Authorisation Infrastructure for On-Demand Grid and Network Resource Provisioning

Yuri Demchenko[*], Mihai Cristea[*], Cees de Laat[*], Evangelos Haleplidis[#],

* University of Amsterdam, System and Network Engineering Group
{demch, cristea, delaat}@science.uva.nl
University of Patras
ehalep@gmail.com

Abstract. The paper presents the Authorisation (AuthZ) infrastructure for combined multidomain on-demand Grid and network resource provisioning which we refer to as the Complex Resource Provisioning (CRP). The proposed CRP model provides a common abstraction of the resource provisioning process and is used as a basis for defining the major AuthZ mechanisms and components that extend the generic AAA AuthZ framework to support CRP (GAAA-CRP), in particular using XML-based AuthZ tickets and tokens to support access control and signalling during different CRP stages. The proposed GAAA-CRP framework is implemented as the GAAA Toolkit pluggable library and allows integration with the Grid and network service and control plane middleware. The proposed authorisation infrastructure allows using in-band binary tokens to extend network access control granularity to data plane and support binding applications to dataflows. The paper discusses the use of the ForCES network management model to achieve interoperability with the network control plane and define the GAAA-NRP interfaces to network control plane. This research was conducted as a part of the EU Phosphorus project.

Keywords: Complex Resource Provisioning (CRP), Multidomain Network Resource Provisioning, AAA Authorisation Framework, Authorisation session, Token Based Networking (TBN), ForCES.

1. Introduction

High performance distributed Grid applications that deal with high volume of processing and visualisation data require dedicated high-speed network infrastructure provisioned on-demand. Currently large Grid projects and Cloud Computing providers use their own dedicated network infrastructure that can handle the required data throughput but typically are over-provisioned. Any network upgrade or reconfiguration still requires human interaction to change or negotiate a new Service Level Agreement and involve network engineers to configure the network. Need for combined computer-network resources provisioning and optimisation will increase with emerging Cloud Computing that has stronger commercial focus than Grid computing.

Most of Grid usage scenarios can benefit from combined Grid and network resource provisioning that besides improving performance can address such issues as (application centric) manageability, consistency of the security services and currently becoming important energy efficiency. The combined Grid/computer and network resource provisioning requires that a number of services and network resources controlling systems interoperate at different stages of the whole provisioning process. However in current practice different systems and provisioning stages are not connected in one workflow and can not keep provisioning and security context, what is resulted in a lot of manual work and many decision points that require human involvement.

In this paper we extend the proposed earlier the Network Resource Provisioning (NRP) model [1] to the more general Complex Resource Provisioning (CRP) model that provides a common framework for combined Grid/computer resources and network infrastructure provisioning and allows for integrating existing systems/services and technologies into common provisioning workflow that include such stages as reservation, deployment, access, and additionally decommissioning, that require different security and access control services and mechanisms.

Security and authorisation services to support CRP should have high granularity, capable of dynamic invocation at different networking layers, and support all stages of the provisioned resources lifecycle. The proposed GAAA-CRP infrastructure and services are designed in such a way that they can be used at all networking layers (dataflow plane, control plane and service plane) and allow easy integration with Grid middleware and application layer security. For this purpose, special mechanisms are proposed to manage inter-layer and inter-domain security context.

The paper is organized as follows. Section 2 describes the proposed general CRP model that separates resource reservation, resource deployment, and resource access stages. This section also summarises common requirements to AuthZ services/infrastructure to support different provisioning and AuthZ scenarios in distributed dynamic environment. Section 3 discusses the use of the AuthZ tickets and tokens for signalling and access control in multidomain CRP. Section 4 provides suggestions how the ForCES and Token Based Networking (TBN) can be used to achieve higher granularity of the control of the provisioned network paths. Section 5 briefly presents our ongoing implementation, and finally section 6 provides a short summary and suggests future developments.

2. CRP model and GAAA-CRP Authorisation infrastructure

The typical on-demand resource provisioning process includes four major stages, as follows: (1) resource reservation; (2) deployment (or activation); (3) resource access/consumption, and additionally; (4) resource de-commissioning after it was used. In its own turn, the reservation stage (1) typically includes three basic steps: resource lookup; complex resource composition (including alternatives), and reservation of individual resources.

