Defining Inter-Cloud Architecture

Yuri Demchenko, Rudolf Strijkers, Canh Ngo, Mihai Cristea, Mattijs Ghijsen, Cees de Laat
System and Network Engineering Group
University of Amsterdam
Amsterdam, The Netherlands
e-mail: {y.demchenko, strijkers, c.t.ngo, M.L.Cristea, M.Ghijsen, delaat}@uva.nl

***Abstract*—**

This poster paper presents on-going research to develop the Inter-Cloud Architecture that should address problems with multi-domain heterogeneous Cloud based applications integration and interoperability, including integration and interoperability with legacy infrastructure services. The Cloud technologies actually represent a new step in evolutional computing and communication technologies development chain. Cloud technologies are evolving as a common way of infrastructure services and resources virtualisation and provisioning on-demand. In this way they bring applications and infrastructure services mobility and physical/hardware platform independency to the existing distributed computing and networking technologies. The paper refers to existing standards in Cloud Computing, in particular, recently published NIST Cloud Computing Reference Architecture (CCRA). The proposed Inter-Cloud Architecture combines commonly adopted Cloud service models such as IaaS, PaaS, SaaS in one multilayer model with corresponding inter-layer interfaces. The paper also briefly presents the architectural framework for Cloud based infrastructure services provisioned on-demand being developed by authors that allows optimised provisioning of both computing, storage and networking resources. The proposed architecture is intended to provide a conceptual model for developing Inter-Cloud middleware and in this way will facilitate Clouds interoperability and integration.

Keywords-Inter-Cloud Architecture; Cloud Computing Reference Architecture Interoperability; Architectural framework for Cloud infrastructure services provisioned on-demand; Cloud middleware.

I. INTRODUCTION

Cloud computing technologies [1, 2] are emerging as infrastructure services for provisioning computing and storage resources on-demand in a simple and uniform way and may involve multi-provider and multi-domain resources, including integration with the legacy services and infrastructures. In this way, Clouds introduce a new type of services and a new abstraction layer for general infrastructure services virtualisation (similar to utilities).

Current development of the Cloud technologies demonstrate movement to developing inter-Cloud models, architectures and integration tools that could allow integrating Cloud based infrastructure services into existing enterprise and campus infrastructures, on one hand, and provide common/interoperable environment for moving existing infrastructures and infrastructure services to virtualised Cloud environment. More complex and community oriented use of Cloud infrastructure services will require developing new service provisioning and security models that could allow creating complex project and group oriented infrastructures provisioned on-demand and across multiple providers.

The paper presents on-going research at the University of Amsterdam to develop the Inter-Cloud Architecture (ICA) that should address problems with multi-domain heterogeneous Cloud based applications integration and interoperability, including integration and interoperability with legacy infrastructure services, and to facilitate interoperable and measurable intra-provider infrastructures and Clouds federation. The papers refers to the architectural framework for provisioning Cloud Infrastructure Services On-Demand [3] being developed by authors in a number of currently running projects such as GEANT3 [4] and GEYSERS [5] that provides a basis for defining the proposed Inter-Cloud architecture.

II. MOTIVATION FOR ICA DEFINITION

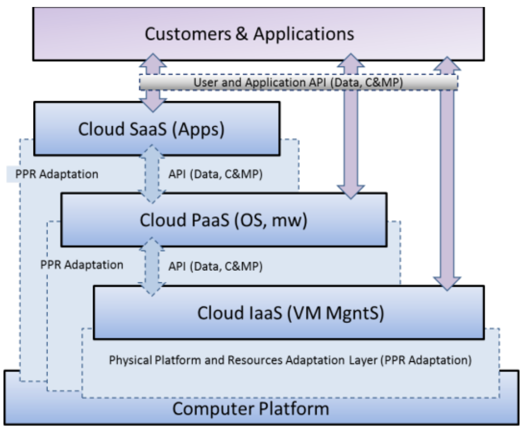
The Cloud definition by NIST [1] since its publication few years ago became commonly accepted: "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

The following five basic Clouds characteristics include: on-demand self-service, broad network access, resource pooling, rapid elasticity, measured Service.

