DEFINING GENERIC ARCHITECTURE FOR CLOUD IAAS PROVISIONING MODEL

Yuri Demchenko, Cees de Laat
System and Network Engineering Group, University of Amsterdam, Amsterdam, The Netherlands

Aleksej Mavrin
Verizon Nederland B.V., Amsterdam, The Netherlands

Keywords: Cloud computing, Infrastructure as a Service (IaaS), on-Demand Infrastructure Services Provisioning (ISOD), Composable Services Architecture (CSA), GEANT Multidomain Bus (GEMBus).

Abstract: Infrastructure as a Service (IaaS) is one of the provisioning models for Clouds as defined in the NIST Clouds definition. Although widely used, current IaaS implementations and solutions doesn't have common and well defined architecture model. The paper attempts to define a generic architecture for IaaS based on current research by authors in developing novel architectural framework for Infrastructure Services On-Demand (ISOD) provisioning that is originated from the telecommunication and networking area and allows for combined network and IT resources provisioning. The paper proposes the Composable Services Architecture (CSA) for dynamically configurable virtualised services. The proposed CSA includes such important component as the Services Delivery Framework (CSA SDF) that defines the services provisioning workflow and supporting infrastructure for provisioned services lifecycle management. The CSA SDF extends existing lifecycle management frameworks with additional stages such as "Registration and Synchronisation" and "Provisioning Session Binding" that specifically target such scenarios as the provisioned services recovery or re-planning/migration and provide necessary mechanisms for consistent security services provisioning as an important component of the provisioned on-demand infrastructure. The paper also describes the GEMBus (GEANT Multidomain Bus) that is considered as a CSA middleware platform. The presented architecture is the result of the on-going cooperative effort of the two EU projects GEANT3 JRA3 Composable Services and GEYSERS.

1 INTRODUCTION

Cloud technologies (NIST, GFD.150) are emerging as infrastructure services for provisioning computing and storage resources, and expectedly they will evolve into the general IT resources. Cloud Computing can be considered as a natural evolution of the Grid Computing technologies to more open infrastructure-based services.

The growing interest and adoption of the Cloud based service provisioning, operational and business models are facilitated by the fact that different groups of users, vendors, providers and operators see new opportunities related to their own specific needs and interests: scientist and specialist users like simplicity of setup and elasticity of execution/computation environment; computer scientists and programmers see a new environment for using and developing new programming models; providers, and in particular telecommunication providers, and operators expect to benefit from new provisioning models and new market for complex infrastructure services.

The current Cloud services implement 3 basic provisioning models (as defined in the NIST document (NIST)): Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). IaaS suggests involving different types of resources, of which one is the network connectivity with guaranteed Quality of Service (QoS), and availability of user controlled infrastructure operation via dynamically created control and management planes/functions. Most of currently available Cloud providers are positioned as

Demchenko Y., de Laat C. and Mavrin A..
DEFINING GENERIC ARCHITECTURE FOR CLOUD IAAS PROVISIONING MODEL.
DOI: 10.5220/0003392600790085
In *Proceedings of the 1st International Conference on Cloud Computing and Services Science* (CLOSER-2011), pages 79-85
ISBN: 978-989-8425-52-2

PaaS; although some of them may claim to do IaaS but in practice this turns to be rather PaaS like in case of Amazon EC2 Cloud or RackSpace Cloud as in fact they deliver single provider services based on their private network infrastructure.

Despite a rapid grow, Cloud technologies are lack of well-defined architectural framework(s) and operational models. Current industry standardisation activity is primarily focused on functional interoperability of deployment platforms and components supporting basic/core usage scenarios such as applications and Virtual Machine (VM) images creation, deployment and management, services request and execution management.

This paper attempts to define a generic architecture for the Cloud Infrastructure as a Service provisioning model which is defined as the Composable Services Architecture (CSA). The CSA extends virtualisation and dynamic service provisioning concepts to complex infrastructure services that are composed of multiple composable component services and resources. The CSA is SOA based and built upon industry adopted Enterprise Service Bus (ESB).

The presented architecture is the result of the ongoing cooperative effort of the two EU projects GEANT3 JRA3 Composable Services (GEANT Project) and GEYSERS (GEYSERS Project) and currently considered to be contributed to the Open Grid Forum (OGF) standardisation activity. Current development is based on previous works by the authors in the framework of the EGEE and Phosphorus projects (EGEE Project), (Phosphorus Project) that have been resulted in proposing the general Complex Resource Provisioning (CRP) model (Demchenko et al., 2009) that includes such main stages as reservation, deployment, access, and decommissioning.