The reservation stage may require the execution of complex procedures that may also request individual resources authorisation. This process can be controlled by an advance reservation system [2] or a meta-scheduling system [3]; it is driven by the

provisioning workflow and may also include Service Level Agreement (SLA) negotiation [4]. At the deployment stage, the reserved resources are bound to a reservation ID, which we refer to as the Global Reservation Identifier (GRI). The decommissioning stage is considered as an important stage in the whole resource provisioning workflow from the provider point of view and should include such important actions as global provisioning/access session termination and user/process logout, log information sealing, accounting and billing.

The rationale behind defining different CRP workflow stages is that they may require and can use different security models for policy enforcement, trust and security context management, but may need to use common dynamic security context.

In the discussed CRP model we suggest that the resources are organised in domains that are defined (as associations of entities) by a common policy or a single administration, with common namespaces and semantics, shared trust, etc. In this case, the domain related security context may include:

- static security context such as domain based policy authority reference, trust anchors, all bound by the domain ID and/or domain trust anchor [19];
- dynamic or session related security context bound to the GRI and optionally to a Local Reservation ID (LRI).

In general, domains can be hierarchical, flat or have irregular topology, but all these cases require the same basic functionality from the access control infrastructure to manage domain and session related security context. In the remainder of the paper we will refer to the typical use case of the network domains that are connected as chain (sequentially) providing connectivity between a user and an application.

Figure 1 illustrates major interacting components in the multi-domain CRP using example of provisioning multidomain network connectivity between a User and a Destination resource or application. Each networking domain is presented as

- Network Elements (NE) (related to the network Data plane);
- Network Resource Provisioning Systems (NRPS) acting as a Domain Controller (DC) (typically related to the Control plane);
- Inter-Domain Controller (IDC) managing cross-domain infrastructure operation, often referred to as Network Service Plane (NSP).

Access to the resource or service is controlled by the DC or NRPS and protected by the generic Authentication, Authorisation, Accounting (AAA) service that enforces a resource access control policy. The following functional elements comprise the proposed authorisation infrastructure for CRP which we will refer to as GAAA-CRP:

- Policy Enforcement Point (PEP), Policy Decision Point (PDP), and Policy Authority Point (PAP) as major functional components of the Generic AAA AuthZ infrastructure (GAAA-AuthZ) [5].
- Token Validation Services (TVS) that allow efficient authorisation decision enforcement when accessing reserved resources.

Depending on the basic GAAA-AuthZ sequence (push, pull or agent) [4], the requestor can send a resource access request to the resource (which in our case is represented by NRPS) or an AuthZ decision request to the designated AAA server which in this case will act as a PDP. The PDP identifies the applicable policy or policy set and retrieves them from the PAP, collects the required context information, evaluates the request against the policy, and makes the decision whether to grant access or not.

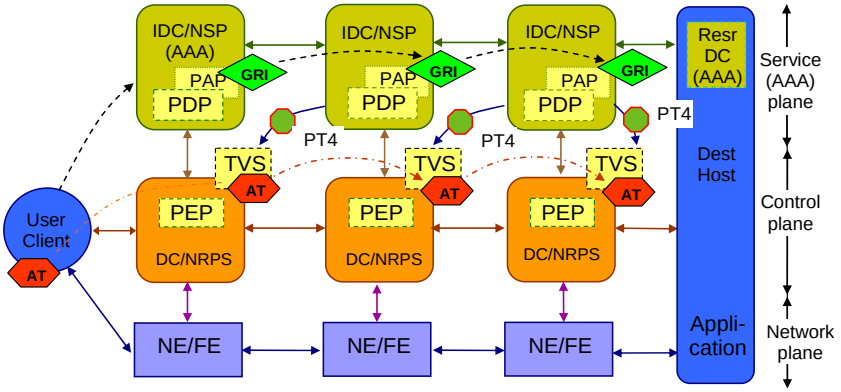

Figure 1. Components involved in multidomain network resource provisioning.
CRP stages reservation, deployment and access are presented by the flows correspondingly GRI (forward from the user to the resource), pilot tokens PT4 (backward), and access tokens AT (forward).

Depending on the used authorisation and attribute management models, some attributes for the policy evaluation can be either provided in the request or collected by the PDP itself. It is essential in the Grid/Web services based service oriented environment that AuthN credentials or assertions are presented as a security context in the AuthZ decision request and are evaluated before sending request to PDP.

