Collaboration and the Knowledge Economy: Issues, Applications, Case Studies
Paul Cunningham and Miriam Cunningham (Eds)
IOS Press, 2008 Amsterdam
ISBN 978–1–58603–924-0

Dynamic SLA and Trust for Next Generation Business Models in Grid

Francesco D'ANDRIA[1], Sandra JIMÉNEZ[1], Mélanie BIETTE[1],
David BROSSARD[2], Lenni MADSED[2] Francesco Orciuoli[3]
[1]*Atos Origin - Research and Innovation, Diagonal 200, Barcelona, 08011, Spain*
Tel: +34 93 4861818, Fax: +34 93 4860766,
Email: {francesco.dandria, sandra.jimenez, melanie.biette}@atosorigin.com
[2]*British Telecommunications plc, Adastral Park, Martlesham Heath, IP5 3RE, United Kingdom Tel: +44 1473 621 560, Email: {david.brossard, lenni.madsen}@bt.com*
[3]*DIIMA, Univeristy of Salerno, via ponte don Melillo, Fisciano, 84084, Fisciano (SA)*
Email: orciuoli@diima.unisa.it

Abstract: This paper looks at the changing nature and relationship of trust and Service Level Agreements (SLA) in emerging business environments based on Grid technologies. This paper presents a new approach for managing SLA contracts in new highly dynamic scenarios tackled in BEinGRID's business experiment on virtual hosting environments (VHE) applied to online gaming. A new negotiation mechanism based on the advertisement/discovery of business services is presented. Finally, we present the application of our work to a concrete business scenario related to the on-line game application provision, offering as well an overview of the main business benefits assessed during the evaluation of the components.

1. Introduction

BE09, one of the Business Experiments of the BEinGRID project, aims to design a new technological and advanced Virtual Hosting Environment (VHE) built on top of a Service Oriented Infrastructure (SOI) for Application Service Provider (ASP).

This Grid-based SOA environment will offer an advanced method of achieving cross-enterprise integration of application services and the virtualization of the computational environment where these services are hosted and executed. Whilst leveraging the middleware to create the virtual environment, BE09 assures the quality of the services and the possibility to create dynamic federations of users, services and servers, offering as well a high level of security and trust. Applied to our case of study of on-line games, it will decrease costs in the development of gaming infrastructure.

Grid technologies have a high level of awareness and are well-known in the scientific and academic fields but there still exist obstacles that must be overcome before Grid commercialisation, such as the concepts of trust and quality of service assurance. Service Level Agreements (SLAs) have been introduced in Grid Computing to overcome this weakness by introducing a contract between the provider and the customer.

Traditionally in computing, notions surrounding trust and Service Level Agreements have been interlinked with service provisioning but have been kept almost separate in notion. SLA has commonly been associated with contracts of acceptable use and service provision that often remain stable throughout the lifetime of an agreement, whereas trust has often been defined on an individual basis. An example of this is the selection of a typical service provision in the computing industry of web hosting. Here customers may make their choice based on their personal assessment with respect to several criteria, such as a personal recommendation of the supplier. Once this selection has been made the

Copyright © 2008 The Authors

customer may enter an agreement with the provider, determining some user requirements such as acceptable use and service provision issues, as server uptime, response time, etc.

To adapt the traditional notions of trust and SLA to the new development of Service Oriented Architectures (SOA), and in particular to the innovation within the area of service composition led by the Grid community, is a significant challenge which is up-to-date with the market trends. A survey of Gartner added that by 2010, 80 percent of application software revenue growth, including licenses and subscription fees, will come from products based on SOA and that 25 percent of application demands will be delivered through real-time infrastructure (RTI)/IT utility [1] [8]. Furthermore William Fellows, one of the founders of The 451 Group, added that utility models and SOAs within enterprises will be the key market with companies testing outsourced Grid services [2]. Here the automated selection of services has to be represented in more solid terms and trust therefore has to be tangible. Additionally, as multiple services are orchestrated and workflows consisting of multiple services are composed, SLAs often become multidimensional and susceptible to rapid change depending on various environmental factors.

