26.07.07

Deliverable DJ5.4.1: Advanced Technologies Overview

Deliverable DJ5.4.1

Contractual Date:	28/02/07
Actual Date:	26/07/07
Contract Number:	511082
Instrument type:	Integrated Infrastructure Initiative (I3)
Activity:	JRA5
Work Item:	4
Nature of Deliverable:	R (Report)
Dissemination Level	PU (Public)
Lead Partner	DFN
Document Code	GN2-07-142v3

Authors: T. Kersting (DFN, main editor), K. Baumann (SWITCH), O. Canovas (University Murcia), T. Chown (University Southampton), N. Goncalves (FCCN), J. Howlett (Ukerna), M. O'Leary (Ukerna), T. Lenggenhager (SWITCH), D.R. Lopez (RedIRIS), G. Lopez (University Murcia), S. Neinert (University Stuttgart), C. Rodriguez (RedIRIS), A.A. Solberg (Feide), M. Stanica (DFN), K. Wierenga (SURFnet) S. Winter (Restena)

Abstract

This document reviews technical developments in the JRA5 area that are not yet part of the project plan. It provides an overview of these Items, and also an evaluation of their applicability to the remaining period of the GÉANT2 project. The evaluation indicates whether to include these developments in future plans.

Table of Contents

0 Executive Summary i

1 Introduction 1

2 Roaming related technologies 3
- 2.1 Diameter 3
  - 2.1.1 Transport Layer 4
  - 2.1.2 Message Format 4
  - 2.1.3 Networking 4
  - 2.1.4 Roaming 5
  - 2.1.5 Error Handling 5
  - 2.1.6 Transition 5
  - 2.1.7 Extensions 5
  - 2.1.8 Hindrances 6
- 2.2 RadSec 6
  - 2.2.1 RadSec on different platforms 7
  - 2.2.2 RadSec Internet Draft 8
  - 2.2.3 Dynamic Discovery 8
- 2.3 WiMAX 8
  - 2.3.1 Impact on JRA5 9
- 2.4 Mobile networks 9
  - 2.4.1 IMS 9
  - 2.4.2 EAP-SIM 13
- 2.5 Evaluation from the JRA5 perspective 17

3 AAI related technologies 18
- 3.1 Federated Identity Management architectures 18
  - 3.1.1 Shibboleth 2.0 18
  - 3.1.2 Shibboleth 2.1 21
  - 3.1.3 Liberty Alliance 21
  - 3.1.4 CardSpace 22
  - 3.1.5 OpenID 23
- 3.2 Authorisation Management Systems 24

3.2.1 X-RBAC 24
3.2.2 ShARPe 24
3.2.3 Signet and Grouper 25
3.3 Standards 26
3.3.1 SAML 2.0 26
3.3.2 XACML 29
3.3.3 Resource Description Framework (RDF) 31
3.4 Evaluation from the JRA5 perspective 34

4 SSO related technologies 35
4.1 NAS-SAML 35
4.2 Internet2 proposal for SAML integration in RADIUS 37
4.3 Evaluation from the JRA5 perspective 39

5 Comprehensive technologies 40
5.1 IPv6 40
5.1.1 IPv6 in general 40
5.1.2 Current Status of IPv6 developments and deployments 40
5.1.3 MIPv6 42
5.1.4 Recommended Actions 43
5.2 Evaluation from the JRA5 perspective 44

6 Value-added services 46
6.1 SLP 46
6.1.1 Logical overview 46
6.1.2 SLP roles 47
6.1.3 Network protocol 47
6.1.4 Security 48
6.1.5 Adoption 48
6.2 SIP/IP Telephony 48
6.3 IP Telephony + eduroam 49
6.3.1 Voice over eduroam Challenges 50
6.4 Virtual organisations 51
6.4.1 Characteristics of a Virtual Organisation Framework 52
6.4.2 A Scenario for a Virtual Organisation – The Project 53
6.4.3 VO Membership Service 53
6.4.4 Confederation Aspects of VOs 54

6.4.5 Conclusions on Virtual Organisations 55
6.5 Evaluation from the JRA5 perspective 55

7 Conclusions 56

8 References 57

9 Acronyms 58

Table of Figures

Figure 1: IMS Architecture Diagram 12
Figure 2: EAP-SIM Processing 14
Figure 3: Architecture of a XACML-based infrastructure 29
Figure 4: PolicySet of an XACML Policy 30
Figure 5: NAS-SAML Components 36
Figure 6: Proposed architecture for SAML integration into RADIUS 38

0 Executive Summary

Following the JRA5 work plan, Work Item 4, “Integration of Advanced Technologies”, began in Year 3. The objective is to review the recent results of technical developments in the JRA5 area that are not yet part of the project plan. The intention is not only to provide an overview of these items but also to evaluate them with regard to their applicability to the remaining period of the GÉANT2 project. The evaluation will indicate whether to include them in future plans for roaming, AAI or uSSO, or even into more than one work item (called comprehensive technologies, with IPv6 as the most popular example).

An application-oriented view is also added, under the heading “value-added services”, to identify services that can be used based on the results of the JRA5 infrastructures, raising the profile of the JRA5 results in the process.

The list of results described in the document is by no means exhaustive. It would be useful to repeat this technology analysis at a later stage in Year 4.

The results of this evaluation are summarised in the table below:

Area	Technology	Conclusions	Page
Roaming			
	Diameter	Great promise for future authentication transport protocol, but not yet ready. Future development to be monitored.	3
	RadSec	Of great current use, with deployment in eduroam being planned.	6
	WiMAX	Development to be monitored, but future impact probably low. However, this may change as the technology matures and more end user devices become available.	8
	IMS EAP-SIM	Low potential, as eduroam is solely an IP network. Promising, but hindered by low deployment of SIM card readers.	9 13
AAI			
	Federated Identity Management	Shibboleth 2.0 will be automatically supported by eduGAIN (the confederated identity solution developed in work item 2 of JRA5) by its future transition to SAML 2.0 (see below).	18
	Authorisation Management	eduGAIN’s structure safeguards against future developments in this area (for example with Signet, Grouper or ShARPe).	24
	Standards	Development is planned for SAML 2.0 to be used as the future protocol for eduGAIN.	26
SSO			

Area	Technology	Conclusions	Page
	NAS-SAML	Identified as a potential solution to both push- and pull-models for SSO.	35
	Internet2 with RADIUS	A SAML profile for integrating network access based on RADIUS is currently being discussed under the auspices of MACE.	37
Comprehensive Technologies			
	IPv6/MIPv6	No current requirement for deploying IPv6 connectivity, though should be considered in near future at eduroam “hotspots”. MIPv6 not available in production-quality, and will only have significant benefit at “hotspots”.	40 42
Value-added services			
	SLP	SLP is considered useful for visiting users. Should be followed on a federational level, but not highly prioritised.	46
	IP Telephony + eduroam	Given the mobile nature of eduroam, problems considering the interruption of VoIP service when changing Access Points still need to be solved as well as various QoS issues.	49
	Virtual organisations	Still a lot of open questions before knowing for sure if the concept can be supported by eduGAIN.	51

1 Introduction

The pilot infrastructures of roaming and AAI were (and are being) built using the technology available at the time of writing the specification. The purpose of work item 4, “Integration of advanced technologies” is to investigate and evaluate technologies that have become available since the project started. If such a technology could add value to the Roaming and/or AAI, and has been successfully tested, the plan is to incorporate it into the existing infrastructures.

In reality, some technologies that became available during the duration of the project have already been either integrated or evaluated in the other work items. The result of this is that the technologies described here are less mature than originally foreseen. This means that many of the advanced technologies in this document may be investigated and tested in work item 4, but not integrated into the results of the other work items.

Work item 4 is aimed at new pilots and advanced experiments with new technologies and applications. Work on Roaming, AAI and SSO will continue within a broader participants group and gather feedback from practical experiences.

This document gives an overview of the technologies and concepts identified as candidates for further study. The first two sections of this deliverable analyse developments related to roaming and AAI. Some of them (such as RadSec, SAML2.0 and Shibboleth) are already addressed in the JRA5 work plan. These will be incorporated into the respective infrastructures as soon as they become available.

For uSSO, the Diameter and NAS-SAML research work is of great value. Technologies like XACML are applied in the DAMe subproject and may be of further interest for JRA5 work.

A very important candidate for an extended roaming testbed is mobile IPv6. As this protocol is built to support mobility, it enables enhanced usage for authentication and authorisation. Current implementations based on the latest Internet-Drafts have not yet been stable enough to be considered for the roaming set-up. Experiments are planned as part of an UKERNA project that will provide more insight.

IMS will allow for the smooth integration of new IP based services with previous 3GPP architectures and makes content based applications more attractive to mobile devices. The coverage of mobile networks might become an important extension to WLAN based solutions.

The Session Initiation Protocol (SIP), developed by the IETF as a standard for signalling, will also be used in IMS and in mobile networks. It provides a perfect scenario for tests with the performance sensibility (QoS) of real-time audio transmission, or as a basis for other value added services (such as instant messaging and location based services). This will close one of the gaps that currently still exist between the “Internet” and mobile networks.

2 Roaming related technologies

The development of the eduroam infrastructure provides the opportunity to start an eduroam service. However, further research will focus on the improvement of the current architecture and the extension of the protocols used, in order to enable the transfer of attributes as an important step towards the integration of roaming and AAI.

2.1 Diameter

Diameter is a protocol for the Authentication, Authorisation and Accounting (AAA) of users accessing an IP network. The base protocol was specified in RFC 3588 in September 2003. It is meant to be a successor to the widely used Remote Authentication Dial In User Service (RADIUS) protocol, as specified in RFC 2865 in June 2000. RADIUS was designed as a dial-up protocol for services like the Point-to-Point protocol (PPP) or Serial Line Internet Protocol (SLIP). Diameter, by contrast, is a full-blown AAA protocol and supports new access technologies presenting higher demands. One of the more important aspects here is that roaming users moving across the borders of administrative domains require special treatment. Several weak points of RADIUS are improved in Diameter, as explained in the following description.

The base protocol offers reliable transport mechanisms, specifies the format of messages and how they are exchanged. It defines the connection management for peering and routing and the management of application sessions. Diameter applications are protocols based on the Diameter protocol; they extend it for concrete use cases. Examples are the Diameter Network Access Server Application as defined in RFC 4005 application or the Diameter Mobile IPv4 Application as in RFC 4004.

Although Diameter is still mainly a client-server protocol, it has aspects of a peer-to-peer protocol. A Diameter server can start connections and trigger actions. Only when having this feature is it possible to explicitly disconnect clients, as they may have exceeded some quota or there may be abuse by the user. With RADIUS this is possible if RFC 3576: Dynamic Authorisation Extensions to Remote Authentication Dial In User Service (RADIUS) is supported.

In the following sections, properties of the Diameter protocol are analysed and compared to corresponding RADIUS properties, and possible effects on the eduroam infrastructure are shown.

2.1.1 Transport Layer

Diameter can be used on top of the Transmission Control Protocol (TCP) or Stream Control Transmission Protocol (SCTP), both of which are connection-oriented. These are more reliable and robust than the connectionless User Datagram Protocol (UDP) that RADIUS is based upon. TCP and SCTP both provide integrity of the data using checksums, ensure ordered delivery of messages, detect duplicates, and automatically resend messages when acknowledgements are missing. Unreliable delivery of messages is acceptable for Voice Over IP or computer games, where a single message is less important compared to efficiency and delay, but with an AAA system every single message is required and none can be omitted.

Diameter also offers flow control and congestion avoidance to deal with high network loads. Although SCTP is nowhere near as common as TCP, one special feature of SCTP is that it also supports multiple data streams in one connection (see RFC 2960). Both flow control and congestion avoidance will probably be helpful as traffic in the AAA infrastructure continues to grow, especially if there are any central elements like the root servers in current eduroam.

2.1.2 Message Format

The maximum length of one complete RADIUS message is 4096 bytes. This might be enough for transmitting user names and passwords, but it might not be enough for transmitting XML documents as with SAML or XACML. Diameter uses a length field of three bytes that would allow a maximum message length of 16 megabytes. The length of single attribute-value pairs (AVPs) is limited to 255 in RADIUS, as only one byte is used as a length field. In Diameter three bytes are used for the AVP length field. This message format allows transporting much larger payloads in one message.

Therefore, Diameter messages (and especially their headers) are designed for efficient processing by hardware as they have a 32-bit alignment.

2.1.3 Networking

RADIUS links between servers can only be configured manually by setting shared secrets. Because of this, creating networks of a large number of interlinked RADIUS servers does not allow expansion, and structural changes of a once established infrastructure will be laborious. Shared secrets are associated with a specific IP address, so all IP addresses have to be static.

Diameter offers manually configured server lists, as well as dynamic discovery of peers. Using DNS SRV, NAPTR and SLPv2, links can be established dynamically if the newly found server can be authorised. This feature is optional and might not be supported by all implementations. If it is available, it could reduce administrative efforts and be helpful in moving away from a hierarchical structure, as is the current situation in eduroam.

2.1.4 Roaming

Diameter offers better support for roaming. It has true routing capabilities, including routing tables, duplicate detection and loop detection, so it is more adequate for setting up networks of connected servers in a federated AAA infrastructure. With RADIUS, the proxying feature has to be “exploited”, resulting in a non-optimal treelike structure and the danger of unintentionally configuring loops. Similar to RADIUS proxying capability, Diameter offers forwarding of certain messages, but it differentiates proxying and relaying of messages. In the second case the content of messages will not be modified. This is not guaranteed when using RADIUS proxying. In addition to routing by realm, routing by host is also possible.

2.1.5 Error Handling

As installing and configuring RADIUS servers is a complex task (also true with Diameter), mistakes are probable. If these servers are connected and integrated in one large, European-wide network, complexity increases even more, and therefore so does the potential for error.

In RADIUS, messages are often silently discarded in case of an error. This is sensible for dial-up users, as the only mistake they could make is entering a wrong user name or password, and an administrator can look in the log files of his server to determine the cause of an error. However if an administrator wants to connect his server to one from another country, thereby connecting it to hundreds of other servers, it would be helpful if he could receive error messages. This is the case in Diameter, where clearly defined messages are returned for most error cases.

Another problem with RADIUS servers occurs when they are not reachable for any reason. The RADIUS RFC considers keep-alive messages harmful since they add to network load, and when clients request access the same effect would be acquired anyway. Diameter exchanges these messages to detect such errors pro-actively, because access requests have a user waiting for a response.

2.1.6 Transition

To simplify transition from RADIUS to Diameter the structure of the messages is very similar. They each consist of a header plus a sequence of AVPs. Also the first 256 AVP numbers correspond to RADIUS AVPs, so translation agents can convert between the two protocols readily. However, Diameter messages and also AVPs can be much longer than in RADIUS, so translation is not a trivial task. Together with the Network Access Server Application, a Diameter server can take the place of a RADIUS server, for example to provide AAA services for dial-in PPP users.

2.1.7 Extensions

RADIUS can be extended in two ways:

- New attributes can be added without disturbing the rest of the protocol. But as the protocol explicitly allows modification of attributes (with some explicit exceptions), there is no guarantee on how new attributes would be treated by a specific server.
- New modules can be added, as is the case with EAP-modules for example. The original code need not be modified; the new modules can be configured and integrated via the normal configuration files.

With Diameter it is possible to configure new attributes, and especially tagged vendor-specific attributes and commands can also be added. There are several predefined data formats to use with new attributes. Also it allows for “applications” to be used, such as protocols or solutions that operate on top of the base Diameter protocol. This provides a clearly described extension point for adding new functionalities. New applications can work with Diameter servers as they are able to negotiate capabilities, like whether they are using TCP or SCTP, which protocol version is used, and also what applications are supported.