Based on a positive AuthZ decision (in one domain) the AuthZ ticket (AuthzTicket), containing AuthZ decision and context, can be generated by the PDP or PEP and communicated to the next domain where it can be processed as a security context for the policy evaluation in that domain.

In order to get access to the reserved resources (at the access stage) the requestor needs to present the reservation credentials that can be in a form of an AuthZ ticket (AuthzTicket) or an AuthZ token (AuthzToken) which will be evaluated by the PEP with support of TVS for ticket or token evaluation, to grant access to the reserved network elements or the resource. In more complex provisioning scenarios the TVS infrastructure can additionally support an interdomain trust management infrastructure for off-band token and token key distribution between domains that typically takes place at the deployment stage when access credentials or tokens are bound to the confirmed GRI by means of shared or dynamically created interdomain trust infrastructure. Token and token key generation and validation model can use either shared secret or PKI based trust model.

The TVS as a special GAAA-CRP component to support token-based signalling and policy enforcement mechanism is briefly described below.

It is an important convention for the consistent CRP operation that GRI is created at the beginning and sent to all polled/requested domains when running (advance) reservation process. Then in case of a confirmed reservation, the DC/NRPS will store the GRI and bind it to the committed resources. In addition, a domain can also associate internally the GRI with the Local Reservation Identifier (LRI). The proposed TVS and token management model allows for hierarchical and chained GRI-LRI generation and validation.

Correspondingly we define the following sessions in the overall CRP process: provisioning session that includes all stages; reservation session, and access session. All of them should share the same GRI and AuthZ context.

The proposed GAAA-CRP infrastructure includes the following access control mechanisms and components that extend the generic GAAA-AuthZ model described in [4] with the specific functionality for on-demand CRP, in particular:

- AuthZ session management to support complex AuthZ decision and multiple resources access, including multiple resources belonging to different administrative and security domains.
- AuthZ tickets with extended functionality to support AuthZ session management, delegation and obligated policy decisions.
- Access and pilot tokens used for interdomain reservation process management access control as part of the policy enforcement mechanisms that can be used in the control plane and in-band.
- Policy obligations to support usable/accountable resource access/usage and additionally global and local user account mapping widely used in Grid based applications and supercomputing.

The solutions proposed in the GAAA-CRP framework are based on using such structural components and solutions as the Token Validation Service, the Obligation Handling Reference Model (OHRM) [6], and the XACML attributes and policy profile for multidomain CRP that can combine earlier defined XACML-Grid and XACML-NRP profiles [7, 8].

3. Using Tickets and Tokens for Signalling and Access Control and Token Validation Service

In the proposed AuthZ architecture the tokens are used for access control and signalling at different CRP stages and considered as a flexible and powerful mechanism for communicating and signalling security context between domains. Tokens are abstract constructs/entities that refer to the related session context stored in the domains or by services. The GAAA-CRP uses three major types of the provisioning or AuthZ session credentials:

- AuthZ tickets that allow expressing and communicating the full/extended AuthZ session context and in this way could be used as access credentials.
- Access tokens that are used as AuthZ/access session credentials and refer to the stored reservation context.
- Pilot tokens that provide flexible functionality for managing the AuthZ session and the whole provisioning process.

Access tokens are used in rather traditional manner and described in details in [9]. Pilot token can be fully integrated into the existing network Control Plane interdomain protocols such as RSVP and GMPLS and in particular can be used as a container for AuthZ ticket in interdomain communication.

Although the tokens share a common data-model, they are different in the operational model and in the way they are generated and processed. The following elements and attributes are common for all tokens: GRI, DomainID, TokenID,

TokenValue, - that allow unique token's identification and validation. More details about the token datamodel and processing can be found in the recent authors' paper [10].

In the proposed GAAA-CRP the token handling functionality is outsourced to the Token Validation Service (TVS) that supports different token handling models and store token and session related context.

Basic TVS functionality allows checking if a service/resource requesting subject or other entity, that possess current token, has permission to access/use a resource based on advance reservation to which this token refers. During its operation the TVS checks if a presented token has reference to a previously reserved resource and a request information conforms to a reservation conditions.