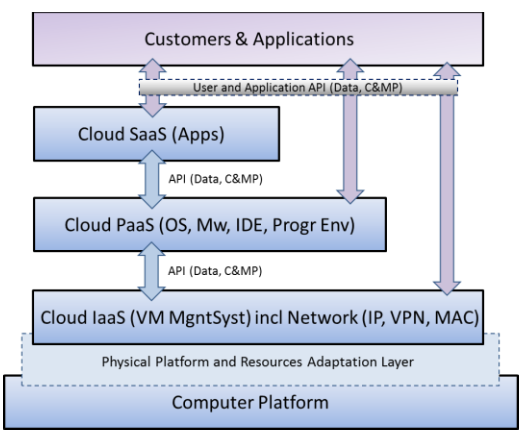
Three basic service provisioning models are defined: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). However, in fact when used as part enterprise infrastructure, all types of Cloud services require some kind of infrastructure support. The difference is whose concern the infrastructure provisioning and operation.

In particular case of Cloud IaaS, the capabilities are provided to the consumer to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

Figure 1, a illustrates current relation between basic Cloud service models IaaS, PaaS, SaaS that expose in most cases standard based interface to user services or applications but actually use proprietary interface to the physical provider platform. However in case of multiple heterogeneous Cloud services integration in one integrated infrastructure or application, Cloud services from different service models and layers need to interact. This motivates definition of the Inter-Cloud Architecture that is depicted on Figure1, b as multilayer architecture with interlayer interfaces.



(a) Current relation between Cloud service models



(b) current relation between Cloud service models

Figure 1. Inter-Cloud Architecture for Cloud interoperability and integration.

III. GENERAL USE CASES FOR ICA

The two basic use cases for Inter-Cloud architecture can be considered: large project-oriented scientific infrastructure provisioning including dedicated transport network infrastructure, and periodic semester based educational course that requires computer laboratory facilities to setup, operated and suspended till the next semester. Both cases should allow the whole infrastructure of computers, storage, network and other utilities to be provisioned on-demand, physical platform independent and allow integration with local persistent utilities and legacy services and applications.

Figures 2 illustrates the typical e-Science infrastructure that includes Grid and Cloud based computing and storage resources, instruments, control and monitoring system, visualization system, and users represented by user clients. The diagram also reflects that there may be different types of connecting network links: high-speed and low-speed which both can be permanent for the project or provisioned on-demand.

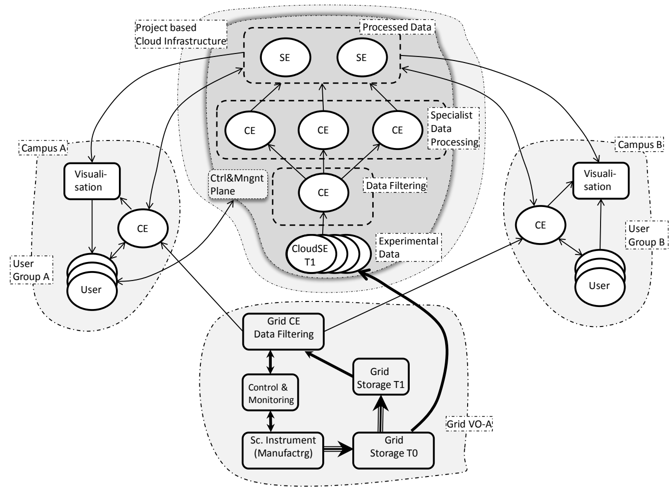


Figure 2. Project oriented collaborative infrastructure containing Grid based Scientific Instrument managed by Grid VO-A, 2 campuses A and B, and Cloud based infrastructure provisioned on-demand.