2.1.8 Hindrances

Diameter is still quite a new protocol, for example the Network Access Server Application was defined in the form of an RFC in August 2005. Therefore, implementations of the protocol are uncommon at the moment. Current access points mostly only use RADIUS, not Diameter. Also, there are no server implementations available that are ready for production use.

The open-source implementation Open Diameter is currently available in Version 1.0.7h. It supports the base protocol, EAP and PANA, but the Diameter NASREQ Application API is still an alpha version (NASREQ is a name for the Network Access Server Application; it was used in the drafts preceding RFC 4005). Also, a translation agent between RADIUS and Diameter is not yet available, making transition to Diameter difficult. Diameter client and server daemon programs are planned to be included in Release 1.0.8.

Not even the new version 3.16 of the commercial Radiator fully supports the base protocol, but merely implements a translation agent to support the Network Access Server Application. In March 2006 OSC was still planning on releasing a full Diameter server when the market would be ready, but this has not happened yet.

2.2 RadSec

RadSec is a new transport profile for RADIUS packets, providing a reliable transport for RADIUS (by using TCP) and a strong cryptographic security layer (by using TLS over the TCP connection). Optionally, other authentication servers can be discovered dynamically, thereby rendering the classical RADIUS IP address and shared secret bindings obsolete. These features make RadSec a lightweight, easy to implement and maintain alternative to the Diameter protocol. Also transition from RADIUS to RadSec is much easier than from RADIUS to Diameter since the core protocol remains unchanged. A detailed description of the original RadSec implementation is found in the deliverable Inter-NREN Roaming Architecture: Description and Development Items [DJ5.1.4], sections 3.2 and 3.4.

Many of the advantages offered by Diameter can also be achieved with RadSec. Among those explained in the previous section, the following can be obtained by using RadSec:

- Transition: Transitioning from RADIUS to RadSec is negligible, as the payload is identical bit-by-bit; only the means of transporting the payload change.
- Reliable Transport (TCP): SCTP is not specified for the RadSec Internet Draft since the SCTP protocol is currently not mature enough to be considered for a real-world deployment. This may change in later RadSec versions.
- Flow control and congestion Avoidance: These properties come built-in into TCP stacks (RFC-2581).
- Dynamic Peer Discovery and independence of IP-address-to-shared-secret bindings.
- Pro-active error detection. The state machine of RFC3539 section 3.4.1 is optionally supported using the RADIUS Status-Server message.

JRA5 has already selected RadSec as a candidate for future enhancements of the eduroam infrastructure, after the evaluation presented in chapter 4 of deliverable DJ5.1.4. One of the consequences of the evaluation is that it is advisable to produce alternative implementations of the protocol as well as to write a formal specification of the protocol. These two objectives should be pursued in Work item 1. A summary of the efforts that are being undertaken is given in the sections below.

2.2.1 RadSec on different platforms

During the initial JRA5 evaluation of RadSec for DJ5.1.4, only a single implementation existed: the product Radiator from Open Systems Consultants (OSC).

Radiator is a cross-platform software because it is based on the scripting language Perl and will run on any platform provided Perl is available. The drawback of this implementation is that it is commercial software – a license of several hundred Euros is required to use it on a permanent basis. This is likely to be unacceptable for some educational institutions.

JRA5 provided a second implementation based on ANSI-conformant C code. The aim was to create a lightweight proxy that can be installed alongside arbitrary RADIUS implementations, thus providing RadSec capabilities to each RADIUS device independently of vendor, computing capacities or platform. A prototype is already available[1] and proven to interoperate with Radiator. It is small enough to be installed even on NAS devices, which makes it possible to enable the reliable, secure RadSec transport on the entire authentication chain between NAS and home authentication server. One NAS device candidate is the Linksys WRT-54g series Access Point. This can be used with an embedded flavour of Linux as firmware and is thus able to run the RadSec proxy.

[1] http://software.uninett.no/radsecproxy/

2.2.2 RadSec Internet Draft

The original RadSec whitepaper by OSC specified RadSec less precisely than is usually expected from a network protocol. Furthermore, the whitepaper and the implementation left many degrees of freedom for a Radiator server administrator. One example is that TLS encryption was optional, which led to potential configuration errors with a high impact on security (message payload could be transmitted in clear text).

In a combined effort of JRA5 and the TERENA Task Force Mobility, an attempt to standardise RadSec at the Internet Engineering Task Force is underway. The specification is reviewed by a group of experts, including major RADIUS server software implementers.

The Internet draft specifies the protocol more precisely than the original vendor's whitepaper and puts more constraints on the protocol usage. It will therefore define RadSec version 2. The changes are going to be so that the existing Radiator RadSec implementation is already conformant to version 2 with the right set of configuration options. The JRA5 implementation will be conformant to version 2 from the beginning.

It is expected that the RadSec Internet Draft will be presented to the IETF at the 69th IETF meeting.

2.2.3 Dynamic Discovery

RadSec does not specify an authoritative means of finding a peer dynamically. This is not necessary because the server authorisation model uses server certificates to determine if a discovered peer is authorised to perform user authentications or not. Therefore, any means of looking up peer addresses is appropriate, be it a lookup based on insecure, distributed databases (most notably DNS) or secure lookups (DNSSec, LDAP or the eduGAIN MDS).

The existing Radiator implementation has a DNS resolution mechanism, which operates similar to the one defined in the Diameter RFC3588. The JRA5 implementation is envisaged to have the same DNS resolution mechanism, plus the facility to plug in other lookup mechanisms, such as the eduGAIN MDS.

This flexibility in peer resolution mechanisms outdoes Diameter's peer resolution capabilities. According to RFC3588, Diameter resolution is limited to manual configuration, Service Location Protocol v2 (SLPv2) or DNS resolution.

2.3 WiMAX

WiMAX is a technology standard that enables wireless broadband communication between fixed stations. The technology can be used, for instance, to connect households to the Internet in areas where no fixed connection is available, or to provide backhaul for Wi-Fi hotspots. Therefore WiMAX can be seen as an alternative to DSL or cable. There is also a variant, which enables support for mobile clients.

The name "WiMAX" is a term introduced by the WiMAX forum. WiMAX covers the IEEE 802.16 specifications for WLANs. Today a typical WiMAX set-up consists of a wireless base station that is connected to a fixed network surrounded by multiple fixed-wireless clients. Clients can interact with the base station to gain access to resources such as the Internet. The base station decides which client may transmit (following a scheduling algorithm), so it is not necessary for clients to receive each other's signals. The implication is that clients can be further away from the base station. This makes WiMAX the preferred technology for rural areas.

WiMAX equipment operates in the 2-11 GHz range or 10-66 GHz range, depending on the specific sub-standards. Sub-standards include, for example, IEEE 802.16d and 802.16e that allow communication with a data rate up to 70 Mbps in total in the 2-11 GHz frequency range. IEEE 802.16d is suitable for fixed-wireless applications and the IEEE 802.16e standard adds capability for full mobile support.

A license is needed for transmission of radio signals in the Wireless Local Loop frequency bands. For licensed WiMAX, radiation powers above 100 mW are allowed. This means that coverage is extended, but battery issues at the client side may limit actual useful range. The WiMAX IEEE 802.16e standard adds mobility requirements, including features to lower battery usage.

2.3.1 Impact on JRA5

WiMAX uses an authentication and encryption scheme similar to the Wi-Fi WPA2 standard. The difference is that WiMAX security will use the PKM-EAP method with 3DES or AES encryption. WPA2 typically uses EAP-TTLS, PEAP or EAP-TLS with AES encryption. In order to use eduroam authentication for a WiMAX network it is imperative that tunnelled protocols be supported. Vendors have stated that they will support this, but it is at this point unclear what operators will offer.

2.4 Mobile networks

2.4.1 IMS

2.4.1.1 *Introduction*

IP Multimedia Subsystem (IMS) is a set of specifications that describes the Next Generation Networking (NGN) architecture for implementing IP based telephony and multimedia services. IMS defines a complete architecture and framework that enables the convergence of voice, video, data and mobile network technology over an IP-based infrastructure.

IMS was initially defined by the Third Generation Partnership Project (3GPP), which is a collaboration agreement among a number of telecommunications standards bodies as part of their standardisation work for supporting GSM networks and radio technology evolution. IMS was first introduced in 3GPP Release 5, where “Session Initiated Protocol“ (SIP), defined by the Internet Engineering Task Force (IETF), was chosen as the

main protocol for IMS. It has been further enhanced in Release 6 and 7 of 3GPP to include additional features like presence and group management, interoperation with WLAN and CS based systems, and Fixed Broadband access.

Another standards body, Third Generation Partnership Project 2 (3GPP2), also standardised their own IMS. 3GPP2 developed North American and Asian Cellular Radio-telecommunication Intersystem Operations into a third-generation system. The initial release of 3GPP2 specifications on IMS largely adopts from 3GPP Release 5. The two IMS networks defined by these two organisations are fairly similar, but not exactly the same. 3GPP2 added appropriate adjustments for their specific issues. Nevertheless, the purpose of both organisations is to ensure that IMS applications will work consistently across different network infrastructures.

In addition to the 3GPP and 3GPP2, Open Mobile Alliance (OMA) plays an important role on specifying and developing IMS service standardisations. The services defined by OMA are built on top of IMS infrastructure, such as Instant Messaging (IM), Presence service and Group Management Service.

2.4.1.2 *IMS architecture*

IMS architecture supports a wide range of services that are enabled based on SIP protocols. As can be seen from the Figure 1, an IMS architecture delivers multimedia services that can be accessed by a user from various devices via an IP network or traditional telephony system. It also allows the service creator the ability to combine services from CS and PS domains in the same session, and for sessions to be dynamically modified at any time (e.g. adding a video component to existing voice session). The underlying network architecture can be divided into three layers (Device Layer, Transport Layer and Control Layer), plus the service layer, and will be introduced from bottom to top respectively.

- Device Layer: The IMS architecture provides a variety of choices for users to choose end-point devices. The IMS devices such as computers, mobile phones, PDAs and digital phones are able to connect to the IMS infrastructure via the network. Other types of devices, such as traditional analogue telephones, although they are not able to connect to an IP network directly, are able to establish the connection with these devices via a PSTN Gateway.
- Transport Layer: The transport layer is responsible for initiating and terminating SIP sessions and providing the conversion of data transmitted between analogue/digital formats and an IP packet format. IMS devices connect to the IP network in the transport layer via a variety of transmission media, including Wi-Fi, DSL, Cable, SIP, GPRS (General Packet Radio Service) and WCDMA (Wideband Code Division Multiple Access, a type of 3G cellular network). In addition, the transport layer allows IMS devices to make and receive calls to and from the PSTN network or other circuit-switched networks via the PSTN gateway.
- Control Layer: The Call session Control Function (CSCF), which is a general name that refers to SIP servers or proxies, is one of the core elements in the control layer. CSCF handles SIP registration of the end points and processes SIP signal messaging of the appropriate application server in the service layer. Another element in the control layer is the Home Subscriber Server (HSS) database that stores the unique service profile for each end user. The service profile may include a user's IP address, telephone records, buddy lists, voice mail greetings and so on. By centralising a user's information in

HSS, service providers can create unified personal directories and centralised user data administration across all services provided in IMS.

- Service Layer: On top of the IMS network architecture is the service layer. The three layers described above provide an integrated and standardised network platform to allow service providers to offer a variety of multimedia services in the service layer. Application servers run all the services. The application servers are not only responsible for hosting and executing the services, but also host multiple services, for example, telephony and messaging services run on an application server; one advantage of this flexibility is to reduce the workload of the control layer. There are many application servers providing different services and three core application servers of IMS will be highlighted below.
  - Presence server: A “Presence server” provides the services to collect, manage and distribute the real time availability and the means for communicating among users. It allows users to both publish their presence information and subscribe to the service in order to receive notification of changes by other users.
  - Group List Management server: A “Group List Management server“ provides services that allow users or administrators the ability to manage, create, modify, delete and search the network-based group definition and the associated lists of members. It also maintains the access permissions and other specific properties associated with the groups and the members. It is also used to provide buddy lists for instant messaging or other services.
  - Instant Messaging Server: An “Instant Messaging server“ provides a communication service that allows users to send and receive messages instantly. Users are able to deliver the messages containing rich text, images, audio, video or the combination of these over an IP network. It has been widely uses in today’s Internet community and IMS will bring the same service experience into the mobile world.

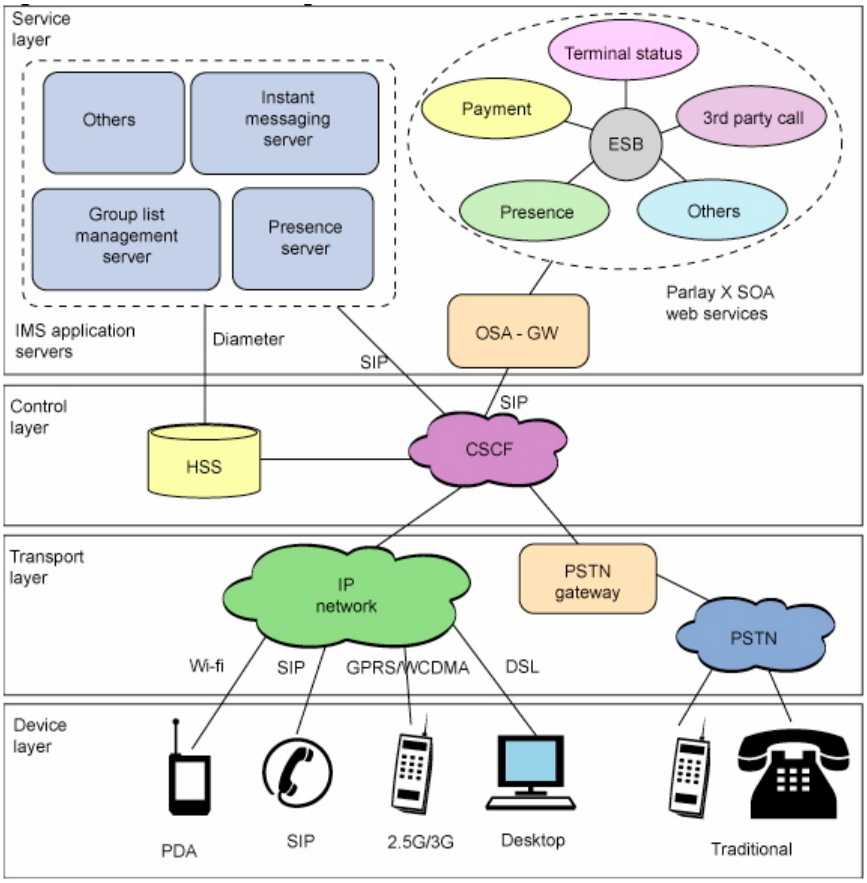

Figure 1: IMS Architecture Diagram

2.4.1.3 *Quality of Service*

Real-time Mobile IP communication is difficult due to fluctuating bandwidths, which severely affect the transmission of IP packets through the network. In normal IP networks, IP transport would be what is known as “best effort“, meaning that the network will do its best to ensure the required bandwidths, but there is no guarantee. The result is that real-time mobile IP services function poorly or not at all (i.e. voice quality is poor or garbled, video “jitter“, etc.) depending on the bandwidth availability and network congestion.