The paper is organized as follows. Section 2 analyses the typical infrastructure for e-Science applications that includes computing, storage, visualisation and their connection to network infrastructures. Section 3 presents the proposed Composable Services Architecture and section 4 describes the proposed Services Delivery Framework (CSA SDF) that defines the on-demand service provisioning sequence and workflow. Section 5 provides information about the ongoing development of the GEMBus that is considered as a middleware and enabling technology for the dynamically provisioned composable services integration. Section 6 provides general requirements and suggestions for building consisting security and refers to other ongoing works by the authors.

2 ON-DEMAND INFRASTRUCTURE SERVICES PROVISIONING

Two general use cases for on-demand infrastructure services provisioning can be considered for defining basic requirements to Cloud IaaS: large scientific infrastructure and transport network infrastructure provisioning. These use cases represent the two different perspectives in developing infrastructure services – users and application developers perspective, on one side, and providers perspective, on the other side. Users are interested in uniform and simple access to the resource and the services that are exposed as Cloud/Grid resources and can be easily integrated into the scientific or business workflow. Infrastructure providers are interested in infrastructure resource pooling and virtualisation to simplify their on-demand provisioning and extend their service offering and business model to Virtual Infrastructure provisioning.

Figure 1 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.

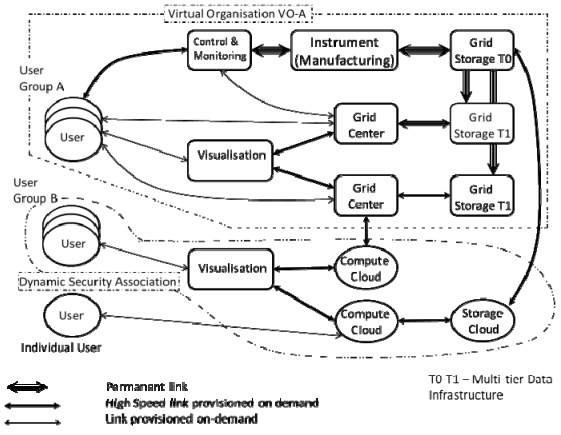

Figure 1: Components of the typical e-Science infrastructure involving multidomain and multi-tier Grid and Cloud resources and network infrastructure.

The figure illustrates possible different relations between users/actors in Clouds and Grids. Grid architecture is built around so-called Virtual Organisations (VO) that are defined as collaborative associations/federations of the member organisations

to share complex resources which are however remaining under control and in ownership of the VO member organisations. Grids natively supports project oriented collaborative environment with sharing resources committed/donated by the VO members. Clouds are considered as a technology for provisioning complex computing resources, applications and increasingly infrastructure resources and services on-demand on the pay-per-use base.

Typically business relations between provider and customer are expressed in the Service Level Agreement (SLA) that defines the services provided by the provider, including security services that are provided as a part of the provider Cloud environment. The offered/provided services are uniform and cannot be modified or configured by user what creates problems for their integration into the existing user infrastructure or building effective project based collaborative environment. With wider adoption of the Cloud infrastructure services and their integration into organisational IT infrastructure the demand for dynamically configurable/ manageable composable services will be growing. The solution for mentioned problems can be seen in provisioning manageable, dynamically configured services that support all stages of on-demand infrastructure services provisioning. This problem is being researched as a part of the GEANT3 JRA3 Composable Services activity (GEANT Project) and GEYSERS project Logical Infrastructure Composition Layer (LICL) definition and design (GEYSERS Project).

3 THE COMPOSABLE SERVICES ARCHITECTURE (CSA)

The Infrastructure as a Service provisioning involves dynamic creation of infrastructure consisting of different types of resources together with necessary control and management planes, all provisioned on-demand. The proposed CSA provides a framework for the design and operation of the composite/complex services provisioned on-demand. It is based on the component services virtualisation, which in its own turn is based on the logical abstraction of the (physical) component services and their dynamic composition. Composite services may also use the Orchestration service provisioned as a CSA infrastructure service to operate composite service specific workflow.

Figure 2 shows the major functional components of the proposed CSA and their interaction. The central part of the architecture is the CSA middleware that should ensure smooth service operation during all stages of the composable services lifecycle.