The figure can also illustrate a typical usecase when a high performance infrastructure is used by two or more cooperative users/researcher groups in different locations. In order to fulfill their task (e.g. cooperative image processing and analysis) they require a number of resources and services to process raw data on distributed Grid or Cloud data centers, analyse intermediate data on specialist applications and finally deliver the result data to the users/scientists. This use case includes all basic components of the typical e-Science research process: data collection, initial data mining and filtering, analysis with special scientific applications, and finally presentation and visualisation to the users.

IV. ABSTRACT MODEL FOR CLOUD BASED INFRASTRUCTURE SERVICES PROVISIONING

Figure 2 below illustrates the abstraction of the typical project or group oriented Virtual Infrastructure (VI) provisioning process that includes both computing resources and supporting network that commonly referred as infrastructure services. The figure also shows the main actors involved into this process, such as Physical Infrastructure Provider (PIP), Virtual Infrastructure Provider (VIP), Virtual Infrastructure Operator (VIO).

The required supporting infrastructure services are depictured on the left side of the picture and includes functional components and services used to support normal operation of all mentioned actors. The Virtual Infrastructure Composition and Management (VICM) layer includes the Logical Abstraction Layer and the VI/VR Adaptation Layer facing correspondingly lower PIP and upper Application layer. VICM related functionality is described below as related to the proposed Composable Services Architecture (CSA).

The proposed architecture is a SOA (Service Oriented Architecture) based and uses the same basic operation principle as known and widely used SOA frameworks, what also provides a direct mapping to the possible VICM implementation platforms such as Enterprise Services Bus (ESB) [6] or OSGi framework [7].

The infrastructure provisioning process, also referred to as Service Delivery Framework (SDF), is adopted from the TeleManagement Forum SDF [8] with necessary extensions to allow dynamic services provisioning. It includes the following main stages: (1) infrastructure creation request sent to VIO or VIP that may include both required resources and network infrastructure to support distributed target user groups and/or consuming applications; (2) infrastructure planning and advance reservation; (3) infrastructure deployment including services synchronization and initiation; (4) operation stage, and (5) infrastructure decommissioning. The SDF combines in one provisioning workflow all processes that are run by different supporting systems and executed by different actors.

Physical Resources (PR), including IT resources and network, are provided by Physical Infrastructure Providers (PIP). In order to be included into VI composition and provisioning by the VIP they need to be abstracted to Logical Resource (LR) that will undergo a number of abstract transformations including possibly interactive negotiation with the PIP. The composed VI need to be deployed to the PIP which will create virtualised physical resources (VPR) that may be a part, a pool, or a combination of the resources provided by PIP.

The deployment process includes distribution of common VI context, configuration of VPR at PIP, advance reservation and scheduling, and virtualised infrastructure services synchronization and initiation, to make them available to Application layer consumers.

The proposed abstract model provides a basis for ICA definition and allows outsourcing the provisioned VI operation to the VI Operator (VIO) who is from the user/consumer point of view provides valuable services of the required resources consolidation - both IT and networks, and takes a burden of managing the provisioned services.