In a basic scenario, the TVS operates locally and checks a local reservation table directly or indirectly using a reservation ID (e.g. in a form of GRI). It is also suggested that in a multi-domain scenario each domain may maintain its Local Reservation ID (LRI) and provides its mapping to the GRI. In more advanced scenario the TVS should allow creation of a TVS infrastructure to support tokens and token related keys distribution to support dynamic resource, users or providers federations.

For the purpose of authenticating token origin, the pilot token value is calculated of the concatenated strings DomainId, GRI, and TokenId. This approach provides a simple protection mechanism against pilot token duplication in the framework of the same reservation/authorisation session.

The following expressions are used to calculate the TokenValue for the access token and pilot token:

```
TokenValue = HMAC(concat(DomainId, GRI, TokenId), TokenKey)
```

When using pilot tokens for signalling during interdomain resource reservation, TVS can combine token validation from the previous domain and generation of a new token with the local domain attributes and credentials.

4. Fine-grained Policy Enforcement at Networking Layer

4.1. In-band policy enforcement with TBN

The proposed GAAA-CRP architecture is easily integrated with the Token Based Networking (TBN) technology being developed at University of Amsterdam [11] to achieve in-band policy enforcement at dataflow layer. The TBN allows binding dataflows to users or applications by labeling application specific traffic, in particularly, our IPv4 implementation uses IPoption field to add a binary token to each IP packet. The token value is calculated similar to the XML token value by applying HMAC-SHA1 transformation to concatenated binary strings of the masked IP packet payload and GRI.

The TBN infrastructure consists of Token Based IP Switch (TBS-IP) that is controlled by inter-domain controllers in each domain. The TBS includes such major components as Token Builder (TB) and TVS that provides a similar functionality as defined in the GAAA-CRP framework. The applications' traffic is first tokenised by

the *TB* of a local domain (e.g., a campus network), after which it is enforced by the *TBS-IP* at each domain along the end-to-end path.

Tokens are used to label dataflows and can be made independent of upper layer protocols. In this way the token can be regarded as an aggregation identifier to a network service. The following four types of aggregation identifiers that can be combined are defined:

- identifier to link a service to the NE (e.g., a multi-cast, or transcoding);
- identifier that defines the service consumer (e.g., the grid application);
- identifier that defines the serviced object (e.g., the network stream);
- identifier that defines the QoS (security, authorisation, deterministic property, etc.).

The semantics that is referred to by a token (e.g., a certain routing behaviour) can be hard-coded into a TBS or dynamically programmed via TVS. Hence, a token provides a generic way to match/link applications to their associated network services. Tokens can be either embedded in the application generated traffic or encapsulated in protocols where embedding is not supported, such as in public networks.

To provide necessary performance for multi-Gigabit networks, TBS-IP is implemented using Intel IXDP2850 network processor that has a number of built-in hardware cryptographic cores to perform basic cryptographic functions such as required for TBN operation HMAC, SHA1, digital signature and encryption [11, 12].

TBS-IP control plane relies on a master-slave communication using ForCES protocol described in details in the next section.

It is important to mention that the TBN functionality can support Multi-Level Security (MLS) model [13] by labelling and encrypting dataflows between security critical applications at data and control planes while GAAA-CRP model allows flexible policy based reservations and access control at service-plane.

4.2. Using ForCES for network management at control and data planes

ForCES stands for Forwarding and Control Element Separation and is an upcoming IETF standard [14, 15]. ForCES defines a framework and associated protocol to standardize information exchange between the control and forwarding plane that comprise of Forwarding Elements (FE) and Control Elements (CE) correspondingly.

The basic building blocks of the ForCES model are the Logical Function Blocks (LFBs) described in an XML format. The ForCES protocol [15] works in a master-slave mode in which FEs are slaves and CEs are masters. The protocol includes commands for transport of LFB configuration information, association setup, status, and event notifications, etc. The protocol provides an open API for configuring and monitoring the Forwarding Plane in a standard manner. Grouping a number of LFBs, can create a higher layer service like TBS-IP in our case or a firewall. Similarly any security method at networking layer can be described using the ForCES model.

The ForCES standard framework defines the transport mapping layer (TML) to transfer the ForCES messages from the CE to the FE and vice versa. Currently defined is the SCTP TML that uses SCP protocol for secure messages exchange [16].