Composable Services Middleware (CSA-MW) provides common interaction environment for both (physical) component services and complex/composite services, built with them. Besides exchanging messages, CSA-MW also contains/provides a set of basic/general infrastructure services required to support reliable and secure (composite) services delivery and operation:

- Service Lifecycle Metadata Service (MD SLC) that stores the services metadata and in particular the services state and the provisioning session context.
- Registry service that contains information about all component services and dynamically created composite services. The Registry should support automatic services registration.
- Logging service that can be also combined with the monitoring service.
- Middleware Security services that ensure secure operation of the CSA/middleware.

Note, both logging and security services can be also provided as component services that can be composed with other services in a regular way.

The CSA defines also Logical Abstraction Layer for component services and resources that is necessary part of creating services pool and virtualisation. Another functional layer is the Services Composition layer that allows presentation of the composed/composite services as regular services to the consumer.

The Control and Management plane provides necessary functionality for managing composed services during their normal operation. It may include Orchestration service to coordinate component services operation, in a simple case it may be standard workflow management system.

CSA defines a special adaptation layer to support dynamically provisioned Control and Management plane interaction with the component services which to be included into the CSA infrastructure must implement adaptation layer interfaces that are capable of supporting major CSA provisioning stages, in particular, service identification, services configuration and metadata including security context, and provisioning session management.

The application or service middleware layer can be defined as an additional upper CSA architecture layer to provide application – infrastructure interface

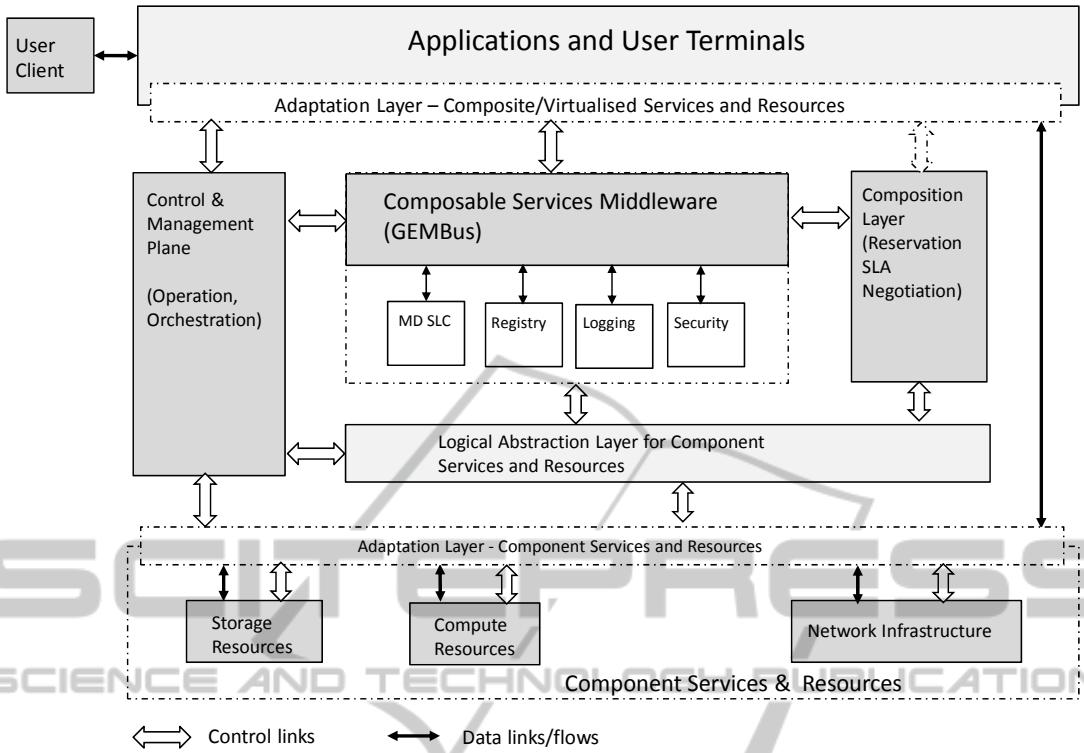

Figure 2: Composable Service Architecture and main functional components.

with the underlying virtual infrastructure provisioned by CSA, but this functionality is typically addressed by the applications provider themselves.