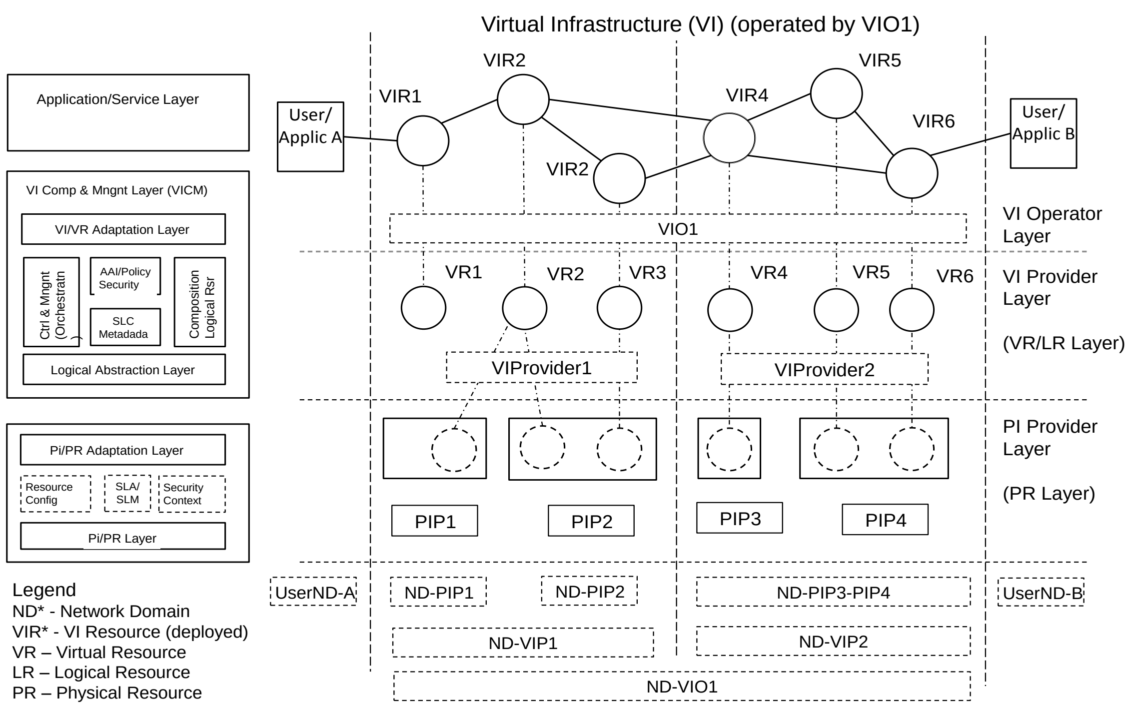


Figure 2. Main actors, functional layers and processes in on-demand infrastructure services provisioning

V. INFRASTRUCTURE SERVICES MODELING FRAMEWORK

The Infrastructure Services Modeling Framework (ISMF) provides a basis for virtualization and management of infrastructure resources, including description, discovery, modeling, composition, and monitoring. In this paper we mainly focus on the description of resources and the lifecycle of these resources. The described model in this section is being developed in the GEYSERS project [5].

A. *Resource Modeling*

The two main descriptive elements of the ISMF are the infrastructure topology and descriptions of resources in that topology. Besides these main ingredients, the ISMF also allows for describing QoS attributes of resources, energy related attributes, and attributes needed for access control.

The main requirements for the ISMF are, that it should allow for describing Physical Resources (PR) as well as Virtual Resources (VR). Describing physical aspects of a resource means that a great level of detail in the description is required while describing a virtual resource may require a more abstract view. Furthermore, the ISMF should allow for manipulation of resource descriptions such as partitioning and aggregation. Resources on which manipulation takes place, and resources that are the outcome of manipulation are called Logical Resources (LR).

The ISMF is based on semantic web technology. This means that the description format will be based on the Web Ontology Language (OWL) [9]. This approach ensures the ISMF is extensible and allows for easy abstraction of resources by adding or omitting resource description elements. Furthermore, this approach has enabled us to re-use the Network Description Language [10] to describe infrastructure topologies.

B. *Virtual Resource Lifecycle*

Figure 3 illustrates relations between different resource presentations along the provisioning process that can also be defined as the Virtual Resource lifecycle.