This ability to combine both trust and SLA is a central focus for business applications for Grid-based SOA environments. SLA will foster Grid adoption into the market by improving commercial relationships between end-users and providers. As this paper will illustrate, this fact challenges current technological provision and human understanding of both trust and SLA within a Grid-based SOA environment. The rest of this paper is organized as follows: Section 2 shows the main goals to be achieved, the identification of the problems and the intended results. Section 3 analyses some of the design decisions made and technologies used in implementing our framework. Section 4 describes the technical aspects of the VHE, while the SLA layer description and technical details are detailed in section 5. Section 6 introduces the business case used to validate the of our VHE implementation along with a description of the on-line game application scenario. Finally, Section 7 provides a short summary and some concluding remarks.

2. Objectives

The goal of this paper is to show a way of implementing trust based on the concept of VHE and secure federation, as well as the implementation of an SLA management layer, enabling automatic resource selection, in a Grid-based SOA environment. In order to support the environment we envision many issues need to be addressed, which are detailed in this section. As a case study, the BE09 examines an online gaming platform where has been demonstrated a convergence between Web Services (WS-* approach) and service-oriented Grid computing architecture (the OGSA approach). BE09 offers an emerging market-led proposition targeting higher eBusiness and IT services to enterprises to deliver high quality infrastructure and systems that support the enterprise's move to SOA.

In this business context (see section 6 for further details), the management of an SLA with agreed QoS parameters between a Service Customer and some Service Provider(s) is an essential requirement. The lack of appropriate mechanisms for SLA negotiation, monitoring and evaluation is an important barrier to the uptake of Grids by companies and industry in distributed e-business environments. This also highlights the real risk in the decision and system management since an SLA violation results in paying a penalty fee. In this paper it is described the problem of managing SLA e-contracts in a Grid scenario and we outline a design that considers aspects such as contract negotiation, monitoring, evaluation and application of suitable policies when QoS is not guaranteed. In particular, we will focus on those that are more related to business requirements, including:

- Automatic resources "negotiation" through an "SLA-based" advertisement and discovery mechanism.

Copyright © 2008 The Authors

- Monitoring of agreements, considering network related Quality of Service (QoS) and the network availability itself as a relevant component of the value chain for service provision.
- Platform independent agreement evaluation against the Service Level Objectives (SLO) inside the contract at run-time

As for trust, it can be established through a federation layer discussed in [12] and [13] mainly a security token service (STS) and the governance layer which controls it. Each partner wanting to take part in the VHE exchanges business cards – XML documents that contain a given partner's public key – via their federation establishment interface. Once each partner's STS is thus configured, the circle of trust becomes effectively active. Please refer to [12] and [13] for a more in-depth analysis of the security and federation principles.

3. Methodology Used

The approach mainly follows an XML-based consolidated specification WSLA [4], although we have also considered aspects of WS-Agreement specification [5] in the definition of the SLA template and SLA contracts associated to a service. We offer in addition a new negotiation model based on the advertisement and discovery of an SLA contracts supported by a UDDI[i] technology on top of a hierarchical registry architecture.
The SLA Management subsystem includes SLA monitoring and enforcement mechanisms, taking a supervisory role to verify that the negotiated contract conditions of all running services are met and to take corrective actions when a violation of the SLA contract is detected. Following the WS-Notification specification [9], the different components are able to send alerts about any abnormal situation to the interested modules. We present also, an approach to map the application QoS requirements in infrastructure QoS parameters.

From a security perspective, the approach follows a web-service based angle where communication takes place via SOAP and an ensemble of security services (mainly the Policy Enforcement Point PEP, Security Token Service STS, and Policy Decision Point PDP) provides authentication, authorization and XML threat protection. The combination and configuration of these services is largely discussed in [12] and [13]. These services brought together provide what can be called a B2B gateway, a thin governable layer that provides highly dynamic & flexible security mechanisms for web service security.

4. Virtual Hosting Environment

The VHE offers capabilities for establishing and maintaining federation and VO, capability for service virtualization, for distributed access management, for location and discovery, and also some of the hardware resources, offering a secured environment.

The VHE is layered into different components as follows:

1. The VO Toolkit
2. The Governance Gateway
3. The Core Host Environment

Each layer addresses different concerns. Firstly the **VO Toolkit** addresses CDL [14]-based collaboration establishment, selection of partners and roles, and federation creation. The VO Toolkit comes in three versions:

- The *host edition* which stores information about partners, their business cards, the services they can offer, existing VOs, and virtualized resources. It acts as a UDDI or can be plugged to a third party UDDI.
- The *initiator edition* which is responsible for the creation of collaborations, the sending of email invitation to members, the selection and prioritization of members for each role to be fulfilled in a given collaboration; it sends each member a list of

Copyright © 2008 The Authors

participating members' business cards as well as the unique identifier of the VO currently being created.