The Quality of Service (QoS) mechanisms were developed in order to overcome these issues, and provide some type of guaranteed level of transmission instead of “best effort“. QoS ensures that critical elements of IP transmission (such as transmission rate, gateway delay and error rates) can be measured, improved and guaranteed in advance. Users are able to specify the level of quality they require depending on the type of service and the users’ circumstances.

The “intelligence” required to enable QoS within a mobile IP network is specified in IMS in the Policy Decision Function (PDF). The PDF interacts with and controls the underlying packet network.

2.4.1.4 *Security for "early IMS" implementations*

It is envisaged that security defined in TS 33.203 (3GPP specification "3G security: Access security for IP-based services") may not be available for a while, especially because of the lack of USIM/ISIM interfaces and prevalence of devices that support IPv4. For this situation, to provide some protection against the most significant threats, 3GPP defines some security mechanisms. These are informally known as "early IMS security", in TR 33.978.

2.4.1.5 *Impact on JRA5*

IMS is nothing new in the sense that the underlying technical concepts of IMS have been discussed in standardisation and technology circles for some time. IMS has had a relatively low profile to date, mainly being relegated to technical discussion forums, whose main concern was in defining IMS network in technical terms. In terms of deployment there is no direct relation to JRA5 as IMS is a pure telco technology. It might be interesting as an additional access to the network, especially considering the broad availability of cellular networks, but chances of the paradigm of "likewise support" as used in eduroam are estimated to be very minor. The devices used for IMS might be interesting in terms of Authentication because of the USIM/ISIM (Universal Subscriber Identity Module/IP Multimedia Services Identity Module) interfaces.

2.4.2 EAP-SIM

2.4.2.1 *Introduction EAP, EAP-SIM*

EAP stands for Extensible Authentication Protocol and is primarily developed as a Point-to-Point Protocol (PPP); today we are speaking about a generic Authentication Protocol according to RFC3748. EAP-SIM is a mechanism for mutual authentication and session key agreement using the Global System for Mobile Communications (GSM) Subscriber Identity Module (SIM).

EAP-SIM in WLAN's

EAP-SIM can be used for fast connect in hotspots environment. A hotspot operator also provides access to the GSM network through standard WLAN configuration consisting of clients (supplicants) with 802.1x WLAN, EAP-SIM software, an authenticator, Wireless Access Point (AP), and an authentication server (an AAA/RADIUS server). During the authentication process, the AAA/RADIUS server uses information on the SIM card to access the GSM authentication centre by a GSM/MAP/SS7 gateway. If the SIM card and EAP-SIM software are able to validate the GSM triplets, the AAA/RADIUS server requests that the AP grants network access to the users. The AP connects the user's device to the WLAN and forwards accounting-information to the AAA/RADIUS-server to indicate that the connection has been completed. Such information can be pooled into a database for billing purposes. Thus the use of EAP-SIM in WLAN's is an emerging topic.

In accordance to the "Meetinghouse, Network Access Security" several usage scenarios are emerging, for example USB/PCMCIA SIM access, SIM access from GPRS PC cards, SIM access from Mobile phone using Bluetooth, SIM reader hardwired within the Notebook and so on.

Process of EAP-SIM in details

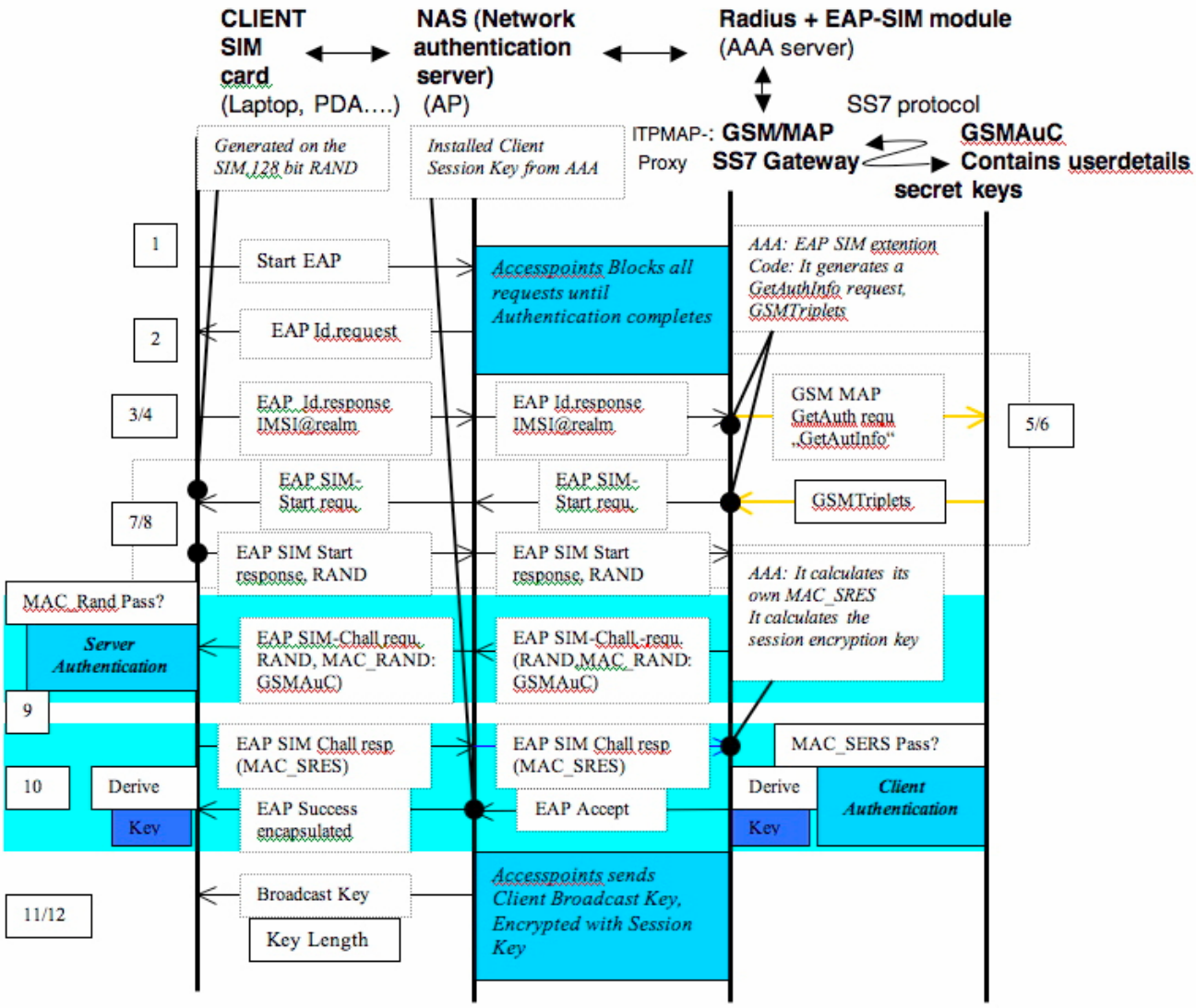

Figure 2: EAP-SIM Processing

1. EAPOW-Start-message from the client starts the authentication protocol and indicates to the AP that the client wants to authenticate using EAP.

2. In response, the access point sends an EAP-Identity Request message to the client. At this point, the client has not yet been assigned an IP address, and the AP blocks all messages from the client except for those necessary for authentication (EAP, EAP-SIM protocol messages).

3. The client responds to the access point's request with an EAP-Identity-Response-message containing the user's network identity. This identity is read from the SIM-card. (e.g. 0<IMSI>@<realm> where IMSI

is the International Mobile Subscriber Identity – as used in GSM networks). Further <realm> is the operator's domain name string (e.g. switch.ch).

4. The AP forwards the EAP Identity Response to the AAA server using a RADIUS-protocol-message.

5. The AAA server determines that the user intends to use EAP SIM authentication based on its configuration parameters or on the identity passed to it and invokes its EAP SIM extension code. This code then starts the EAP SIM extension protocol by sending an EAP SIM Start request back to the Client. It may also generate a GetAuthInfo message to the (GSM) AuC requesting a number of GSM triplets; this step may be delayed until after a response to the EAP-SIM-Start-message is received to ensure that the client indeed supports the EAP-SIM protocol.

   Depending on the realm contained in the identity string, the AAA-request might need to be proxied from the local AAA server to the service provider's AAA-server.

6. The GetAuthInfo message is routed to the Internet Transfer Point Mobile Application Part (ITPMAP-Proxy), which acts as a gateway to the service provider's SS7 network. The ITP translates the request into a standard GSM-MAP-GetAuth-request before sending it to the (GSM) AuC.

7. On Receiving the EAP-SIM-Start-request, the client reads a 128 bit random number generated on the SIM and passes it back to the AAA server in the EAP-SIM-Start-response.

8. Once the AAA server has received the client's EAP-SIM-Start-response and the response from the (GSM) AuC containing a sufficient number of GSM-triplets (typical two to three), it then constructs an EAP-SIM-Challenge-message that contains the random numbers (RAND) received from the (GSM) AuC and a 160 bit message-authentication code (MAC_RAND).

9. The client passes the EAP-SIM-Challenge-request to the SIM card, which first calculates its own MAC_RAND. The AAA server is validated if the result matches the MAC_RAND received from the server. Only in that case, does the SIM calculate the GSM result (SRES) and encryption key (Kc) for each of the RANDs it received, as well as a 160-bit message-authentication code (MAC_SRES) based on these results and the user identity. Only MAC_SRES is returned to the AAA server (and therefore exposed on the radio link) in the EAP-SIM-Challenge-response. The SIM also calculates cryptographic keying material, using a secure hash function on the user identity and the GSM-encryption-keys, for the derivation of session encryption keys.

10. When the AAA server receives the client's EAP SIM Challenge response, it calculates its own MAC_SRES and compares it to the one received from the client. If both match, the client is authenticated and the AAA server also calculates the session encryption keys. It then sends a RADIUS ACCEPT message to the AP, which contains an encapsulated EAP Success message and the (encrypted) client session key.

11. The AP installs the session key for the client's association ID and forwards the EAP Success message to the client. It then sends its broadcast key, encrypted with the client's session key, in an EAP Key

message to the client. It also unblocks the data path for the client so that IP traffic can flow between the client and the rest of the network.

12. Upon receiving the EAP Success message, the EAP SIM supplicant returns the session encryption key calculated by the SIM to the EAP framework, which installs it on the client's WLAN radio card.

The client is now able to securely send and receive network traffic.

The advantages of EAP-SIM over conventional SIM

EAP-SIM has the following advantages with respect to conventional SIM:

- Extended functionality of the AAA server (better authentication methods, authorisation and accounting) over an AuC (Authentication Centre)
- Mutual authentication of EAP-SIM (see figure 1, far right column).

2.4.2.2 *Evaluation from the JRA5 perspective*

Possible Involvement in JRA5

The focus for an EAP-SIM-Implementation should be the sharing of the infrastructure between the Edu-, and Commercial-World, based on a multiprovider-enabled-architecture. Multiprovider-enabled-architecture means that different providers use the same SSID to allow their customers to gain access to the Internet at the campus. "Seamless Roaming" should be an additional requirement. The user should not be involved with the choice of the used technology like UMTS, EDGE, GPRS, HSDPA, etc.

The idea of an educational institution taking the role of a SIM-provider is not impossible, but rather improbable, because of the interests of building up an edu-network-infrastructure are not primarily commercially oriented.

Similarity to eduroam

With eduroam, all components required to realise an EAP-SIM-implementation are in place; IEEE 802.1x enabled AP's and Radius Servers like FreeRadius, Radiator and CiscoACS, performing the authentication-process. The challenge to implement EAP-SIM is rather to design a scalable architecture including an intelligent approach of managing (EAP-)SIM-AuthN-Traffic. It is necessary to differentiate between higher educational institutions as a SIM-provider, an EAP-SIM-infrastructure totally managed by a GSM-provider on the campus and an EAP-SIM-implementation in the context of a multiprovider-enabled-architecture and a centralised proxy (Radius Server) processing the EAP-SIM-AuthN-Traffic in responsibility of the NREN. The second approach can be considered to be the easiest way to introduce EAP-SIM on the campus, in that version all (EAP-)SIM specific components for handling EAP-SIM-AuthN-Traffic are in the responsibility of the GSM-provider. This solution is very restricted in terms of the provider and finally of the scalability.

Thus the future of EAP-SIM is situated in a scaling solution based on a multiprovider-enabled-architecture.

Different EAP-methods on the same infrastructure

All hardware-components, including the configuration to manage EAP(-SIM), are the same for all EAP methods. The major difficulty in using different EAP methods on the same infrastructure consists of realm-based-routing of the Radius-request by the Radius Server. This problem seems to be solved for Radiator, (see White paper: EAP-SIM Support) equipped with a GSM/MAP/SS7-gateway. For FreeRadius the integration of an EAP-SIM-module is still in progress. The argument of which EAP-method should be used, depends on how scalable the infrastructure has to be.

The experience with EAP-SIM shows different kinds of requirements. Wireless Internet Service Providers are more interested in different EAP-methods whereas Classical GSM-Providers like Swisscom are closed focused on EAP-SIM.

2.4.2.3 *Remarks and Conclusion*

EAP-SIM is a good example of an authentication mechanism that not only provides client authentication, but also establishes encryption keys for data traffic. In addition, the SIM card is widely considered as a user authentication mechanism, since it is actually a device authentication method that is rarely removed from the hardware.

2.5 Evaluation from the JRA5 perspective

The Diameter protocol shows a great promise as a future authentication credential transport protocol. It is currently limited by insufficient implementations, but its further development and deployment should be monitored thoroughly by JRA5. As an intermediate solution, RadSec is of large importance to the eduroam community. Its deployment in the eduroam production environment is currently being planned. A standardisation effort is underway and might help adoption of RadSec in other servers or even NAS devices.

It is yet unclear how deployments of Wi-Fi and WiMAX can augment or replace each other. WiMAX is a technology that should be closely monitored. The impact in the near future is probably low, but may rise significantly as the technology matures and more end user devices become available.

The impact of IMS appears to be low because it is a collection of technologies that are used by telecommunications providers, whereas eduroam is a pure IP network based service. However, future developments may open up the possibility that a user authenticates to the eduroam network with credentials he has acquired by a different IMS service. So, new authentication protocols within EAP may evolve. One of these EAP authentication protocols that may have an impact is EAP-SIM to identify users by means of their SIM card. This approach shows promise, but is currently hindered by lack of large-scale deployments of SIM card readers on campuses. If a campus overcomes this hindrance, EAP-SIM may prove to be a valuable authentication mechanism.

3 AAI related technologies

The term AAI is very much confined to the NREN community. In the commercial and governmental worlds, these kinds of technology are included in the broader field of identity management. Nevertheless, NRENs have been one of the key players in this arena, and many of the concepts and technologies that are now being deployed worldwide have their roots in the academic networking developments and practices. This allows the research area to be optimistic about the possibility of not only keeping pace with the evolution of these technologies but also of directly influencing it.

3.1 Federated Identity Management architectures

3.1.1 Shibboleth 2.0

The information contained in this chapter is all derived from the Shibboleth 2 Roadmap as presented in detail on the Shibboleth Wiki: https://spaces.internet2.edu/display/SHIB/ShibTwoRoadmap.

Shibboleth 2.0 will be based on the new OpenSAML 2.0 library and will add support for SAML 2.0 Authentication, Single Logout, and Artifact resolution request/response pairs. There will also be a new discovery service added, and attribute push will become the preferred flow.