4 CSA SERVICE DELIVERY FRAMEWORK

The proposed CSA Service Delivery Framework (CSA SDF) implements the Service Delivery Framework defined by the TeleManagement Forum (TMF) (TMF) and extends it with additional stages to address secure composable services operation and integration into the heterogeneous multidomain Cloud IaaS environment. Figure 3 illustrates the main service provisioning stages:

Service Request (including SLA negotiation). The SLA can describe QoS and security requirements of the negotiated infrastructure service along with information that facilitates authentication of service requests from users. This stage also includes generation of the Global Reservation ID (GRI) that will serve as a provisioning session identifier and will bind all other stages and related security context.

Composition/Reservation that also includes **Reservation Session Binding** with GRI what provides support for complex reservation process in potentially multidomain multi-provider environment. This stage may require access control and SLA/policy enforcement.

Deployment, including services **Registration and Synchronisation.** Deployment stage begins after all component resources have been reserved and includes distribution of the common composed service context (including security context) and binding the reserved resources or services to the GRI as a common provisioning session ID. The Registration and Synchronisation stage (that however can be considered as optional) specifically targets possible scenarios with the provisioned services migration or re-planning. In a simple case, the Registration stage binds the local resource or hosting platform run-time process ID to the GRI as a provisioning session ID.

Operation (including Monitoring)**.** This is the main operational stage of the provisioned on demand composable services. Monitoring is an important functionality of this stage to ensure service availability and secure operation, including SLA enforcement.

Decommissioning stage ensures that all sessions are terminated, data are cleaned up and session security context is recycled. Decommissioning stage can also provide information to or initiate services usage accounting.

The two additional (sub-)stages can be initiated from the Operation stage and/or based on the running composed service or component services state, such as their availability or failure:

Re-composition or **Re-planning** that should allow incremental infrastructure changes.

Recovery/Migration can be initiated both the user and the provider. This process can use MD-SLC to initiate full or partial resources re-synchronisation, it may also require re-composition.

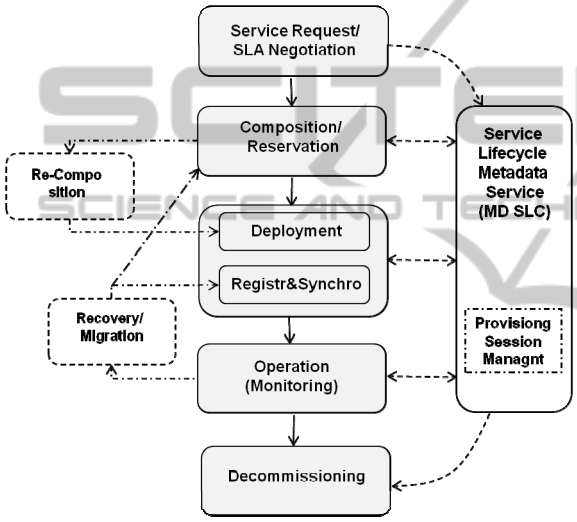

Figure 3: On-demand Composable Services Provisioning Workflow.

5 GEMBUS AS A CSA MDDLEWARE PLATFORM

GEANT Multidomain Bus (GEMBus) is being developed as a middleware for Composable Services (GEANT3 Project) in the framework of the GÉANT3 project aimed for creating a new generation of the pan-European academic and research network GÉANT. GEMBus incorporates the SOA services management paradigm (OASIS) in on-demand service provisioning. The Composable Services Architecture will span over different service interactions, from the infrastructure up to application elements and will provide functionality to define, discover, access and combine services in the GÉANT environment. The goal of GEMBus is to establish seamless access to the network infrastructure and the services deployed upon it, using direct collaboration between network and applications, and therefore providing more complex community-oriented services through their composition.

GEMBus is built upon the industry accepted Enterprise Service Bus (ESB) (Chappell, 2004) and extends it with the necessary functional components and design patterns to support multidomain services and applications. Figure 4 illustrates the suggested GEMBus architecture. GEMBus infrastructure includes three main groups of functionalities:

- GEMBus Messaging Infrastructure (GMI) that includes the messaging backbone and other message handling services such as routing, configuration services, secure messaging, and event handler/interceptors. GMI supports both SOAP-based and RESTful services conforming to Representational State Transfer architecture (REST) (Pautasso et al., 2008).
- GEMBus infrastructure services that support reliable and secure composable services operation and the whole services provisioning process. These include such services as service registries, composition and orchestration, security and access control, logging and monitoring, and the also important Lifecycle Metadata Service.
- Component services, although typically provided by independent parties, will need to implement special GEMBus adaptors or use special "plug-in sockets" that allow their integration into the GEMBus/CSA infrastructure.