- The *member editions* (one per participating member) which are responsible for handling requests from the initiator, accepting / refusing invitations, and liaising with the underlying member infrastructure in particular the gateway governance layer which contains a federation manager. When a member accepts an invitation and the initiator edition operates the VO, each member edition will then talk to its governance gateway to configure the corresponding infrastructure and in particular the security component responsible for the establishment of trust, i.e. the security token service (STS); it relays the information received from the initiator to the federation manager.

The **Governance Gateway** handles two main aspects and therefore exposes two main interfaces:

- The *gateway federation manager* which is responsible for the initial security infrastructure configuration. Based on contextual data, the federation manager will load an infrastructure profile which determines which security services to use. If there is an STS (which is needed in federation establishment), the federation manager will configure it. The federation manager uses the VO ID and business cards received from its member edition to configure the STS in a two step process: firstly create the actual federation within the STS and secondly push the business cards to the given federation in the STS. A business card is an XML document containing information about a partner and that partner's public key.
- The *gateway virtualization service* focuses on the instantiation of services and contextualization of the instance's exposure. It also manages the core hosting environment where the Service Instantiator lies. The Service Instantiator is responsible for the host selection and actual instantiation of the service whereas the virtualization service is responsible for the contextualization and exposure of the instance with a given security context and within the pre-established federation.

Finally, the Core Host Environment handles the hosting and running of the actual applications. It runs the Service Instantiator which creates contextualized instances of the hosted applications. These instances are then given to the virtualization service for further contextualization and for an associated infrastructure configuration that provides security and QoS related services. The core host environment allows a lower level of control focusing more on memory and CPU usage, handling possible load balancing between different hosts.

4.1. The Enforcement of Trust

The trust established by means of the VHE, VO Toolkit and eventually STS can now be used to safely and securely exchange messages between the different partners of the trust realm. What is interesting to note is that the trust definition may be asymmetric: if there are four partners (A, B, C, D), one partner (A) may know the other three (B, C, D) but the last partner may only know A, hence ignoring the other partners.

Security is enforced by means of a policy enforcement point (PEP) which intercepts outgoing and incoming messages. The PEP analyzes the contextual information contained in the SOAP message, e.g. the addressing headers, to locate a security policy to enforce. Once found, the PEP executes the said policy. The latter contains an order to invoke the STS to either:

1. issue a SAML token in the case of an outgoing message,
2. or validate a SAML token in the case of an incoming message.

The SAML token is used as the VO wide identity of the requester. An STS can only issue a token if the requestor is defined in the STS and if the requestor wants to use the

Copyright © 2008 The Authors

service of a partner defined in the same STS (by means of its business card as aforementioned). In addition, if there any claims / attributes for the given requestor stored in the STS, the latter can include them in the SAML token. The PEP receives the SAML token from the STS as well as the proof-of-possession key which contains the symmetric key that will be used to encrypt the PEP-PEP message exchange. The SAML token is encrypted using the public key of the targeted partner (which is stored at the STS within the given partner's business card) and added to the SOAP message. The PEP encrypts the outgoing message and sends it onwards. Confidentiality & privacy is therefore ensured.

On the server side, the incoming message is intercepted. The receiving PEP sends the encrypted SAML token to the STS which validates it and returns the proof of possession key to the PEP which then uses it to decrypt the message body.

To further enhance the security model, a policy decision point can be added to provide attribute-based access control where the attributes are managed by and stored at the STS.

5. SLA as a VHE Added Value

Our SLA management strategy considers two well differentiated phases: the advertisement and discovery of the couple Business Service / SLA Contract and the monitoring and evaluation of its fulfilment at run-time. Here we show the process following the BE09 scenario of on-line games described in section 6.

5.1. SLA as Capability to Advertise and Discover Business Service

VHE offers capabilities to advertise and discover business service providing information about QoS through the locator and discovery subsystem (L&D). L&D subsystem has a four-layered hierarchical architecture with registries at three different levels:

- At the bottom level (Service Level) there are the services that have to be published.
- At Host level, there is the Host Instances Registry (HIR), which runs on all hosts. It is a registry that using the WS-Notification mechanism collects information about the deployed business services and forwards it to the Gateway Instances Registry (GIR).
- At the Gateway level, on the Gateway Machines, runs the GIR, managing the information of the business services inside a Hosting Environment (HE).
- Finally, at Virtual Organization level, we have the Locator Service, as the top high-level registry for all the VO's available business services.