- Shibboleth 2.0 will fully interoperate with Shibboleth 1.3.
- Certificates will need to be embedded in the metadata to take advantage of many SAML 2.0 features.
- SAML 2.0 Single Logout will be implemented with both front-channel browser and back-channel SOAP calls. Applications that maintain session information will need modification to handle logout requests, but it's not yet entirely determined how.
- SAML 2.0 Authentication will include some basic support for the 'AuthnContext' concept, an XML schema to express both technical and procedural aspects of user authentication and identity proofing.

3.1.1.1 *IdP Changes*

- The Shibboleth 2.0 identity provider will be a Java web application meeting the Servlet 2.3 specification.
- The *Protocol Endpoints* are the entry points into the IdP. They are responsible for receiving and dispatching requests to an appropriate protocol handler based on the path of the requested URL.
- *Protocol Handlers* receive messages and process them with the help of further subsystems.
- The *Session Manager* is responsible for tracking information about a user's session. This includes when the user first logged in, the authentication mechanisms that the user has authenticated with, and the services the user is logged in to.
- *Authentication Handlers* are used by authentication protocol handlers in order to authenticate a user.
- The *Attribute Resolver* pulls data from organisational data sources into the system through data connectors, transformed by attribute definitions, and converted into a protocol specific representation by attribute encoders. In addition to user attributes, the Attribute Resolver is also responsible for creating user name identifiers and translating name identifiers into a principal (user) name through a special principal connector.
- Based on the relying party, the *Attribute Filtering Engine* filters attributes and values as well as name identifiers. All of this is controlled by one or more policy files. These files may be signed allowing one or more to be safely administrated by an external entity, such as a federation management body.
- The service component model allows for services to be created as extensions and hooked into the IdP start up process. Examples may be a caching system shared by multiple components, a health monitor, or a special audit and logging service.

3.1.1.2 *C++ SP changes*

- The export of assertions to the surrounding environment will be modified.
- Apache attribute export will be performed by default with sub process environment variables rather than HTTP header variables.
- Attribute caching functionality will be simplified by removing unused and confusing options for refreshing and expiring attributes.
- There will be support for clustering sessions and other state information using ODBC-compliant databases.
- Query functionality will be better modularised for potential re-use in the future.
- IIS GUI integration may be improved.

3.1.1.3 *Java SP Changes*

- The Java SP will be implemented as a Java servlet 2.3 filter. Its subsystems will be strung together using Spring 2.0 and all of this rests on top of OpenSAML 2.
- The *Servlet Filter* is a message routing engine that will inspect the URL path component of the incoming request and hand that request off to the correct profile handler.
- A *Profile Handler* processes an incoming request using the rules of a given profile. Some example profiles are SAML 1 SSO and SAML 2 Logout. Specifically, the profiles will leverage the OpenSAML 2 message decoders to unpack the message, receive the message payload (normally a SAML message),

process the message, prepare a response if necessary, and use a message encoder to pack the response into the outgoing response object.

- The *Session Manager* manages the user's session. SAML 2 incorporates the notion of a "global" session – a super-session encompassing all the sessions in all the applications a user is currently logged into (including the identity provider a user came from). The session manager deals with creating, maintaining, and destroying the SP session and communicating this state with the IdP, if necessary. Hooks may be added to deal with application-specific rules for session management.
- The *Attribute Resolver* is responsible for pulling user specific attributes from different sources using data connectors. The first data connector to be implemented will pull this information from SAML attribute statements received from the user's IdP, but other sources such as relational databases and LDAP directories could provide information as well.
- The *Attribute Filtering Engine* filters each set of attributes coming in from a data connector. Afterwards, the raw attributes will be run through a series of attribute definitions that will create the final attributes to be released to the protected resource. These definitions may create, delete, or transform attributes in any manner. Finally the attribute service will make the final attribute set available to the application as either 'HttpServletRequest' attributes or 'HttpSession' attributes. The developers are currently working to make sure that data connectors and attribute definitions are shareable between the IdP and SP.

3.1.1.4 *Discovery Service (also known as WAYF) Changes*

- In Shibboleth 1, a service provider sends a user to the discovery service and the discovery service in turn sends the user to the appropriate IdP.
- In Shibboleth 2, the service provider will send the user to the discovery service which will in turn send the user back to the service provider though the purpose of the discovery service remains the same: identifying the user's identity provider. This new discovery service protocol will be submitted to OASIS.
- The Shibboleth 2 Discovery Service will be written as a Java web application using the servlet 2.3 specification. It will present the user with a list of identity providers to choose from. In order to make this list more manageable, it will allow custom filter plug-ins to be provided, these plug-ins may remove SPs from the list of entities, use the SAML IdP cookie to bubble up recent choices, or use any other mechanism they can implement to filter the list presented to the user. In addition the discovery service will be able to read metadata from multiple sources (which may in turn represent multiple federations) and present that information to the user in a number of different ways.

3.1.1.5 *OpenSAML Changes*

- Complete rewrite based on a generic XML processing library
- Support for SAML 2.0 as well as 1.0 and 1.1
- XML Digital Signature and Encryption Support
- Incorporation and extension of Shibboleth trust code

3.1.2 Shibboleth 2.1

Work in 2.1 will primarily focus on adding support for SAML 2.0 NameIDManage, NameIDMapping request/response pairs and support for mechanisms like WS-Security in securing SAML SOAP requests.

Additional features may include standards-based extensions to support the integration of web services, such as web-based portals using WSRP services, library meta-search, etc.

3.1.3 Liberty Alliance

Liberty Alliance, also known as Project Liberty, is an industry standards consortium developing specifications for federated identity management. In essence Liberty Alliance is specifying two standard suites of direct interest to JRA5, ID-FF and ID-WSF.

3.1.3.1 *ID-FF – The Liberty Identity Federation Framework*

The ID-FF standard is a direct improvement of the SAML 1.X standard. The ID-FF standard enhances the SAML 1.1 standard with new functionality; like session handling, improved logout and authentication contexts. The most important addition is the introduction of Liberty's concept of **federated identities**. The concept is about mapping existing service user accounts with identities at the IdP by using a persistent pseudonym for the user that is different for each service. This concept encourages user privacy by not requiring a global unique identifier for users across services. ID-FF also introduces a protocol to manage updates and removal of those pseudonyms, referred to as "NameIdentifier Management".

Even though ID-FF introduces enhancements to SAML 1.X, the introduction of ID-FF increased incompatibility between identity management software implementations because it is not compatible with SAML. Luckily OASIS and Liberty Alliance joined forces in the specification of SAML 2.0. OASIS incorporated most of the ID-FF features in SAML 2.0, including federated identities, and Liberty Alliance defines SAML 2.0 as the successor of ID-FF 1.2.

As Feide, the Norwegian federation, is built on software supporting ID-FF 1.1, ID-FF 1.2 and SAML 2.0, the Feide Bridging Element can be re-used to connect any ID-FF-based SP or IdP to eduGAIN.

Although SAML 2.0 specifies the federated identity concept, we will see several SAML 2.0 entities that do not support this functionality. One example is Shibboleth, which is not planning to implement "NameID Management" before version 2.1.

3.1.3.2 *ID-WSF – Liberty's Web-services Framework.*

The ID-FF protocol specifies communication between users, IdPs and SPs. In contrast, ID-WSF focuses on communication between services. ID-WSF is a supporting-protocol included in the headers of SOAP messages

between ID-WSF enabled services. ID-WSF does not specify any end-user services itself, but rather defines a platform and architecture to build identity-enabled services upon. The concept of federated identities is important in ID-WSF, allowing services to exchange information about principals without having known identifier in common.

For the more advanced and distributed services using eduGAIN, ID-WSF has been considered a candidate. However ID-WSF is a complex framework and the list of open source implementation is not yet impressive. The need to have a ready solution within a short period of time has resulted in a simpler solution for these services.

Currently there are multiple open source initiatives implementing ID-WSF toolkits. These are still in an early state, but JRA5 should monitor the progress.

3.1.4 CardSpace

Windows CardSpace, formerly known by its codename InfoCard, is a framework developed by Microsoft that securely stores digital identities of a person, and provides a unified interface for choosing the identity for a particular transaction, such as logging in to a website. Windows CardSpace is a central part of Microsoft's effort to create an Identity Metasystem, or a unified, secure and interoperable identity layer for the Internet.

When a CardSpace-enabled application or website wishes to obtain personal information about the user, the application or site demands a particular set of claims or a particular token type from the user. CardSpace then appears, taking over the display of the computer and represents the stored identities as virtual information cards. The user selects the card to use and the CardSpace software contacts the issuer of the identity to obtain a digitally signed XML token that contains the requested information.

CardSpace allows users to create self-issued identities for themselves, which can contain one or more of around 15 fields of telephone-book quality identity information. Other transactions may require a managed identity issued by a trusted identity provider, such as a bank, employer or a governmental agency.

Windows CardSpace is built on top of Web Services Protocol Stack, an open set of XML-based protocols, including WS-Security, WS-Trust, WS-MetadataExchange and WS-SecurityPolicy. This means that any technology or platform that supports WS-* protocols can integrate with CardSpace. In order to accept information cards, a website developer simply needs to declare an HTML <OBJECT> tag that specifies the claims the website is demanding from the user and then implement code to decrypt the returned token and extract the claim values. If an Identity Provider wants to issue tokens, they must provide a means by which a user can obtain a managed card and provide a Security Token Service (STS) that handles WS-Trust requests and returns an appropriate encrypted & signed token. If an IdP does not wish to build an STS, they will be able to obtain one from a variety of vendors including PingID, BMC, Sun or Microsoft, as well as other companies or organisations.

Because it is token-agnostic, CardSpace does not compete directly with other Internet identity architectures like OpenID and Liberty Alliance. In some ways the three approaches to identity can be seen as complementary.

In February 2006, IBM and Novell announced that they will support the Higgins trust framework to provide a development framework that subsumes support for the Web Services Protocol Stack underlying CardSpace within a broader, extensible support for various other identity-related technologies, such as SAML and OpenID. (Note that initial reports in the mainstream media positioning Higgins as a direct competitor for InfoCard were somewhat misconstrued.)

Microsoft included Windows CardSpace within its new Operating System Windows Vista. CardSpace is also available as part of Microsoft's .NET Framework 3.0 for Windows XP and Windows Server 2003.

3.1.5 OpenID

OpenID is different from the other identity frameworks discussed above, in the sense that it does not attempt to establish a trust metric. Instead, OpenID is user-centric and allows the user to establish an identity from an owned URL (for example a personal blog). A user will map their identity (the URL) to an Identity Provider of their choice, by adding a username and a reference to the IdP in the HTML meta-element on her webpage.

Different from a SAML SP, an OpenID consumer will not trust the IdP directly; neither will an OpenID consumer try to map the resolved identity to a real life person. Instead OpenID attempts to be a tool for the user to prove that they are the same person that created a service provider account (user profile). The service provider will benefit from a simpler registration mechanism and does not need to implement password storage and retrieval. Registration and login will also be simpler for the user, and the user will benefit from SSO functionality, and remembering only one username and password.

OpenID is pretty new, but up and coming. Several large companies have welcomed the framework and announced upcoming support for it in their products. A few of these are Microsoft, AOL and digg.com.

Many national identity management federations are considering ways to meet OpenID. Is OpenID a direct competitor to a SAML-based federation? Can OpenID be leveraged in the existing federations? How big is the overlap in functionality between a SAML federation and OpenID?

Answers may differ between federations, but after an open discussion within JRA5, some partners seem to welcome OpenID with open arms. On 5th March 2007 RedIRIS announced a plan for offering OpenID Identity Provider functionality within their federation. A week later, Feide demonstrated a proof of concept OpenID to SAML 2.0 bridge, which allows all Feide users to connect their OpenID identity to Feide as an OpenID IdP.

JRA5 will keep an eye on the “OpenID movement”, and if considered reasonable to do so, will look at the possibility of introducing it in the eduGAIN framework.

3.2 Authorisation Management Systems

3.2.1 X-RBAC

X-RBAC defines a policy specification language for access control. This language is RBAC environment-oriented. It defines several elements for protecting resources and allows definition of relationships between policies for multi-domain environments. X-RBAC defines a language for user, roles and permissions representation.

Once every element is defined, the model allows the establishment of relationships between them. It defines the relationship between users and roles (roles assigned to a user). It also defines the relationship between roles and permissions (permissions assigned to one specific role). Also, it suggests the set of states one role can go through (disable, enable, active).

The relationships between roles in multi-domains can be defined for two different scenarios: loosely coupled system multi-domain environments and federate multi-domain environments.

The first one defines the relationship between roles in environments where independent organisations or systems need to share information in a dynamic way during a period of time. There are two approaches: to use a predefined set of relationships between roles for the inter-domain access management, and to use certificates when a trust management infrastructure is necessary.

In a federated environment, usually one domain is designed as master and the others as local domains or slaves. The master manages the access to the rest of the system through a global policy that is associated to the different local policies. X-RBAC can define RBAC policies and mediation policies in a loosely coupled multi-domain system and federated multi-domain environment.

However, it presents some disadvantages, such as the fact that it does not implement a manageable frame for SSO environments and does not define a query-response protocol for authorisation messages.

3.2.2 ShARPe

ShARPe is a GUI tool for managing the creation and maintenance of user attributes for Shibboleth. It has been developed as part of the collaboration between MAMS (an Australian AAI project) and Shibboleth. The main aim is the management of Shibboleth's Attribute Release Policies. ShARPe allows administrator and users to manage their attributes in an easy way.

Sharpe is defined as an extension of Shibboleth providing other functionalities that are not present in Shibboleth. For example, it allows users to manage their own ARPs, the definition of policies for groups of users, and the mapping of attributes between different schemas.

The tool has three components: SharpeCore, WebSharpe, and Autograph.

- The central part of Sharpe is SharpeCore. These tools help with the support of the Identity Provider's configurations files.
- WebSharpe is an optional Web interface that allows users and administrator to interact with Sharpe.
- Autograph is another interface to manage user idCards. One user can have several idCards to be used with different Service Providers. Users can specify which idCard has to be used according to the service they want to obtain. If one user belongs to one university and they are also a company's employee, they will show the university idCard when using the university services (specifying their staff or student role, for example). On the other hand, when they want to access to the company's network they will display another idCard (specifying their position in the company). Each idCard will take into account different attributes.

3.2.3 Signet and Grouper

The Signet Privilege Management System was designed to address the challenges of managing what people can access and giving control of that process to those in the departments who make the decisions.

Because the implementation of access rules is typically scattered among many systems across an institution, coordinating policy and privileges campus-wide is difficult. This leads to individuals having levels of access that are inappropriate for their positions. This contributes to loss in their productivity and increase in security risks.

Signet holds information about protected resources and services that someone might access or use, such as purchasing or learning management systems. It then enables the resource owner to define the specific interactions (called privileges) with that resource, such as purchasing materials or updating grades for a homework assignment. Appropriate campus decision-makers assign or delegate these privileges to individuals, guided by policy rules coded into the system.

Signet separates the management of the policy and business rules from the technical system, so a change in system technology does not affect the access of those using it. Integrating Signet with an identity management system provides a way to manage the lifecycle of a person's access privileges, enabling automatic revocation of privileges based on status and affiliation changes. Enabling Role-Based Authority can be achieved by combining Signet and Grouper. While Signet manages the information associated with what people can do and supports delegated administration of privileges, it doesn't manage groups or membership of groups such as the list of department chairs or business majors. To help with this aspect of providing role-base authority, campuses can use the Grouper Group Toolkit that offers a consolidated way to automate or enable a delegated management of groups and related membership. Together, they provide a solution for role-based authority.