The following functionalities are essential to enable GEMBus operation in the multidomain heterogeneous service provisioning environment:

- Service registries supporting service registration and discovery. Registries are considered as an important component to allow cross-domain heterogeneous services integration and metadata management during the whole services lifecycle.
- Security, access control, and logging should provide consistent services and security context management during the whole provisioned services lifecycle.
- Service Composition and Orchestration models and mechanisms should allow integration with the higher level scientific or business workflow.

The GEMBus and GMI in particular are built on the top of the standard Apache/Fuse messaging infrastructure that includes the following components (Fuse ESB, Apache ServiceMix):

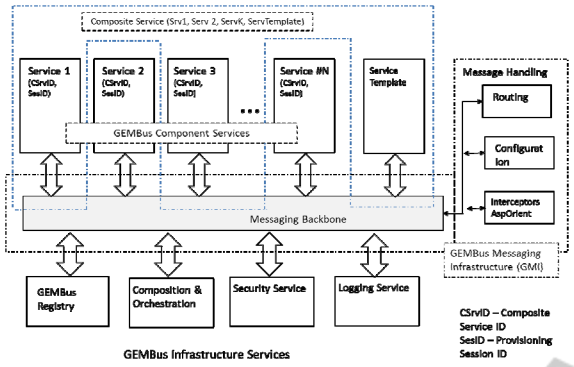

Figure 4: GEMBus infrastructure, including component services, service template, infrastructure services, and core message-processing services.

- Fuse Message Broker (Apache ActiveMQ) messaging processor
- Fuse Mediation Router (Apache Camel) normalised message router

The GEMBus services and applications can be deployed on the standard Fuse or Apache ESB server. As component services that can be integrated with the standard OSGi (OSGi) and Spring (Spring Security) compliant service development frameworks and platforms such as Fuse Services Framework/Apache CXF and Fuse ESB/Apache ServiceMix.

6 CSA SECURITY INFRASTRUCTURE

Providing consistent security services in CSA and GEMBus is of primary importance due to potentially multi-provider and multi-tenant nature of Clouds IaaS environment. The CSA security infrastructure (CSA-Security) should address two aspects of the IaaS operation and dynamic security services provisioning:

- Provide security infrastructure for secure IaaS operation, including access control and SLA and policy enforcement for all interacting roles and components in CSA, secure messaging and transport services.
- Provisioning dynamic security services, including creation and management of the dynamic security associations, as a part of the provisioned complex/composite services or virtual infrastructures.

The first task is a traditional task in security engineering, while dynamic provisioning of managed security services remains a problem and requires additional research.

An important issue in building consistent security services for dynamically provisioned virtual infrastructures is the Security Services Lifecycle Management (SSLM) that extends the described above CSA SDF service lifecycle management model with additional sub-stages and functions to bind dynamic security context to the general provisioning session and Cloud virtualisation and hosting platform in such a way that to ensure all operations on the virtual infrastructure and user data to be secured during their whole lifecycle. It is described in details in earlier authors' paper (Demchenko et al., 2010).

CSA-Security and GEMBus should implement the following basic infrastructure security services to ensure normal operation of the virtual infrastructure:

- Access control (e.g. Authentication, Authorisation, Identity Management)
- Policy and SLA enforcement
- Data, messaging and communication security
- Additionally, auditing/logging and accounting.

CSA-Security should implement multi-layer security services including transport, messaging and application/data security, and additionally network layer security for distributed VPN based enterprise domains. Security and security services in the CSA and GEMBus design are applied at different layers and can be called from different functional components using standard/common security services interface. Security services are governed by related security policies.

Security services should support the whole provisioned/composable services lifecycle and consequently support session-related security context management.

Security services can be designed as: pluggable services operating at the messaging layer; OSGi bundles that can be dynamically added as composable services to other composable services to form an instant virtual infrastructure; or exposed as Web services and be integrated with other services by means of higher level workflow management systems.

7 SUMMARY AND FUTURE DEVELOPMENTS

This paper presents the ongoing research on developing architecture and framework for dynamically provisioned and reconfigurable

infrastructure services to support modern e-Science and high-technology industry applications that require both high-performance computing resources (provisioned as Grids or Clouds) and high-speed dedicated transport network.