When a new business service is deployed on a host, it is published together with an associated SLA pre-contract in a process that spreads its information across the three different levels of the L&D subsystem (architecture showed in the Figure 1).

Copyright © 2008 The Authors

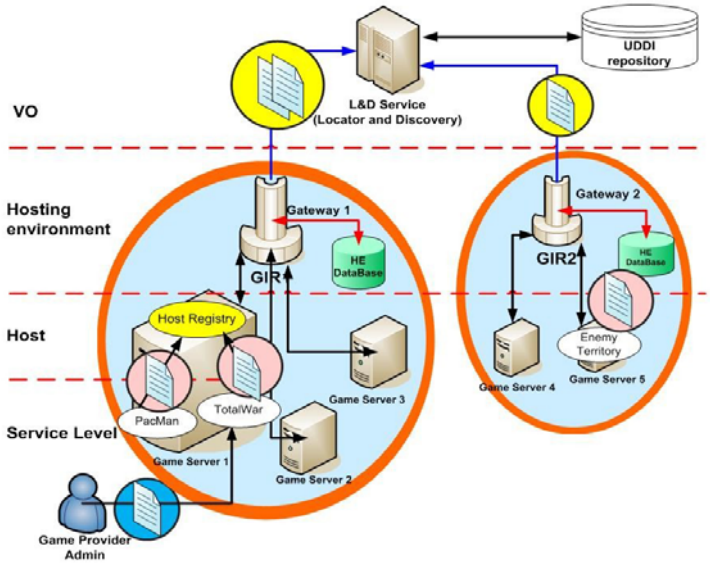

Figure 1: Hierarchical architecture of the L&D Subsystem

In our vision the L&D subsystem extends the classical UDDI directory functionalities in two aspects:

- Allows publishing business services against the directory through an automatic mechanism.
- Allows classifying business on the basis of metadata that describes QoS information contained in the associated SLA pre-contract.

In our scenario when a Game Provider deploys a new game, he also publishes an SLA Template (or SLA pre-contract) associated to that game with some specific QoS that should be guaranteed (see the contract in the pink circle in Figure 1). These QoS parameters cover infrastructure, performance and network parameters, such as CPU use, latency or memory, which will be called low level (LL) parameters from now on. Afterwards, the Game Provider defines an XML-based mapping policy, which, according to a static equivalence established by the Game Platform, maps the LL QoS parameters included in the SLA Template into high level (HL), human understandable, QoS parameters.

```
<?xml version="1.0" ?>
- <mapping_description xmlns="http://www.beingrid.org">
  - <ServiceDescription name="EnemyTerritory">
    - <serviceMap GraphicResolution="Medium" AvailableResources="Medium">
      - <LowLevelSLAParameters>
        - <SLAParameter>
            <Name>CPUUse</Name>
            <Value>70</Value>
            <Predicate>Less</Predicate>
          </SLAParameter>
        - <SLAParameter>
            <Name>Latency</Name>
            <Value>300</Value>
            <Predicate>Less</Predicate>
          </SLAParameter>
        - <SLAParameter>
            <Name>Memory</Name>
            <Value>3000000000</Value>
            <Predicate>Greater</Predicate>
          </SLAParameter>
        </LowLevelSLAParameters>
      </serviceMap>
    </ServiceDescription>
</mapping_description>
```

Figure 2: XML Based Mapping Policy

Copyright © 2008 The Authors

At search time when the on-line game (OLG) clients (Gamers) want to look for a service (game), the "human understandable", HL QoS parameters are specified as search criteria: e.g. Graphic Resolution or Available Resources. Using the mapping capability provided by the VHE the L&D, the Service Directory is queried for potential Service Providers that are able to offer the most suitable service to the client as showed in the Figure 3.

Finally the business service (a match for a given game, in BE09's scenario) and its associated SLA Contract is delivered to the Gamer.