3.3 Standards

3.3.1 SAML 2.0

3.3.1.1 *A unified solution for identity federations*

SAML (Security Assertion Markup Language) is an OASIS http://www.oasis-open.org standard for the formation and exchange of authentication, attribute, and authorisation data using XML. SAML promotes interoperability between disparate security systems, by standardising how identity-related security information can be communicated between policy domains.

Until recently, federated identity has had to put up with the difficulty of having too many standards. Organisations looking to deploy federations had to negotiate protocol selection with each federation partner, having to deal with five protocols: OASIS SAML 1.0 and 1.1, Liberty Alliance ID-FF 1.1 and 1.2 and Shibboleth. Many had to support multiple protocols through mapping and translation techniques that slowed growth and increased the cost of federation deployments.

SAML 2.0 incorporates the critical use cases and features from previous protocols into a single standard, thus making its predecessors obsolete and removing the largest barrier to increased federation adoption: multi-protocol complexity.

3.3.1.2 *Added functionality*

SAML 2.0 represents a major revision of SAML 1.1 and provides significant additional functionality. It unifies the previous disparate federated identity building blocks of SAML 1.1 with input from both higher education's Shibboleth initiative and Liberty's Identity Federation Framework (Liberty ID-FF), and therefore constitutes a critical step towards full convergence for federated identity standards.

Another relevant new feature in SAML 2.0 refers to federated identifier management. Previous versions of SAML did not specify any mechanism by which two sites can agree upon a common identifier to use for a particular user, assuming that this would be established through some out of band exchange. This is covered in SAML 2.0 by defining how two sites can, with the participation of a user, establish one or more identifiers for that user in a dynamic online fashion, as well as manage identifiers once agreed upon.

Whereas its predecessors were not optimised for privacy, SAML 2.0 includes support for deployments in which privacy is a requirement. A relevant example is the definition of the format of a pseudonymous identifier as an opaque string by which two providers can refer to individual principals. User privacy is thus protected by avoiding collusion between multiple providers (as is possible with a global identifier such as an email address). Furthermore, SAML 2.0 includes mechanisms that allow providers to communicate privacy policies or settings from one to the other. The SAML assertions and protocols allow, for instance, communication of a principal's consent to some operation being performed, relevant to their privacy, from one provider where it was obtained to another.

With respect to session management, SAML 2.0 adds support for Single Logout, consisting of the ability to automatically terminate various different sessions established at different sites, based on the users' preferences and the policies of these sites.

3.3.1.3 *Current SAML 2.0 deployments*

OpenSAML

OpenSAML is an open-source toolkit, in Java and C++, produced by Internet2 developers as part of their work on the Shibboleth project. It is able to create objects with the individual information fields that make up a SAML message, build the correct SAML representation, and parse the SAML back into object form, as well as supporting developers implementing application using various SAML profiles and transport bindings.

OpenSAML 2, which is a deployment for the last version of the SAML protocol, is a complete rewrite of OpenSAML 1 built upon a generic library (XMLTooling) for working with XML as objects. OpenSAML is written by the developers of, and is part of, the Shibboleth development effort. Therefore OpenSAML 2 is tied to the release of Shibboleth 2.0, currently scheduled for early 2007.

OpenSSO

The Open Web Single Sign-On Project, also referred to as *OpenSSO*, is an open development effort based on the source code for Sun Java™ System Access Manager, a core identity infrastructure product offered by Sun Microsystems, Inc. The goal of OpenSSO is to provide an extensible foundation for an identity services infrastructure that will facilitate single sign-on (SSO) for web applications hosted on web and application servers in the public domain.

Sun Java System Access Manager and *Sun Java System Federation Manager* are two products designed to support secure and standards-based federation solutions and access to resources. Both of them implement SAML 2.0 among other standards. The next versions of these products will be built from the OpenSSO source.

As part of the OpenSSO project, Sun is developing *Open Federation*, which is an open source project based on the identity federation and web services framework developed for Sun Java System Access Manager and Sun Java System Federation Manager. Open Federation provides an extensible framework to support these features.

Another part of OpenSSO is a SAML 2.0 Service Provider (SP) implementation in PHP called the *Lightbulb*. Its goal is to provide a lightweight means of federating identities, implemented in a scripting language, which will allow users to sign in with a single authentication key and move seamlessly between a variety of interlinked and recombined web services projects. Currently the Lightbulb project is only available for PHP, but it is expected that implementations with Ruby and other scripting languages will also be available.

Lasso

Lasso is a free software C library developed by the French company Entr'ouvert, aiming to implement the Liberty Alliance standards. It defines processes for federated identities, single sign-on and related protocols and is licensed under the GNU General Public License.

Lasso first focused on implementing the Liberty Alliance ID-FF 1.2 protocols; these have been integrated into SAML 2.0, which is now the standard supported by Liberty Alliance. The newest version of Lasso (2.0) has been recently released, including full support of SAML 2.0.

ZXID

The ZXID project aims at providing a full stack implementation of all federated identity management and identity web services protocols. Its initial goal is supporting the SP role, followed by ID-WSF WSC and IdP roles.

ZXID currently includes a C program that implements a SAML 2.0 Service Provider as a CGI script (*zxid*), a C library for supporting SAML 2.0, including federated SSO (*libzxid*), as well as a Perl module, a PHP extension and a Java JNI extension that wrap libzxid. ZXID is free software (under the Apache License 2.0) and is developed by Sampo Kellomäki, co-founder of Symlabs.

Symlabs Federated Identity Access Manager

This free product created by Symlabs provides a multi-protocol federated identity solution encompassing ID-FF 1.2, ID-WSF 1.1, and WS-Federation in Identity Provider, Service Provider, and Enhanced Client roles. Symlabs complements this solution for Identity Management front end with its Directory Extender 3.0 product which addresses ID firewall and back end requirements. Symlabs Federated Identity Access Manager was certified for SAML 2.0 conformance for all available profiles.

3.3.1.4 *Integration into JRA5*

Currently the OASIS SAML 2.0 Metadata Specification is used in eduGAINMeta, which is part of eduGAIN – the confederated identity solution developed in work item 2 of JRA5. The eduGAINMeta API uses OpenSAML and allows dynamic metadata exchanges between eduGAIN federation components (BEs and FPPs) and the eduGAIN Metadata Service (MDS). The library was constructed directly using OpenSAML 2.0 code, in parallel with development efforts for the OpenSAML project.

The eduGAIN core, for which programming had started earlier based on the previous version of SAML, is awaiting the official release of OpenSAML 2.0 before migrating to SAML 2.0. Currently, this is planned to start in Summer 2007.

Concerning the eduGAIN Bridging Elements, there are currently different stages of SAML 2.0 compatibility, depending on the national federations to which they belong. The Norwegian AAI Feide went into production in June 2007 with a SAML 2.0 federation, including the Feide-eduGAIN BE, using the Sun Java System Access Manager and Federation Manager products. In Spain, work is in progress for a direct integration between PAPI and Sun's Access Manager, which will allow the PAPI federation (and BE) to use SAML 2.0 by 2007. Complete integration of this standard inside the federation will occur with the full alignment to Liberty Alliance standards, expected around the third quarter of this year. Federations using Shibboleth are also awaiting the official release of Shibboleth 2.0, consisting of a full implementation of SAML 2.0, currently scheduled for the first part of this year.

As eduGAIN migrates to SAML 2.0, direct connections between SPs and IdPs will become possible, thus putting less weight on the Bridging Elements as Local Federation Adaptors (LFAs). BEs will still be used for federations, which do not use SAML 2.0 internally, in order to perform the translation between local AAI protocols and SAML 2.0.

3.3.2 XACML

XACML (eXtensible Access Control Markup Language), the OASIS proposal for a standard access control language, was defined to represent access control policies in a standard way. XACML is XML-based and includes two different specifications:

- An access control policy language, which defines the set of subjects that can perform particular actions on a subset of resources.
- A representation format to encode access control requests and responses. This is a method for expressing queries about whether a particular access should be allowed and the related answers.

The following diagram depicts the general architecture of a XACML-based infrastructure:

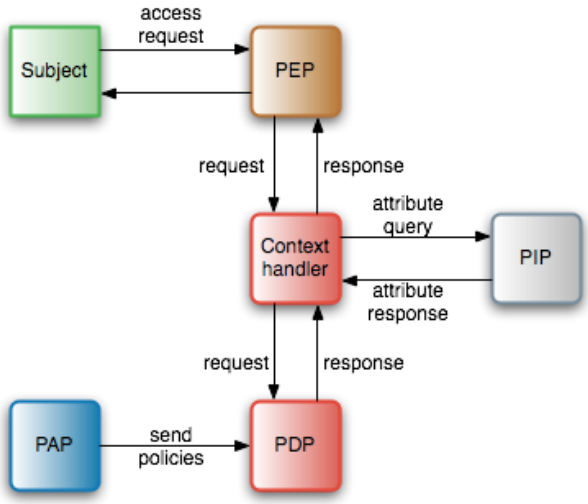

Figure 3: Architecture of a XACML-based infrastructure

Five main components are considered:

- **Policy Administration Point (PAP)**: This is a policy repository, providing them to the Policy Decision Point (PDP).
- **Policy Enforcement Point (PEP)**: This is the connection point between the resource, which requires an access control, and the whole XACML infrastructure. It has the purpose of receiving access requests and, using some components, returning a response message.

- **Policy Information Point (PIP)**: This component gets those attributes that have been required by other components, using external or internal sources.
- **Policy Decision Point (PDP)**: This component makes the authorisation decision using information provided by the rest of the components.
- **Context Handler (CH)**: This is the main interface between the application environment and the Policy Decision Point (PDP). It has to perform the conversion from the format used by that environment to the XACML-based message format.

3.3.2.1 *Context and policies*

A XACML request to the PDP has to contain the following elements:

- **Subject**: The actor that makes a request.
- **Resource**: The service or system component, which the subject wants to access.
- **Action**: Defining which kind of operation the subject requests from the resource.

The PDP has to check the request with its policies. As the following diagram shows, the main element of all XACML policies is a *Policy* or *PolicySet* element. A PolicySet is a container that can hold other Policies or PolicySets, as well as references to other policies (*PolicyIDReference*). A Policy represents a single access control policy, expressed by a set of *Rules*. A Policy or PolicySet may contain multiple policies or Rules, each of which may evaluate to different access control decisions. XACML needs some way of reconciling the decisions each makes, and this is done through a collection of *Combining Algorithms*. An example of those is the *Permit Overrides Algorithm*, which says that if at least an evaluation returns Permit, then the final result is also Permit. A Policy or PolicySet element may also specify a set of *Obligation Attributes*, that is, a set of actions to be performed if the request has been approved.

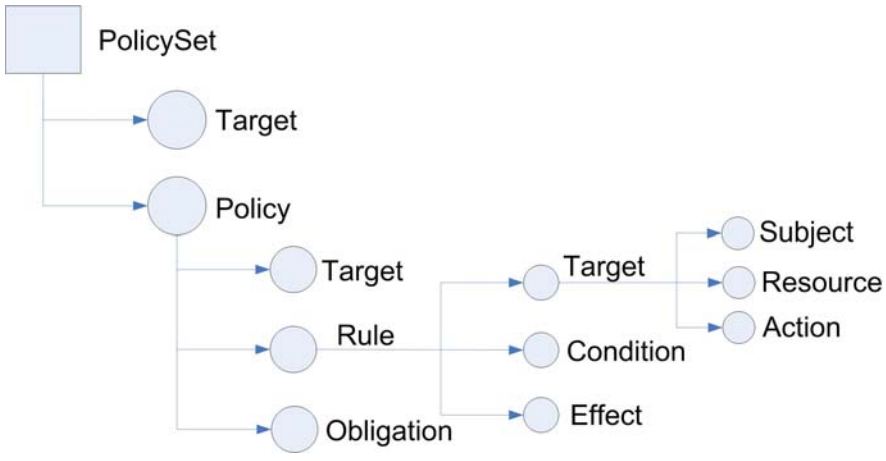

Figure 4: PolicySet of an XACML Policy

XACML provides another feature called *Target*. It is a set of simplified conditions for the Subject, Resource and Action that must be met for a PolicySet, Policy or Rule to apply to a given request. If all the conditions of a

Target are met, then its associated PolicySet, Policy, or Rule is applied to the request. Once a Policy is found, the included rules are evaluated. The main element of a rule is the *Condition*. If the evaluation of the Condition results true, then the Rule's Effect (Permit or Deny) is returned.

The main object that XACML deals with is attributes. Attributes are named values of known types. Specifically, attributes are characteristics of the Subject, Resource, Action or Environment in which the access request is made.

3.3.2.2 *Profiles*

Profiles apply XACML in concrete use cases in a specific environment or to a specific problem that has to be solved. This is a brief overview of the most relevant ones.

- **SAML 2.0 profile**: As described above, PEP has to send a XACML request to the PDP, though XACML does not mandate any particular protocol for that purpose. By means of this profile, XACML-based messages can be transferred inside SAML 2.0 messages.
- **Core and hierarchical role based access control (RBAC) profile**: Most applications use role-based authorisation and access control system rather than authorise individual subjects. This profile defines the usage of XACML to meet the requirements in the ANSI RBAC 'core' and 'hierarchical' standards.
- **Hierarchical resource profile**: This profile is used when rights for resource access are organised in a hierarchy. Being able to express a single policy constraint that will apply to an entire sub-tree of *nodes* in the hierarchy, rather than having to specify a separate constraint for each *node*, increases both the ease of use and the likelihood that the policy will correctly reflect the desired *access controls*. Another special characteristic of *hierarchical resources* is that access to one *node* may depend on the value of another *node.*
- **Multiple resource profile**: This profile describes different ways in which a PEP can request authorisation decisions for multiple resources in a single request context, and how the result of each is represented in the single response context that is returned to the PEP. Also, it describes how a PEP can request a single authorisation decision in response to a request for all the nodes in a hierarchy.

3.3.3 Resource Description Framework (RDF)

Resource Description Framework (RDF) is a family of World Wide Web Consortium (W3C) specifications, originally designed as a metadata model using XML but which has come to be used as a general method of modelling knowledge, through a variety of syntax formats (XML and non-XML).

The RDF metadata model is based upon the idea of making statements about resources in the form of subject-predicate-object expressions, called triples in RDF terminology. The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object. For example, one way to represent the notion "The sky has the colour blue" in RDF is as a triple of specially formatted strings: a subject denoting "the sky", a predicate denoting "has the colour", and an object denoting "blue".

This mechanism for describing resources is a major component in what is proposed by the W3C's Semantic Web activity: an evolutionary stage of the World Wide Web in which automated software can store, exchange, and use machine-readable information distributed throughout the web, in turn enabling users to deal with the information with greater efficiency and certainty. RDF's simple data model and ability to model disparate, abstract concepts has also led to its increasing use in knowledge management applications unrelated to Semantic Web activity.

3.3.3.1 *Ontologies*

A collection of RDF statements intrinsically represents a labelled, directed pseudo-graph. As such, an RDF-based data model is more naturally suited to certain kinds of knowledge representation than the relational model and other ontological models traditionally used in computing today. However, in practice, RDF data is often stored in relational database representations sometimes also called triple stores, and as RDFS and OWL demonstrates, additional ontology languages can be built upon RDF.

3.3.3.2 *Resource identification*

The subject of an RDF statement is a resource, possibly as named by a Uniform Resource Identifier (URI). Some resources are unnamed and are called blank nodes (bnodes) or anonymous nodes. They are not directly identifiable. The predicate is a resource as well, representing a relationship. The object is a resource or a Unicode string literal.