We consider the ForCES network management model as a way to integrate networking Control plane and Data plane into the general CRP process that requires

heterogeneous networks configuration at least at the deployment and decommissioning stages. Recent works to define Web Services interfaces to ForCES devices makes such integration even simpler [17]. In our GAAA-CRP implementation we use ForCES protocol for transferring TBS-IP configuration information from the inter-domain controller to TB and TVS.

The ForCES framework provides a standard way of adding security services to both CE and FE. When used in the CRP/NRP Grid/Networking infrastructure the ForCES security framework [16] can benefit from using the common AuthN/AuthZ infrastructure. In this case the standard GAAA-AuthZ components can be added and related policies defined for the basic ForCES security functions such as endpoints and messages authentication.

5. GAAA-NRP Implementation in GAAA-TK Pluggable Library

All proposed GAAA-AuthZ functionality is currently being implemented in the GAAA Toolkit (GAAA-TK) pluggable Java library in the framework of the Phosphorus project [18]. The library provides also a basis for building AAA/AuthZ server that can act as Domain Central AuthZ Service (DCAS) or operates as a part of the Inter-Domain Controller (IDC) and allows for complex policy driven resource reservation and scheduling scenarios.

The library allows for AuthZ request evaluation with local XACML based PDP or calling out to the external DCAS using the SAML-XACML protocol. Current library implementation [19] supports both XACML-Grid and XACML-NRP policy and attribute profiles as configurable metadata set. For the convenience of application developers, the GAAA-TK provides simple XACML policy generation tools.

The TVS component is implemented as a part of the general GAAA-TK library but can also be used separately. It provides all required functionality to support token based policy enforcement mechanism that can be used at each networking layer and in particular for token based networking. All basic TVS functions are accessible and requested via a Java API. Current TVS implementation supports shared secret and PKI based token key distribution.

The GAAA TK library provides few PEP and TVS methods that support extended AuthZ session management and provide necessary AuthZ token and ticket handling functionality (refer to the GAAA-TK release documentation [20] for the complete API description). The two basic PEP methods provide simple AuthZ session management and allow using AuthZ tickets or access tokens as session credentials, however they differ in either requiring complete request information or using AuthZ ticket or token as only access credentials. Both of these methods can either return a valid AuthZ ticket or token, or “Deny” value.

6. Summary and Future Research

This paper presented the results of the ongoing research and development of the generic AAA AuthZ architecture in application to two inter-related research domains:

on-demand optical network resource provisioning and Grid based Collaborative Environment that can use the same Complex Resource Provisioning model.

The proposed AuthZ infrastructure will allow easy integration with the Grid middleware and applications what is ensured by using common Grid/network resource provisioning model that defines specific operational security models for the three major stages in the general resource provisioning: reservation, deployment or activation, and access or use. The current implementation of the GAAA-NRP authorisation infrastructure and GAAA-TK library in the Phosphorus project multidomain networking testbed provides a good basis for further research on improving efficiency of the provisioning and authorisation sessions management and extending functionality of the session management mechanisms such as discussed in this paper AuthZ tickets, access and pilot tokens.

The authors will continue research into developing security and trust models for the GAAA-CRP and CRP to define requirements for key management in multidomain environment. Currently proposed and implemented TVS infrastructure uses a shared secret security model that has known manageability problems.

The authors believe that the proposed solutions for AuthZ session management in on-demand resource provisioning will provide a good basis for further discussion among Grid and networking specialists.

Acknowledgements

This work is supported by the FP6 EU funded Integrated project PHOSPHORUS (Lambda User Controlled Infrastructure for European Research) IST-034115.)