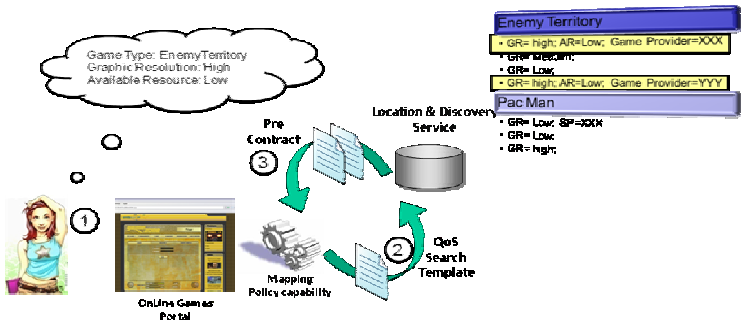

Figure 3: Mapping Policy Capability at search time

5.2. SLA as Capability to Evaluate and Monitor Business Contract

After the business service has been delivered, it is necessary to ensure that the contractual terms are respected. It is done though the monitoring and evaluation subsystem (M&E).

M&E subsystem is logically divided into three main blocks:

- **Application-specific monitoring**: offers the ability to retrieve at run-time information about the users participating in the match, and other general information about the match like its lifecycle, number of users playing the match and some game statistics;
- **Infrastructure monitoring:** offers the ability to monitor resources virtualized as grid services. In BE09 it will be possible to monitor parameters like the CPU cycle and the memory consumed by the match (service instance) at runtime;
- **Evaluation layer:** offers the ability to collect (through the two above mentioned modules) the monitored values in order to verify whether the measurements are within the thresholds defined in the SLA contract assigned to every player. Whenever the execution of a match does not satisfy these SLA conditions, the module will launch a notification event (using a WS-Notification mechanism) about this anomalous behavior.

Figure 4 gives a high level view of our monitoring and evaluation idea.

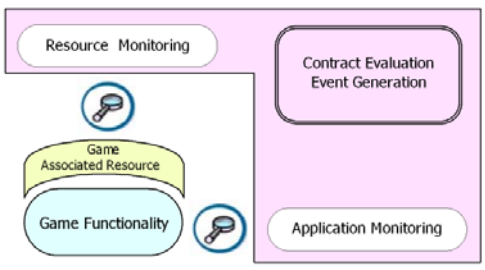

Figure 4: Monitoring and SLA Contract Evaluation at Run-Time Phase

6. Business Case Description

Overall, through BE09, it is intended to validate the technical and economical viability of a VHE implementation that builds on commercial middleware components and SOA.

Copyright © 2008 The Authors

Validation will be achieved by implementing an On-Line Gaming (OLG) platform (provided by Andago) on the top of a VHE specialisation for on-line collaborative gaming.

VHE offers a strategic competitive advantage to IT Services and Communications companies in face of SOA platform vendors and IT integration companies. Indeed once the VHE infrastructure is in place, the companies increase their ability to deliver cost effectively, exploit short concept-to-market timescales and also benefit from economies of scale. We expect that our decision to leverage SOA and Grid concepts realised on a commercial-off-the-shelf and Open standards Web Services layer will push the wide adoption of the Web Services technology[ii]. The validation scenario used in BE09 is about enabling the provision of high-end on-line collaborative gaming environments. On-line gaming is fast conquering this new market and more than 30% of gaming development companies is currently focusing on developing high performance platforms for the new on-line trend. VHE provides the organisations new means to configure new services faster and to improve efficiency, as their fast business environment changes.

6.1. On-Line Game Application Scenario

As said before, Andago is an application provider which offers an on-line collaborative game platform, and wants to cut the cost of their provision by outsourcing hardware resources and non-operational services. By doing this, Andago will also benefit from a more flexible platform that would allow them to load-balance servers they use based on user demand, and also to choose over time which providers they want to contract for using their *security, SLA* and other non-operational services.

In order to do so, Andago contacts a VHE operator (located in UK). Through its VHE, the VHE operator offers access to all the capabilities that Andago needs to achieve the above. Some of these capabilities are provided by the VHE operator itself, other by different infrastructure providers. Namely, the VHE operator offers capabilities for establishing and maintaining federation and VO, capability for service virtualization, for distributed access management, and also some of the HW resources hosted in the UK. In addition to this, there are two infrastructure providers (located in Spain and in Italy) that offer capabilities for SLA monitoring and evaluation, for location and discovery and for advanced VO management. Initially, Andago wants to offer on-demand gaming platform to its user base in Spain. Then as Andago's customer base expands they decide to expand their VO, in order to introduce more servers, and some additional security services. Since Andago is establishing its presence in the UK, they need to find HW in that area in order to ensure they meet the latency requirements to their UK customers. Upon search at L&D subsystem, Andago finds suitable resources offered by the VHE operator located in UK and amends their contract with the VHE operator to include more game servers. This is followed by fixing SLA agreements between Andago and the VHE operator and the application deployment. In order to enhance business intelligence of their network centric application, Andago chooses to use advanced VO management services offered to the VHE by the infrastructure provider located in Italy and distributed access management services offered by the VHE operator. These are network-hosted services that allow Andago to define their own access control policies and manage participation to different federations.