In Semantic Web applications, and in relatively popular applications of RDF like RSS and FOAF (Friend of a Friend), resources tend to be represented by URIs that intentionally denote actual, accessible data on the World Wide Web. But RDF, in general, is not limited to the description of Internet-based resources. In fact, the URI that names a resource does not have to be dereferenceable at all. For example, a URI that begins with "http:" and is used as the subject of an RDF statement does not necessarily have to represent a resource that is accessible via HTTP, nor does it need to represent a tangible, network-accessible resource.

Therefore, it is necessary for producers and consumers of RDF statements to be in agreement on the semantics of resource identifiers. Such agreement is not inherent to RDF itself, although there are some controlled vocabularies in common use, such as Dublin Core Metadata, which is partially mapped to a URI space for use in RDF.

3.3.3.3 *Statement reification and context*

The body of knowledge modelled by a collection of statements may be subjected to reification, in which each statement (that is each triple subject-predicate-object altogether) is assigned a URI and treated as a resource about which additional statements can be made, as in "Jane says that John is the author of document X". Reification is sometimes important in order to deduce a level of confidence or degree of usefulness for each statement.

In a reified RDF database, each original statement, being a resource itself, usually has at least three additional statements made about it: one to assert that its subject is some resource, one to assert that its predicate is some resource, and one to assert that its object is some resource or literal. More statements about the original statement may also exist, depending on the application's needs.

Borrowing from concepts available in logic (and as illustrated in graphical notations, such as conceptual graphs and topic maps), some RDF model implementations acknowledge that it is sometimes useful to group statements according to different criteria, called situations, contexts, or scopes, as discussed in articles by RDF specification co-editor Graham Klyne. For example, a statement can be associated with a context, named by a URI, in order to assert an "is true in" relationship. As another example, it is sometimes convenient to group statements by their source, which can be identified by a URI, such as the URI of a particular RDF/XML document. Then, when updates are made to the source, corresponding statements can be changed in the model, as well.

Implementation of scopes does not necessarily require fully reified statements. Some implementations allow a single scope identifier to be associated with a statement that has not been assigned a URI itself.

In first-order logic, as facilitated by RDF without scopes, the only meta-level relation is negation, but the ability to generally state propositions about nested contexts allows RDF to comprise a meta-language that can be used to define modal and higher-order logic.

3.3.3.4 *Query and inference languages*

Several query languages for RDF graphs have emerged. RDF query languages allow expressions to be written that can be evaluated against a collection of statements in order to produce, for example, a narrowed set of statements, resources, or object values, or to perform comparisons and operations on such items. RDF queries can be used by knowledge management applications as a basis for inference actions.

Modelled loosely after SQL, the query language SPARQL is emerging as the de-facto RDF query language. On the track towards status of W3C Recommendation, it was released as a Candidate Recommendation in April 2006, but is back to Working Draft status since October 2006, due to open issues.

Other notable RDF query and inference languages include:

- RDQL, precursor to SPARQL, SQL-like
- Versa, compact syntax (non–SQL-like), solely implemented in 4Suite (Python)
- XUL has a template element in which to declare rules for matching data in RDF. XUL uses RDF extensively for data binding.

3.3.3.5 *Applications*

- RDF Site Summary; One of several "RSS" languages for publishing information about updates made to a web page; it is often used for disseminating news article summaries and sharing weblog content.

- FOAF (Friend of a Friend): Designed to describe people, their interests and interconnections.
- DOAC (Description of a Career): Supplements FOAF to allow the sharing of résumé information.
- DOAP (Description of a Project): Designed to describe software projects; uses FOAF to identify the people involved.
- SIOC (Semantically-Interlinked Online Communities): Designed to describe online communities and to create connections between Internet-based discussions from message boards, weblogs and mailing lists.
- MusicBrainz: Publishes information about Music Albums.
- Creative Commons: Uses RDF to embed license information in web pages and mp3 files.
- Haystack: Semantic web browser from MIT CS & AI lab.
- Many other RDF schemas are available by searching SchemaWeb
- Microsoft shipped a product called Connected Services Framework, which provides an RDF based Profile Management capabilities.

3.4 Evaluation from the JRA5 perspective

eduGAIN is intended to accommodate changes to the technologies it uses without affecting the user applications. The definition of the eduGAIN interfaces in terms of services provided through abstract operations allow for both a seamless migration of the protocols and profiles that the eduGAIN components use, and an extensions of the facilities offered by them. After the first pilot of the eduGAIN infrastructure is demonstrated and initial deployments start, a transition to the SAML 2.0 protocol set at the core of the infrastructure is planned, without requiring any change from the participating federations or the relaying parties. This transition will provide almost automatic compatibility with Shibboleth 2.

Furthermore, the future evolution of participating federations in order to incorporate new technologies and standards (as could be the case of Signet, Grouper or ShARPe) is not likely to be interfering with eduGAIN components, since the BE concept guarantees a decoupling between the tightly-coupled internal federation processes and the looser trans-federation links provided by eduGAIN. The BE concept provides a simple way for integrating federations based in different technologies, and for providing service to future relaying parties.

The new advent of the so-called *user-centric identity management* (CardSpace, Higgins, …) opens new ways to explore not only the integration of infrastructures based on this concept in eduGAIN, but also the possibility of making eduGAIN itself evolve in this direction. This is also applicable to the use of Semantic Web technologies in component location and identification, and trust link building.

4 SSO related technologies

For JRA5, the term Single Sign On (SSO) includes both sides of the problem; allowing the users to provide a single set of credentials not only to access the network but also to access the applications. The issue is thus extended from allowing access “at anytime from anywhere” to allowing access “at anytime from and to anywhere”. This is one of the goals of the activity and therefore there is a direct participation of JRA5 in the evolution of the related technologies, which essentially pursue the integration of the existing infrastructures for network and application access.

4.1 NAS-SAML

Network access systems have traditionally been based on login/password mechanisms, and other systems, following a more advanced approach for mutual authentication, are based on X.509 identity certificates. These systems are especially useful for organisations, which are concerned with the real identity of the requester. There are other organisations where the different users are classified according to different attributes, such as their administrative tasks, the type of service obtained, or some other internal requirements. In those previous scenarios, the user's identity wasn’t sufficient for granting access to the controlled resource, since additional attributes needed to be known describing, e.g., the role being assigned to the user in order to offer the right service. Therefore, those scenarios need a system able to grant the set of attributes specifying those privileges to the different users. This is the main aim of NAS-SAML.

NAS-SAML is a network access control approach based on the use of an AAA infrastructure to manage the user authentication process and authorisation attributes. This proposal is based on the SAML and XACML standards, which are used for expressing access control policies based on attributes, authorisation statements, and authorisation protocols.

In a NAS-SAML application scenario, as shown in the figure below, every end user belongs to a home domain, where they are given a set of attributes (for example the roles they play or additional information useful for the access control scenario). When the user requests a network connection in a particular domain by means of an 802.1X connection, the request is captured by an AAA server located in the target domain, which then makes a query to obtain the attributes linked to the user from an authority responsible for managing them, located in the user's home domain. Alternatively, following a push approach, the user can present their attributes instead of letting the AAA server recover them. The communication between domains is carried out using the DIAMETER protocol. Finally, the AAA server sends an authorisation decision query to a local PDP (Policy Decision Point)

entity, and that element provides an answer indicating whether the attributes satisfy the resource access control policy. Furthermore, that policy can also establish the set of obligations derived from that decision, for example some QoS parameters, security options, etc. This general scheme works both in single and inter-domain scenarios.

The main components of NAS-SAML are the following:

- *End User*: Entity requesting access to the network. They will try to gain access to a network making use of their attributes defined in the home domain. The user must have a valid identity (valid authentication credentials such as X.509 authentication certificate or login/password).
- *AAA Server*: It is responsible for receiving and processing the authentication, authorisation and accounting information related to the users trying to get access to the network. It is also responsible for querying other entities in order to retrieve any additional information that could be needed in order to perform its task. It is based on DIAMETER technology.

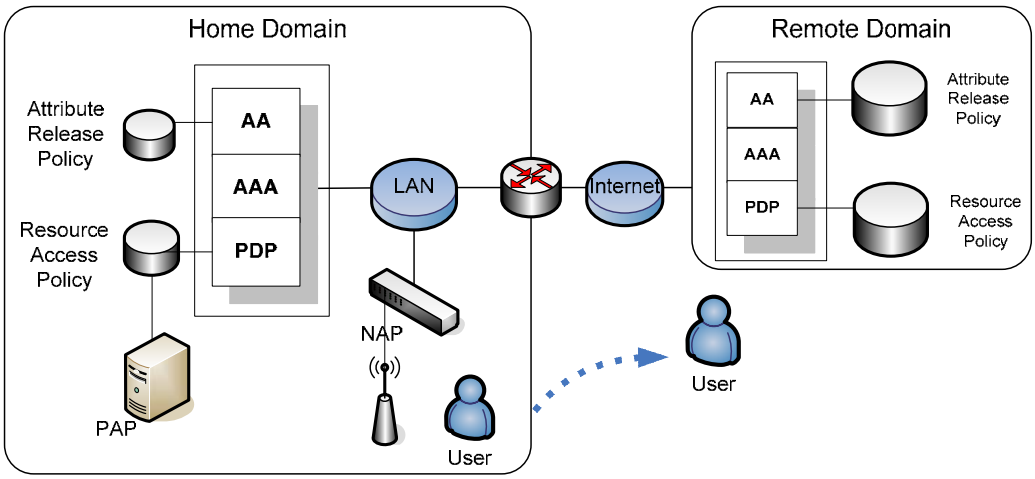

Figure 5: NAS-SAML Components

- *Attribute Authority (AA)*: The AA determines the set of attributes that can be released to a requester in an access control process. The AA has its own Attribute Release Policy, expressed in XACML.
- *Policy Decision Point (PDP)*: Entity responsible for generating the statements related to authorisation decisions based on the user attributes. It has its own Resource Access Policy, expressed in XACML.
- *Policy Administration Point (PAP)*: This entity defines, signs, and publishes the resource access policy.
- *Network Access Point (NAP)*: This element has two main functions: first, it has to forward the client requests to the appropriate AAA server of the remote domain; second, once the AAA has obtained the authorisation decision, the NAP obtains and enforces the properties of the network connection.

Interactions among the different components described above depend on the requirements imposed by the user to gain access to the network. On one hand, the end user can follow a pull approach, which requires the minimum overload and is more suitable for limited terminals, such as PDAs or mobile phones. In this way, all the authorisation tasks are performed by the network infrastructure. On the other hand, following a push model,

the user can present a particular set of attributes to the system, following their preferred disclosure policy. The push model involves a certain support for selecting and transporting attributes from the end user terminal, representing a more intrusive approach. In consequence, NAS-SAML provides solutions to these two different environments.

4.2 Internet2 proposal for SAML integration in RADIUS

A SAML profile for integrating network access based on RADIUS is currently being discussed under the auspices of MACE. The proposal is to continue using the RADIUS protocol for user authentication (as it is currently in use by the eduroam infrastructure), have the visited site use a SAML profile to obtain attributes describing the user, and have the visited site make an access control decision using the SAML attributes. The combination of both technologies poses some issues that must be addressed:

- Minimise the modifications to both the RADIUS and the SAML messages.
- Minimise the number of new RADIUS attributes added to the RADIUS messages.
- Minimise additional steps, which could create delays in the RADIUS interactions.
- Consider that the RADIUS protocol does not provide end-to-end message integrity.
- The intermediate Radius Proxy Servers can see all of the Radius Attribute/Value pairs contained in a message, and could potentially edit them.
- Support must be provided for both SAML 1.1 and SAML 2.0

The currently proposed architecture is shown in the following diagram:

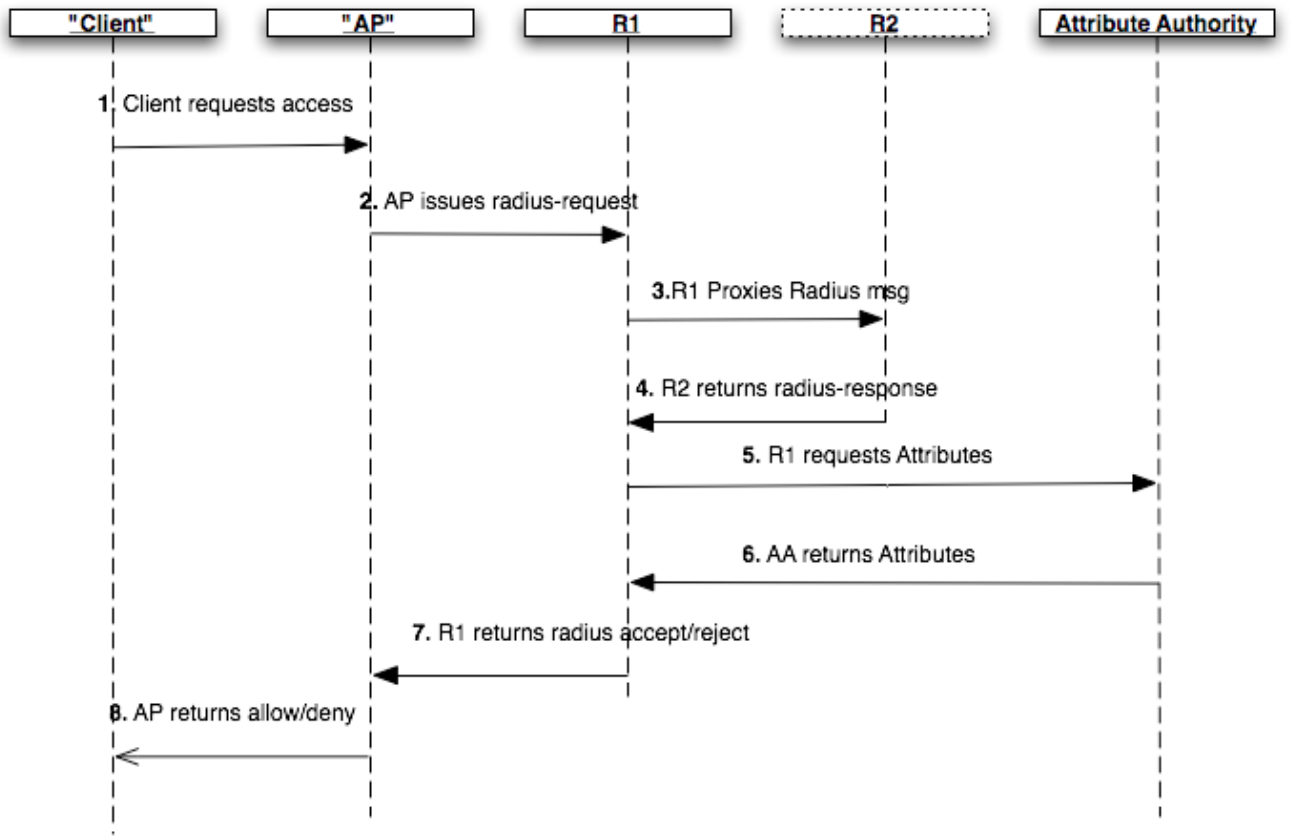

Figure 6: Proposed architecture for SAML integration into RADIUS

The elements considered are:

- **Access Point (AP)**: An entity that participates in an 802.1x message exchange with a client requesting access to a local network, and initiates a RADIUS protocol message exchange with a local RADIUS server in order to authenticate and authorise the requesting user. No changes are being proposed to the messages that APs receive or send.
- **R1 (RADIUS Server)**: R1 is a RADIUS server at the Network Provider site that also functions as a SAML-defined Attribute Requester. It accepts RADIUS protocol messages from a local AP, and optionally proxies them to remote RADIUS servers. It is able to engage in SAML interactions and to make access decisions based on the results of these interactions.
- **R2 (RADIUS Server)**: R2 is a RADIUS Server at the User Home site. It accepts RADIUS protocol messages proxied from R1, and responds with appropriate RADIUS protocol messages. It is able to provide data about the appropriate SAML Attribute Authority that can be used for additional queries.
- **Attribute Authority**: This entity functions according to the Shibboleth SAML profile, processing SAML queries and returning a SAML response message, possibly including one or more assertions containing SAML attributes.