References

1. Demchenko Y, A. Wan, M. Cristea, C. de Laat, Authorisation Infrastructure for On-Demand Network Resource Provisioning, Proceedings The 9th IEEE/ACM International Conference on Grid Computing (Grid 2008), Tsukuba, Japan, Sept. 29 - Oct. 1, 2008. Pp. 95-103. IEEE Catalog Number CFP08GRI-CDR, ISBN 978-1-4244-2579-2.
2. Hafid A., A. Maach, J. Drissi, "A distributed advance reservation system for interconnected slotted optical networks: Design and simulations," Computer Communications, Volume 30 , Issue 5 (March 2007), Pages 1142-1151.
3. MSS Viola Meta Scheduling Service Project. [Online]. Available http://packcs-e0.scai.fhg.de/viola-project/
4. Yuanming, C., W. Wendong, G. Xiangyang, Q. Xirong, "Initiator-Domain-Based SLA Negotiation for Inter-domain QoS-Service Provisioning", Proc. 4th Int. Networking and Services, 2008, 16-21 March 2008. Pp. 165 - 169.
5. Vollbrecht, J., P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Bruijn, C. de Laat, M. Holdrege, D. Spence, "AAA Authorization Framework," Informational RFC 2904, Internet Engineering Task Force, August 2000. ftp://ftp.isi.edu/in-notes/rfc2904.txt
6. Demchenko, Y., C. de Laat, O. Koeroo, H. Sagehaug, Extending XACML Authorisation Model to Support Policy Obligations Handling in Distributed Applications, Proceedings of the 6th International Workshop on Middleware for Grid Computing (MGC 2008),

December 1, 2008, Leuven, Belgium. ISBN:978-1-60558-365-5. [Online] Available from http://portal.acm.org/citation.cfm?id=1462704.1462709
7. "An XACML Attribute and Obligation Profile for Authorization Interoperability in Grids," Joint EGEE, OSG, and Globus document. [Online]. https://edms.cern.ch/document/929867/1
8. Demchenko, Y., C. M. Cristea, de Laat, XACML Policy profile for multidomain Network Resource Provisioning and supporting Authorisation Infrastructure, IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY 2009), July 20-22, 2009, London, UK. Accepted paper.
9. Gommans, L., L. Xu, Y. Demchenko, A. Wan, M. Cristea, R. Meijer, C. de Laat, Multi-Domain Lightpath Authorization using Tokens, Future Generations Computer Systems, Vol 25, issue 2, February 2009, pages 153-160
10. Demchenko, Y., C. de Laat, T. Denys, C. Toinard, Authorisation Session Management in On-Demand Resource Provisioning in Collaborative Applications. COLSEC2009 Workshop, The 2009 International Symposium on Collaborative Technologies and Systems (CTS 2009), May 18-22, 2009, Baltimore, Maryland, USA.
11. "The Token Based Switch: Per-Packet Access Authorisation to Optical Shortcuts", by Mihai-Lucian Cristea, Leon Gommans, Li Xu, and Herbert Bos, in Proceedings of IFIP Networking, Atlanta, GA, USA, May 2007
12. "ForCES Token Based Switch Design and Implementation ", Phosphorus Project Deliverable D4.3.2, September 30, 2008. [Online]. Available: http://www.ist-phosphorus.eu/files/ deliverables/Phosphorus-deliverable-D4.3.2.pdf
13. Alkassar, A., C. Stuble "Security Framework for Integrated Networks", Proc. Military Communications Conference, 2003 (MILCOM2003), 13-16 Oct. 2003, Volume: 2, pp. 802-807. ISBN: 0-7803-8140-8
14. Yang, L., Dantu, R., Anderson, T., and R. Gopal, "Forwarding and Control Element Separation (ForCES) Framework", RFC 3746, April 2004.
15. Dong, L., Doria, A., Gopal, R., HAAS, R., Salim, J., Khosravi, H., and W. Wang, "ForCES Protocol Specification", http://www.ietf.org/id/draft-ietf-forces-protocol-22.txt (work in progress), March 2009.
16. Salim, J. and K. Ogawa, "SCTP based TML (Transport Mapping Layer) for ForCES protocol", http://www.ietf.org/internet-drafts/draft-ietf-forces-sctptml-04.txt (work in progress), July 2009.
17. Haleplidis, E., Haas, R., Denazis, S., Koufopavlou, O., "A Web Service- and ForCES-based Programmable Router Architecture", IWAN2005, France.
18. Phosphorus Project. [Online]. Available: http://www.ist-phosphorus.eu/
19. "GAAA Toolkit pluggable components and XACML policy profile for ONRP", Phosphorus Project Deliverable D4.3.1, September 30, 2008. [Online]. Available: http://www.ist-phosphorus.eu/files/ deliverables/Phosphorus-deliverable-D4.3.1.pdf
20. " Updated GAAA Toolkit library for ONRP (final project release)", Phosphorus Project Deliverable D4.5, March 30, 2009. [Online]. Available: http://www.ist-phosphorus.eu/files/ deliverables/Phosphorus-deliverable-D4.5.pdf