The paper proposes the Composable Services Architecture (CSA) that is intended to provide a conceptual and methodological framework for developing dynamically configurable virtualised infrastructure services.

The proposed CSA is currently being implemented in the framework of the GEANT3 Project as an architectural component of the GEANT Multidomain service bus (GEMBus). The GEMBus extends the industry adopted Enterprise Service Bus (ESB) technology with the additional functionality to support multidomain services provisioning. The GEMBus infrastructure intended to allow dynamic composition of the infrastructure services to support collaboration of the distributed groups of researchers.

The concepts and solutions presented in this paper are intended to be offered as a contribution to the Open Grid Forum (OGF) Research Group on On-Demand Infrastructure Services Provisioning (ISOD-RG) initiated by the authors and the involved projects (for details see materials of ISOD BoF at OGF30 (ISOD BoF) and planned ISOD-RG meeting at OGF31).

The authors believe that concepts proposed in this paper will provide a good basis for the further discussion among researchers about defining an architecture for dynamically configured virtualised infrastructure services as a part of the Clouds IaaS model.

ACKNOWLEDGEMENTS

This work is supported by the FP7 EU funded project GEANT3 (FP7-ICT-238875), and the FP7 EU funded Integrated project The Generalised Architecture for Dynamic Infrastructure Services (GEYSERS, FP7-ICT-248657).

REFERENCES

NIST Definition of Cloud Computing v15. [Online] http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

GFD.150 Using Clouds to Provide Grids Higher-Levels of Abstraction and Explicit Support for Usage Modes. [Online] http://www.ogf.org/documents/ GFD.150.pdf

GEANT Project. http://www.geant.net/pages/ home.aspx

Generalised Architecture for Dynamic Infrastructure Services (GEYSERS Project)-http://www.geysers.eu/

Enabling Grid from E-sciencE (EGEE Project). [Online]. http://www.eu-egee.org/

Phosphorus Project. [Online]. Available: http:// www.ist-phosphorus.eu/

Demchenko, Y., M. Cristea, C. de Laat, E. Haleplidis, Authorisation Infrastructure for On-Demand Grid and Network Resource Provisioning, *Proc. 3rd Intl ICST Conference on Networks for Grid Applications (GridNets 2009),* Athens, Greece, 8-9 Sept 2009. ISBN: 978-963-9799-63-9

TMF Service Delivery Framework. http:// www.tmforum.org/servicedeliveryframework/4664/ home.html

Deliverable DJ3.3.1: Composable Network Services use cases. GEANT3 Project Deliverable. January 11, 2010. http://www.geant.net/Media_Centre/Media_ Library/Media%20Library/GN3-09-198-DJ3_3_1_ Composable_Network_Services_use_cases.pdf

OASIS Reference Architecture Foundation for Service Oriented Architecture 1.0, Committee Draft 2, Oct. 14, 2009. http://docs.oasis-open.org/soa-rm/soa-ra/v1.0/ soa-ra-cd-02.pdf

Chappell, D., "Enterprise service bus", O'Reilly, June 2004. 247 pp.

Pautasso, C., O. Zimmermann, F. Leymann, "RESTful Web Services vs. Big Web Services: Making the Right Architectural Decision", *17th International World Wide Web Conference (WWW2008)*, Beijing, China.

Fuse ESB - OSGi based ESB. - http://fusesource.com/ products/enterprise-servicemix/#documentation

Apache ServiceMix an Open Source ESB. - http:// servicemix.apache.org/home.html

OSGi Service Platform Release 4, Version 4.2. - http://www.osgi.org/Download/Release4V42

Spring Security. Reference Documentation. http:// static.springsource.org/spring-security/site/docs/3.1.x/ reference/springsecurity-single.html

Demchenko, Y., D. R. Lopez, J. A. Garcia Espin, C. de Laat, "Security Services Lifecycle Management in On-Demand Infrastructure Services Provisioning", International Workshop on Cloud Privacy, Security, Risk and Trust *(CPSRT 2010), 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom2010)*, 30 November - 3 December 2010, Indianapolis, USA.

On-Demand Infrastructure Services Provisioning BoF (ISOD BoF), OGF30 meeting, 25 October 2010, Brussels. http://www.gridforum.org/gf/event_ schedule/index.php?id=1961