Finally, once the environment is setup, the on-lime game platform can start provision and delivery of the gaming application to its end users. In brief, when an end user requests the service, this is delivered by executing the deployed application. First, using negotiation mechanism of the SLA subsystem, Andago identifies the appropriate resources (servers) that can deliver needed QoS. Then, it uses service virtualization capability that will manage creation of the application instance on the selected HW – this includes service access configurations, any amendments to the policies agreed between the parties, configuration of SLA monitoring and evaluation capabilities, and the billing. In order to correlate the above

Copyright © 2008 The Authors

activities under the common interaction context (i.e. execution of particular application instance for a given user / set of users), VHE federation management capability is used by virtualization service on behalf of Andago. Once application execution is finished, data logging is performed, and the above configurations are released.

7. Conclusions

This paper has introduced an approach for managing SLA contracts and trust in the new, dynamic Grid scenarios tackled in BE09 project. The problems of controlling context, negotiating and controlling QoS as well as undertaking corrective actions are outlined, using an unambiguous and clear specification of Service Level Agreements, which validation and monitoring can be performed by a SLA management solution integrated with the VHE. From the on-line gaming perspective, the game developers meet a higher security level, enforcing trust in an objective way, at the moment of inserting the game in a game platform within a virtual environment. The SLA management and SLA monitoring help them to save time and money, and to insert efficiency inside the platform and then the service. Concerning the ASPs, the developed infrastructure provides them a dynamic orientation of services. This approach will bring security, management and monitoring of resources to an on-line gaming platform. Moreover the ASPs have the advantage of the VHE where federation gives standardisation to the online gaming sector which decreases the fragmentation of the European market and bring a bigger users mass. Finally the players will see a high improvement in the quality of service through the higher security, flexibility and performance shown by the platform. In terms of QoS parameters, we deal with Grid and network metrics by means of a close interaction of both layers. Grid technologies provide numerous advantages such as optimisation of resources use, higher flexibility and new business models in open and distributed e-business context for industry. The paper indicates the need of working in more flexible but simplified contract negotiation mechanisms in these environments, as well as the need of working of collocation protocols for negotiation.

References

[1] Cearley, David W – Fenn, Jackie – Plummer, Daryl C. (2005), "Gartner's Positions on the Five Hottest IT Topics and Trends in 2005", Gartner

[2] Grid Today, Fellows,W, 2006, http://news.taborcommunications.com/msgget.jsp?mid=545661&xsl=story.xsl

[3] BEinGIRD - Business Experiments in Grid http://www.beingrid.com/ http://www.gridipedia.eu/

[4] Web Service Level Agreement (WSLA) http://www.research.ibm.com/wsla/

[5] WS-Agreement Specification www.ogf.org/documents/GFD.107.pdf

[6] Akogrimo Access to Knowledge through the Grid in a Mobile World http://www.mobilegrids.org/

[7] GrASP http://www.eu-grasp.net/

[8] Gartner Dataquest, May 2005

[9] Web Services Notification http://www-128.ibm.com/developerworks/library/specification/ws-notification

[10] UDDI specification http://www.oasis-open.org/committees/uddi-spec/doc/tcspecs.htm

[11] TrustCoM http://www.eu-trustcom.com

[12] Brossard, David et al "Common Capabilities for Trust & Security in Service Oriented Infrastructures" in the eChallenges 2008 proceedings

[13] Angelo Gaeta, F. Orciuoli, N. Capuano, D. Brossard, T. Dimitrakos. A Service Oriented Architecture to support the federation lifecycle management in a secure B2B environment. In proceedings of e2008.

[14] CDL "choreography description language" http://www.w3.org/TR/ws-cdl-10/.

[i] Universal Description, Discovery and Integration (UDDI) is a platform-independent, XML-based registry for businesses worldwide to list themselves on the Internet [10]

[ii] See also Gartner Research statement [8] about Web services standards and technology adoption in 2006 in the "target market" section

Copyright © 2008 The Authors