4.3 Evaluation from the JRA5 perspective

Work on uSSO for JRA5 is on-going in the DAMe project. The main goal of this project is to define a unified authentication and authorization system for federated services hosted in the eduroam network. Those federated services can range from network access control to distributed services like Grid Computing. The eduroam infrastructure will be extended to use the NAS-SAML architecture, so that user mobility will be controlled by security assertions and policies expressed in standard and extensible languages, such as SAML and XACML. The NAS-SAML and XACML research work is therefore of great value to JRA5.

5 Comprehensive technologies

This section discusses the Internet Protocol version 6 (IPv6), the next-generation packet transport for the Internet, and its influence within JRA5 work items, mostly the Roaming work item. After a short introduction to the new protocol's features, the current state of deployment in NRENs is examined and recommendations are given.

5.1 IPv6

5.1.1 IPv6 in general

The IPv6 protocol is a re-design of the IP protocol version 4, which is currently deployed worldwide. Its new features are too numerous to be listed, but some of the major ones with a probable impact on eduroam deployment are:

- The larger address space (128 bit vs. 32 bit) – leading to the obsolescence of Network Address Translation (NAT) and enabling true end-to-end connection to end-user devices;
- Auto-configuration – easier client configuration without the need for stateful address mapping tables;
- Extensive support for IP layer mobility – enables users to keep their home IP address even when moving around.

These new features come at a cost, however. Equipment and vendor support, routing performance on current equipment, configuration details and network operator staff training on the new protocol's internals are major obstacles to IPv6 deployment.

5.1.2 Current Status of IPv6 developments and deployments

The majority of all NRENs in GÉANT2 have an IPv6 enabled backbone. However, traffic levels of IPv6 compared to IPv4 are still negligible. According to the TERENA Compendium 2006 [COMP], p. 57, a mere 1.3% of NREN traffic in the first quarter of 2006 is IPv6, which is even a decline from the previous quarters. The same source also reports that the number of universities with an IPv6 uplink has grown from 15.4% to 34.8%

within one year. This is seemingly a contradiction – more entities get connected while the traffic drops. The likely explanation is that few universities connect their whole campuses, but instead only prepare their core network. One of the reasons for the low usage level of IPv6 is probably that the network of the connected entities is not extended up to the edge, but only used within the network core. However, coverage needs to be extended up to the end-user devices to make people use the protocol. This means that each and every router in the chain from backbone to edge access routers needs to be IPv6-enabled.

Unfortunately, providing end users with IPv6 connectivity requires introducing a solution for the problem of supplying the end-user device with a DNS server. In IPv4, DNS server configuration can be delivered piggyback-style with the device's address configuration via the Dynamic Host Configuration Protocol (DHCP). In IPv6, there is no need for address configuration and so other options need to be considered. Now, two approaches are in use to supply devices with the DNS resolution service:

- DHCPv6 – The pendant of IPv4's DHCP. DHCPv6 clients do not come built-in into most operating systems, so this option often requires deploying a software package to the users in addition to the DHCPv6 server, which needs to be present in the network. One popular DHCPv6 server and multi-platform client software package is Dibbler [DIB]. The Internet System Consortium (ISC) [ISC] is also working on a DHCPv6 client and server, which are supposed to be released in the coming.

Operating System	Client	Server
Windows Longhorn	built-in	built-in
Windows Vista	built-in	-
Windows XP	separate software available (*)	separate software available
MacOS X	- (possibly KAME port)	-
Solaris	built-in (OpenSolaris Nevada)	-
Linux	software available, may be built-in (distribution-specific)	software available, may be built-in (distribution-specific)

(*) Dibbler currently has some problems with wireless devices

- Fixed Unicast or Anycast DNS – This solution means configuring the client device to send its DNS resolution queries to a well-known IPv6 address. That way, no further configuration or protocol is required. Clients can always use this well-known address and can expect answers to their DNS queries to be delivered from a unicast DNS server or an arbitrary DNS server that belongs to the anycast cloud. The client-side configuration is a one-time effort for every client machine. This solution works well only for a static environment where users do not roam: the DNS address resides within the prefix of the network. If a user roams into a completely different network, it depends on the home DNS server's configuration whether or not a DNS resolution request is answered. Since open DNS servers have been used for numerous DoS amplification attacks, opening a DNS server for world access is generally considered harmful and a roaming hotspot can not rely on a visitor having a pre-configured working DNS server.
- Site-local anycast or multicast addresses: This approach uses a well-known anycast address that is globally identical for DNS resolution. That way, users can configure their device once with a static address and use it at any location. For example, Windows XP by default uses the IPv6 addresses

fec0:0:0:ffff::1 as primary DNS server. However, RFC 3879 has deprecated site-local addresses and using this concept is considered bad practice. Using multicast DNSv6 is possible, but the operations are not yet defined (pending work items at the IETF zeroconf and dnsext working groups) and not in a widespread deployment. Further information on Multicast DNS can be found at [MDNS].

A good overview of the implementation status of IPv6 in various operating systems as well as separate software packages is available at [NECINE]. A comprehensive view of the existing DHCPv6 server and client implementations is provided by the IST project 6net in Deliverable D3.2.3v3 [6net].

5.1.3 MIPv6

5.1.3.1 *Background*

Mobile IPv6 (MIPv6) [RFC3775] is an enhancement to IPv6 that provides mobility support. MIPv6 is a Proposed Internet Standard that has been developed by the IETF's mip6 working group [MIP6C].

MIPv6 allows a mobile node (MN) to cross an IP network boundary without disrupting higher-level protocols such as TCP or UDP. Consequently, applications continue to work as the underlying network media or addressing changes. It allows a node to be addressed by the same IP address while visiting other networks than its Home network.

5.1.3.2 *Brief introduction to MIPv6*

To operate as a MIPv6 node, the mobile node registers with a router, called the Home Agent (HA), on his home network. While in the home network the MN behaves normally. If the MN moves to another network it uses a MIPv6 Binding Update (BU) message to instruct the HA to create a binding between the MN's Home Address and the MN's address on the visited network; this address is known as the Care-of Address (COA) and may be allocated through DHCPv6 or auto-configuration.

When a Correspondent Node (CN) then wishes to contact the MN, it sends IP packets to the Home Address. If the MN is away from its home network, the HA uses a proxy Neighbour Discovery (ND) mechanism to intercept the IP packets and then forward the data to the MN's COA.

To avoid triangular routing, MIPv6 supports Route Optimisation (RO) whereby the MN sends a BU to the CN; however, this requires that the CN supports MIPv6. Authentication of BU messages between the MN and HA may be implemented using the Return Routability Test (RRT).

It is not necessary to modify applications for them to operate over MIPv6, assuming the application already supports IPv6. However, the MN's operating system must support MIPv6.

The following report is recommended to readers unfamiliar with IPv6/MIPv6:

- "IPv6 Deployment Guide" by Martin Dunmore (6DISS, 2005 [DEPLG])

MN and HA support for MIPv6 is currently supported in a number of Unix systems and derivates, such as Linux [MIPv6H]. Microsoft have stated that MN MIPv6 support may be provided within Windows Vista when it ships, or more likely, at a future point as part of a software update [MSMIPv6] (Windows XP and 2003 currently support a pre-RFC version of MIPv6). Cisco supports HA MIPv6 in recent versions of IOS [ISOMv6], although IPSec [RFC3776] is not supported [IOSIPSEC]. MIPv6 has also been demonstrated on Nokia devices running System 40 and 60 [NOKMv6] although this has not yet been deployed in commercial products.

5.1.3.3 *Potential issues*

The "bootstrapping problem"

Within MIPv6, a MN requires at least a home address, a home agent address and a security association with that home agent. The process of setting up these prerequisites is called bootstrapping. The bootstrapping process is potentially vulnerable to attack or disruption:

- A bogus home agent could access or disrupt the MN's traffic.
- A bogus MN could obtain or disrupt roaming service if able to authenticate to the HA.
- Man-in-the-middle attacks between MN and HA could access traffic or steal credentials.
- Denial of service vulnerabilities.

Any successful deployment should mitigate these risks.

The bootstrapping problem is discussed in detail in [BSTR].

Firewalls

Some SPI firewalls break MIPv6 signalling due to asymmetric routing. Also, certain known issues exist with firewall traversal, as documented in [MIAO]. Firewalls are a common feature on most networks, and consequently it will be necessary to identify if and how MIPv6 can co-exist with commonly deployed firewalls.

5.1.4 Recommended Actions

5.1.4.1 *eduroam hotspots*

eduroam is advancing towards a global service. It has already expanded to regions where usage of IPv6 is a lot more common than in Europe (e.g. in the Asia-pacific region). It can be expected that some eduroam users expect an IPv6 connection to be present at their roaming location. Because of that, IPv6 connectivity should be added to eduroam hotspots wherever possible. The hotspot should offer both the IPv4 and IPv6 networking stack.

Deploying IPv6 on new hotspots should be carefully planned and supervised. While a lot of equipment claims to fully support IPv6, practical evidence has shown that throughput may be limited or processor loads on switching

and routing equipment may be significantly higher than for IPv4. Because of that, any deployment of IPv6 at the network edge should be thoroughly stress-tested and resource utilisation along the routing paths should be monitored.

Whenever a hotspot is IPv6-enabled, a DHCPv6 server with the appropriate DNS server information should be present. Firewall rules for IPv6 should allow arbitrary destination addresses for DNS queries for the case that a user tries to use his own (home) DNS server.

In general, eduroam resource providers should always advocate the use of DHCPv6 for their own users so that a seamless roaming experience is guaranteed. Letting the users manually configure a DNS server for their eduroam connection is strongly discouraged.

5.1.4.2 *eduroam Infrastructure and Authentication Servers*

The eduroam infrastructure is currently mostly based on IPv4. Even though it is possible to move to IPv6 on several RADIUS implementations today (Radiator, Navis, RadSec proxy) not all implementations are IPv6-enabled yet (FreeRADIUS). Additionally, since the worldwide IPv6 infrastructure is considerably less mature and less resilient (fewer routes) than the existing IPv4 deployment, it is unlikely that a move to IPv6 for eduroam authentication and proxy servers is necessary or prudent within the GN2 project lifetime.

It is recommended to

1. Move to IPv6 on a server connection only if the implementation permits and the move does not lead to a higher risk of service outage

2. Stay with the existing connections otherwise

3. Conduct tests with IPv6 server interconnections to get a better understanding of code and infrastructure maturity within the GN2 lifetime.

5.1.4.3 *Test Runs for RADIUS/RadSec/Diameter over IPv6*

A later deliverable of Work item 1 (Roaming) should examine the feasibility and stability of using IPv6 in the eduroam infrastructure. This deliverable should also provide an overview about the number of eduroam hotspots, which offer IPv6, as well as the operational experience with IPv6 at these hotspots.

5.2 Evaluation from the JRA5 perspective

There is not an immediate need for deploying IPv6 connectivity for eduroam in Europe. However, given the truly international character of eduroam and a user base in regions where IPv6 use is more common, deployment of IPv6 on eduroam hotspots should be seriously considered by eduroam resource providers, and if possible realised in the near future.

The aforementioned DNS server resolution problem leads to an additional configuration burden, and on most operating systems also a user support burden for eduroam Resource Providers: Since incoming guest users may have either a DHCPv6 client configured or a statically configured anycast DNS, both of these solutions should be present at an eduroam hotspot. This means that the hotspot administrator needs to have a DHCPv6 server implementation, as well as configure his network infrastructure to allow DNS queries to destinations outside the own network. Since a well-behaved home DNS server will not answer the user's queries when they are roaming, user experience on the roaming network may be inferior compared to IPv4.

Mobile IPv6 is currently not available in production-quality implementations. Furthermore, Mobile IPv6 will only have a significant benefit to end-users if it is ubiquitous at the eduroam hotspots. It is not expected that this will happen within the GN2 lifetime. Considering Mobile IPv6 deployment may be a goal for a later follow-up project of GN2.

6 Value-added services

The applications mentioned below are considered useful extensions in the wireless network environment and provide a certain incentive to be used on top of eduroam.

6.1 SLP

The Service Location Protocol (SLP, srvloc) is a service discovery protocol that allows computers and other devices to find services in a local area network without prior configuration. SLP has been designed to expand from small, unmanaged networks to large enterprise networks. It has been defined in RFC 2608 as Standards Track document.

6.1.1 Logical overview

Devices use SLP to announce services on a local network. Each service must have a URL that is used to locate the service. Additionally it can have an unlimited number of name/value pairs, called attributes. Each device must always be in one or more scopes. Scopes are simple strings and are used to group services, comparable to the network neighbourhood in other systems. A device cannot see services that are in different scopes.

For example, the URL of a printer could be:

service:printer:lpr://myprinter/myqueue

This URL describes a queue called "myqueue" on a printer with the host name "myprinter". The protocol used by the printer is LPR. Note that the printer uses a special URL scheme “service:”. "service:" URLs are not required, any URL scheme can be used, but they allow you to search for all services of the same type (e.g. all printers) regardless of the protocol that they use. The first three components of the "service:" URL type ("service:printer:lpr") are also called service type. The first two components ("service:printer") are called abstract service type.

A service template, a formalised description of the URL syntax and the attributes, specifies the definition of a “service:” URL and the allowed attributes for the URL. Service templates are defined in RFC 2609.

SLP allows several query types to locate services and obtain information about them:

- It can search for all services with the same service type or abstract service type.
- The query can be combined with a query for attributes, using LDAP's query language.
- Given its URL, the attributes of a service can be requested. In standard SLP the attributes are not returned in the query result and must be fetched separately. The Attribute List Extension (RFC 3059) fixes this problem.
- A list of all service types can be obtained;
- A list of all existing scopes can be requested.

6.1.2 SLP roles

SLP has three different roles for devices. A device can also have two or all three roles at the same time:

- User Agents (UA) are devices that search for services.
- Service Agents (SA) are devices that announce one or more services.
- Directory Agents (DA) are devices that cache services. They are used in larger networks to reduce the amount of traffic and allow SLP to scale. The existence of DAs in a network is optional, but if a DA is present, UAs and SAs are required to use it instead of communicating directly.

Today most implementations are daemons that can act both as UA and SA. Usually they can be configured to become a DA as well.

6.1.3 Network protocol

SLP is a packet-oriented protocol. Most packets are transmitted using UDP, but TCP can also be used for the transmission of longer packets. Because of the potential unreliability of UDP, SLP repeats all multicasts several times in increasing intervals until an answer has been received. All devices are required to listen on port 427 for UDP packets, SAs and DAs should also listen for TCP on the same port. Multicasting is used extensively by SLP, especially by devices that join a network and need to find other devices.