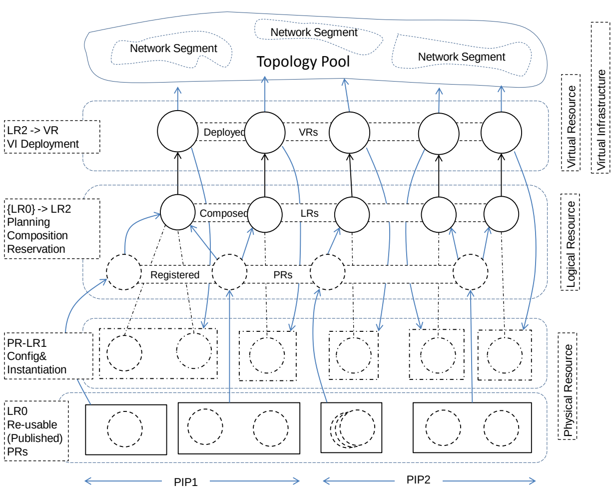


Figure 3. Relation between different resource presentations in relation to different provisioning stages.

The Physical Resource information is published by a PIP to the Registry service serving VICM and VIP. This published information describes a PR. The published LR information presented in the commonly adopted form (using common data or semantic model) is then used by VICM/VIP composition service to create the requested infrastructure using a combination of (instantiated) Virtual Resources and interconnecting them with a network infrastructure. In its own turn the network can be composed of a few network segments run by different network providers.

It is important to mention that physical and virtual resources discussed here are in fact complex software enabled systems with their own operating systems and security services. The VI provisioning process should support the smooth integration into the common federated VI security infrastructure by allowing the definition of a common access control policy. Access decisions made at the VI level should be trusted and validated at the PIP level. This can be achieved by creating dynamic security associations during the provisioning process.

VI. FUTURE DEVELOPMENT

The paper presents on-going research at the University of Amsterdam to develop the Inter-Cloud Architecture (ICA) that should address problems with multi-domain heterogeneous Cloud based applications integration and inter-provider and inter-platform interoperability.

The presented research is planned to be contributed to the Open Grid Forum Research Group on Infrastructure Services On-Demand provisioning (ISOD-RG) [11], where the authors play active role.

ACKNOWLEDGEMENTS

This work is supported by the FP7 EU funded Integrated project The Generalised Architecture for Dynamic Infrastructure Services (GEYSERS, FP7-ICT-248657).

REFERENCES

[1] NIST SP 800-145, “A NIST definition of cloud computing”, [online] Available: http://csrc.nist.gov/publications/drafts/ 800-145/Draft-SP-800-145_cloud-definition.pdf

[2] NIST Cloud Computing Reference Architecture, v1.0. [Online] http://collaborate.nist.gov/twiki-cloud-computing/bin/view/ CloudComputing/ReferenceArchitectureTaxonomy

[3] Generic Architecture for Cloud Infrastructure as a Service (IaaS) Provisioning Model, Release 1. SNE Techn. Report SNE-UVA-2011-03, 15 April 2011. [Online] http://staff.science.uva.nl/~demch/ worksinprogress/sne2011-techreport-2011-03-clouds-iaas-architecture-release1.pdf

[4] GEANT Project. [Online] http://www.geant.net/pages/home.aspx

[5] Generalised Architecture for Dynamic Infrastructure Services (GEYSERS Project). [Online] http://www.geysers.eu/

[6] D. Chappell, ENTERPRISE SERVICE BUS, O’Reilly, June 2004. 247 pp.

[7] OSGi Service Platform Release 4, Version 4.2. [online] Available: http://www.osgi.org/Download/Release4V42

[8] TMF Service Delivery Framework. [Online] http://www.tmforum.org /servicedeliveryframework/4664/home.html

[9] OWL 2 Web Ontology Language [online] Available: http://www.w3.org/TR/owl2-overview/

[10] J. van der Ham, F.Dijkstra, P.Grosso, R. van der Pol, A.Toonk, C. de Laat, "A distributed topology information system for optical networks based on the semantic web", Elsevier Journal on Optical Switching and Networking, Volume 5, Issues 2-3, June 2008, pp 85-93

[11] Open Grid Forum Research Group on Infrastructure Services On-Demand provisioning (ISOD-RG). [Online]. http://www.ogf.org/gf/event_schedule/index.php?event_id=17