is also important given the mix of components which are provisioned and accessed during a typical grid usage scenario, however the complexity of standardizing this led the OGF working group Component Description, Deployment and Lifecycle Management (CDDLM) to abandon efforts to form a standard, despite some progress in the area. Currently the LDAP-based BDII system gained the highest level of adoption.

Testing Framework

As part of the TC GRID interoperability effort, Grid Plugtests have been organized for the past 6 years. These have focused on the execution of standard distributed computing problems over a grid infrastructure with the goal being the fastest solution to a given problem. The STF has been tasked with the refinement of the Grid Plugtest to include a more formal testing framework. This is using five different models for grid interaction: user driven, parallel deployment, gateways, adaptors and translators, and common interfaces. The user driven model requires the end user to explicitly deploy their grid workflow and application onto distributed resources on a per-resource basis, thus replicating effort and creating static or one-off application to resource bindings. The parallel deployment model allows multiple independent users to access the same shared underlying resource with parallel execution of independent applications. Gateways act as forwarding services or proxies for underlying resources and conceal the mapping of an application deployment to the actual physical resource from the end user. Adaptors and translators allow incompatible interfaces and data formats to be joined. Common interfaces allow end users to utilise a common system model and well defined interfaces to dynamically select different resources.

Upcoming Grid Plugtests will allow organizations to test the suitability of their hardware, software, and systems with a standardized grid application deployment and workflow execution.

Making Better Standards

ETSI has embarked on a mission to improve the quality of its standards by introducing guidelines for writing and testing standards and subsequent implementations. A comprehensive website – http://portal.etsi.org/mbs/ – and set of guides have been produced to facilitate the standard writing process. Furthermore, interoperability testing of implementations is a critical step in validating a standard. The Methods for Testing and Specification (MTS) group within ETSI provides further guidance for best practice in developing robust standards. With its expertise in developing and testing standards, ETSI also organizes and hosts numerous "Plugtest" events each year where manufacturers and operators come together to test interoperability of components, software, and systems.

About the Grid Technical Committee

TC GRID's primary goal is to address issues associated with the convergence between IT and Telecommunications, with particular reference to the lack of interoperable grid solutions in situations which involve contributions from both the IT and Telecom industries. This places the focus on scenarios where connectivity goes beyond the local network. The TC GRID activities have an emphasis on interoperable grid applications and services based on global standards and the validation tools to support these standards.

The approach is to actively involve and support grid stakeholders by complementing existing activities. Specifically, TC GRID will address interoperability aspects of end-to-end applications and develop formal test specifications with the aim of assuring interoperability.

About ETSI Specialist Task Forces

STFs are teams of highly-skilled experts working together over a pre-defined period to draft an ETSI standard under the technical guidance of an ETSI Technical Body and with the support of the ETSI Secretariat. The task of the STFs is to accelerate the standardization process in areas of strategic importance and in response to urgent market needs. For more information, please visit:
http://portal.etsi.org/stfs/process/home.asp

The work carried out here is co-financed by the EC/EFTA in response to the EC's ICT Standardisation Work Programme.

About ETSI

ETSI produces globally-applicable standards for Information and Communications Technologies (ICT), including fixed, mobile, radio, converged, broadcast and internet technologies and is officially recognized by the European Commission as a European Standards Organization. ETSI is a not-for-profit organization whose 700 ETSI member organizations benefit from direct participation and are drawn from 60 countries worldwide. For more information, please visit: www.etsi.org

Contact ETSI Grid Technical Committee

http://www.etsi.org
http://portal.etsi.org/grid
Listserv: grid@list.etsi.org
Grid task force: stf331@etsi.org

ETSI Secretariat
650, route des Lucioles
06921 Sophia-Antipolis Cedex
FRANCE
+33 4 92 94 42 00