The operation of SLP differs considerably, depending on whether a Directory Agent (DA) is in the network or not. When a client first joins a network, it multicasts a query for DAs on the network. If no DA answers it assumes it is in a network without DAs. It is also possible to add DAs later, as they multicast a “heartbeat” packet in a predefined interval that will be received by all other devices. When a SA discovers a DA, it is required to register all services at the DA. When a service disappears the SA should notify the DA and deregister it.

In order to send a query in a network without a DA, the UA sends a multicast UDP packet that contains the query. All SAs that contain matches will send a UDP answer to the UA. If the answer is too large to fit into a single UDP packet, the packet will be marked as "overflown" and the UA is free to send the query directly to the SA using TCP, which can transmit packets of any size.

In order to send a query in a network with a DA, the UA will send the query packet to the DA using either UDP or TCP. As every SA must register all services with the DA, the DA is able to fulfil the request completely and simply sends the result back to the UA.

6.1.4 Security

SLP contains a public-key cryptography based security mechanism that allows signing of service announcements. In practice it is rarely used:

- The public keys of every service provider must be installed on every UA. This requirement defeats the original purpose of SLP, being able to locate services without prior configuration.
- Protecting only the services is not enough. Service URLs contain host names or IP addresses, and in a local network it is almost impossible to prevent IP or DNS spoofing. Thus only guaranteeing the authenticity of the URL is not enough if any device can respond to the address.
- As addresses can be spoofed, the authenticity of the device must be proved at a different level anyway, e.g. in the application protocol (e.g. with SSL) or in the packet layer (IPSec). Doing it additionally in SLP does not provide additional security.

6.1.5 Adoption

- SLP is frequently used for locating printers and supported by printing systems such as CUPS.
- SLP is often found in LAN-enabled printers, so that they are discoverable out of the box. Some client print drivers can use this for printer discovery.
- ACN, a protocol being developed for entertainment control, uses SLP to find different devices such as dimmers and intelligent lights.
- Mac OS and Mac OS X up to version 10.1 used SLP to locate file shares and other services. Newer releases of Mac OS X, versions 10.2 onwards, prefer Zeroconf for new features though.
- Novell NetWare clients in a pure IP environment use SLP to locate servers.
- SUSE Linux supports SLP for a variety of services since version 9.1

6.2 SIP/IP Telephony

The two major competing standards for signalling in VoIP or IP telephony, are the IETF standard SIP and the ITU standard H.323.

Both SIP and H.323 may be used for signalling in VoIP applications, although their original focus is very different. While SIP has been designed as a generic transaction protocol for session initiation not bound to any specific media such as audio or video, the focus of H.323 has been to handle voice and multimedia calls, including supplementary telephony services. New product developments, (both in the public and private telecommunications) and VoIP service deployments (also peering/interconnection) use SIP as the default

signalling protocol. Interworking between SIP and H.323 is therefore relevant for those existing carrier VoIP networks and for migrations in multi-site corporate networks and for networks with a mix of these technologies.

Internet protocols and Ethernet LANs were designed primarily for relatively bursty, asynchronous data traffic. As a result, the latency and jitter associated with the transmission of data packets have not been major issues, because they rarely create a noticeable impact in applications like e-mail and Web browsing. Voice transmission, in contrast, requires a steady stream of audio packets to satisfy the user. Excessive packet loss and jitter will result in voice quality degradation with noticeable gaps in the audio stream.

6.3 IP Telephony + eduroam

As the popularity of VoIP (Voice over IP) grows, and PSTN (public switched telephone network) users switch to VoIP in increasing numbers, governments are becoming more interested in regulating VoIP in a manner similar to legacy PSTN services. VoIP is today more than a technology for transport of voice over packet switched IP networks. With emerging standard protocols, like SIP and H.323, that support signalling-over-IP, voice-over-IP has evolved to an alternative IP telephony system.

Meanwhile eduroam is reaching a mature stage and the usage of Wireless LAN (WLAN) technology is taking off due to the mobility and greater flexibility to access the campus network without being connected to the network by a wire.

The marriage of Wireless LAN and IP telephony technologies is Voice over WLAN (VoWLAN), which is a natural extension of VoIP that enables voice communications throughout a corporate campus or other facilities served by eduroam Wireless LAN.

Expectations for mobility are on the rise. Eduroam users who already enjoy wireless e-mail and wireless intranet access are ready to enjoy that access through mobile devices that enable real-time communications. A significant benefit of mixing telephone traffic with data on a WLAN is to provide mobility and make use of a common infrastructure. The support of a common system for both data and voice traffic is generally simpler and less expensive than two separate entities.

Organisations are moving to adopt VoWLAN solutions to enhance communication while leveraging their existing investment in WLAN and VoIP infrastructures. To ensure that those moves are successful, there must be solutions that can provide the necessary reliability, enterprise-class security, and QoS requirements to ensure even the most demanding applications, like wireless voice and video, run smoothly. But while voice and video traffic need preferential treatment, the organisations will benefit by using a single, comprehensive management system for its entire network—voice and data, wireless and wired—to reduce the time and operational cost to manage and provide consistent service to employees regardless of where they are in the network or how they are connecting to it.

6.3.1 Voice over eduroam Challenges

6.3.1.1 *The Bandwidth Challenge*

Usually WLANs provide less bandwidth than wired LANs. Mobile devices must share the bandwidth provided by access points, instead of getting their own dedicated, switched, wired LAN connection.

The ability to add infrastructure to overcome bandwidth limitations is limited by the potential for co-channel interference between neighbouring access points and mobile devices.

6.3.1.2 *The QoS Challenge*

Carrying voice and video over WLANs imposes additional requirements. Voice and video are real-time in nature, and highly latency-sensitive. Network delays that might be imperceptible when surfing the Web can significantly affect voice quality.

To provide high-quality voice, there will be a need to implement QoS over the network, prioritising voice traffic so that it doesn't suffer delays and other problems.

Other open issues such as to guarantee QoS for handover devices, to achieve QoS with minimal power consumption, and QoS–based radio resource management are challenging topics.

6.3.1.3 *The Security Challenge*

Security issues in IP telephony are different and in ways more complex than security for data applications. IP telephony is a complex application involving multiple layers of the protocol stack, requiring interoperability among different new and legacy protocols, and interactions among multiple network elements.

Existing vulnerabilities include eavesdropping, connection hijacking, call fraud, and denial-of-service These take on new forms in a converged network. Voice over wireless may create additional vulnerabilities. Detailed vulnerability analysis of the protocols both in isolation as well as in conjunction with other protocols that they interact with in the converged networks is required to develop appropriate countermeasures. Moreover, VoIP networks are prone to virus and worm spreading through their data network elements. While some solutions have already been proposed for vulnerabilities in protocols such as SIP, H.323 and privacy standards like WEP, more needs to be done. These include encryption of signalling messages and media to address man-in-the-middle attacks. Session Border Controllers are gaining acceptance as voice perimeter controllers in contrast to firewalls for data services.

6.3.1.4 *Mobility Management*

In IEEE 802.11 based Wireless Networks (like eduroam), when the Mobile Station changes its attachment point to the Internet, it may experience an interruption in service, since the Mobile Station cannot send or receive

packets from the time at which it disconnects from one attachment point to the time at which it registers with the new attachment point. Such an interruption would be unacceptable for real-time services such as Voice-over-IP, which demands further optimisation of the mobility management in IEEE 802.11 based wireless LAN.

The challenges of VoWLAN mobility management are to maintain the same QoS, and to ensure security. It requires transferring QoS and security contexts between attachment points and achieving the inter access point handover.

6.4 Virtual organisations

The reason for our interest in VOs is that we expect them to provide a useful framework for coordinated resource sharing in a confederation context, as has already been demonstrated in the grid context.

Note that much of the material for this discussion, is from a paper by Lyle J. Winton[2].

Virtual organisations (VOs) or virtual teams are in no way new concepts and they have been studied for example from organisational and sociological perspectives. In this document we though have a narrow technical focus – how can sharing of eScience resources between a number of institutions in a confederation be supported by the VO concept? For our purposes, we see the virtual organisation as an interesting alternative to setting up a physical organisation (or to reorganise). With the VO approach one may still achieve many of the expected advantages of a physical reorganisation. There may be a variety of reasons to go for the virtual rather than the physical alternative including:

- It makes a quick adaptation to changing needs within an organisation possible.
- It offers a way of organising cooperation between organisations.
- It makes it possible to set up a task force of experts from different geographic locations, with complementing competences etcetera.

There are two points of VOs that might be considered:

- The VO is expected to provide a service infrastructure supporting the everyday work of its members.
- A VO may have eScience resources, such as Grid resources, at its disposal to distribute among its members.

The first point amounts to confederating VO services. Even though an analysis of which everyday VO supporting services to provide is interesting in itself, from a confederation point of view this does not add anything to the eScience use case.

[2] Lyle J. Winton: A Simple Virtual Organisation Model and Practical Implementation, Proceedings of the 2005 Australasian workshop on Grid computing and e-research - Volume 44, Newcastle, New South Wales, Australia, 57 – 65, 2005, ISBN ~ ISSN:1445-1336 , 1-920-68226-0

The interesting question to focus on for JRA5 is: How does the concept of a VO fit in with the confederation concept introduced in JRA5? This question is not fully answered in this document, but will be the result further studies, if this technology is among those chosen.

In this section we first characterise the VO concept and the expectations it is reasonable to have on the framework that supports it, then discuss the Project concept as a use case of confederated authorisation. The section is concluded with a short discussion of confederated authorisation of the operation of an eScience resource and an evaluation of how well eduGAIN can be used for confederated authorisation in the eScience context.

6.4.1 Characteristics of a Virtual Organisation Framework

A VO has members, and distributes available resources to those members. Potential members of a VO are individuals, (segments of) organisations, and other VOs. Although the VO may have resources of its own, it is more often the case that the VO has blocks or other fractions of resources from resource providers at its disposal to distribute among its members. The VO shall have a policy for governance of the VO, membership, and so on, and a policy for how the resources at its disposal shall be distributed among its members.

In addition to these properties (which can be said to define a VO) there are additional requirements on the framework. The Framework manages the VOs and the information about them, and authorises operations on available resources.

Existing Frameworks are certainly very useful, but our analysis shows that they are rather immature and do not satisfy the requirements expected of a VO Framework. We believe reasonable requirements include the following:

- The Framework shall in a scaleable fashion be able to provide VO management services for a large number of VOs.
- Any member of a VO may be a member of any other VOs.
- The Framework shall allow a resource provider to delegate blocks or other fractions of resources to one or more VOs.
- The authorisation policies should be expressed in a formal language.
- It should be possible to base authorisation decisions on VO group and role information.

Finally, there are also security and availability requirements on the VO framework. These requirements are critical, since the delegated resources are protected by the VO framework and the framework must be up and running if the resource is to be available to the VO members. These issues are not discussed further here.

6.4.2 A Scenario for a Virtual Organisation – The Project

The “Project” as an organisational form is an example of a “Virtual Organisation”. In this section we use an eScience project as a concrete use case for the requirements of a VO on eduGAIN.

So let us assume that a project is characterised by the following properties:

- The project is collaboration between organisations, and these collaborating organisations are the partners of the project.
- Each partner is a member of a federation in the confederation.
- Each partner decides which individuals from their organisation are project members.
- The project has a certain “amount” of eScience resources in the confederation at its disposal.
- The eScience resources are provided by resource providers.
- Each resource provider is a member of a federation in the confederation.
- The resource provider decides the policy for allocating eScience resources to the project.
- The project decides the policy for allocating eScience resources to the project partners.
- The partner decides the policy for allocating eScience resources to its project members.

This means that a request by a project member to operate on an eScience resource is authorised by the partner it belongs to, by the project, and by the provider of the eScience resource. The project member may have presented information to all of these confederation components to influence the authorisation decisions. There are concrete approaches to VO frameworks in the Grid middleware packages. We briefly describe one approach (VO Membership Service) in the following section.

6.4.3 VO Membership Service

An existing VO framework is VO Membership Service (VOMS), developed by the European Datagrid Project. It consists of three components:

- A VO server, containing information about individual assertions of user credentials to present to the resource provider as a basis for authorisation.
- A management component.
- A client for generation of additional user credentials.

Credential assertions must be issued because in the VOMS model, authorisation is not delegated to the VO but carried out by the resource provider.

Winton mentions two deficiencies in the VOMS framework:

- The authorisation is not delegated to the VO.
- There is no common IdP associated with the components of the framework, implying that trust anchors of involved entities have to be managed outside the framework.

Another deficiency may be that when VOs are mentioned in connection with VOMS, it is always a rather small number of VOs (10 to 20). We expect every project to be a VO. This means we are talking about thousands of VOs to be managed.

The interactions involved in the use of a VOMS include the following:

- The potential members of a VO shall have a personal X.509 certificate issued by a CA in a PKI trusted by the VOMS.
- The user requests to become a member of a VO managed by the VOMS. A user can be a member of several VOs.
- Once accepted as a member of a VO, user roles for that VO and other relevant information about the user's relation to the VO can be registered in VOMS by its administrators.
- The VOMS issues, on request, authorisation assertions in the form of short-lived certificates, so called proxy certificates, to be presented to the resource providers.

In the VOMS model, all the attributes needed in a proxy certificate must be available in the VOMS VO store. Some of the attributes are probably general identity related attributes, where the original source is the enterprise directory of the user's home organisation, while others are specific to the user's relation to a particular VO. The approach in the VOMS is that all attributes are maintained explicitly. This is difficult to maintain with quality. Efforts to reduce the maintenance burden are certainly welcome. As an example, in the VASH project (VOMS Attributes From Shibboleth), an effort from SWITCH within the EGEE-2 project, a tool for supporting and simplifying this activity has been developed. The responsibility to reinitiate the transfer of attributes remains as a manual procedure. In the long run an event driven solution, where a component may subscribe to attribute updates, would further reduce the maintenance burden.

6.4.4 Confederation Aspects of VOs

What place shall the Virtual Organisation Service have in a confederated AAI like eduGAIN? It certainly is an architectural component in GÉANT2 space (as are Identity Provider and Metadata Service), but is it a component in the confederation or in the federation superstructure? Or is it a new type of member in federations? And what are the other required characteristics of a Project instantiation of a VO (as a GÉANT2 architecture component)?

When choosing to represent a project as a Virtual Organisation, the following factors need to be considered:

- A project is represented as a VO in the confederation.
- Property of a VO is information about project partners, members, available resources, authorisation policies and other relevant information.
- The authority to manage each of these VO attributes shall reflect the authority to control the relation between the VO and the resource, the project and the partner respectively.
- There are several technical approaches to storing, managing and retrieving these relations – the chosen solution must be trusted by all involved.
- Each VO has a unique identity in the confederation.

- The properties of a particular VO are managed by some VO Service, which in turn is hosted by a member of a federation in the confederation and thus has a physical location.
- The function of a VO in the confederation does not depend on the location of the VO, as long as it can be located.
- The information about projects is relevant to at least eScience resource providers and to partners, implying that not only resource providers and partners but also VOs are Attribute providers in the confederation architecture.

6.4.5 Conclusions on Virtual Organisations

This was a brief introduction to Virtual Organisations, and an attempt at identifying how the VO concept fits in with our approach to confederated authentication and authorisation. There remain a lot of questions to answer before we know how VOs can be integrated with eduGAIN:

- Does eduGAIN as it is designed support the general VO model sketched above?
- What shall we require from a VO service system?
- Shall we require a centralised or distributed VO service architecture?
- How can the VOMS model be improved to take advantage of campus AA infrastructures?
- How do we need to improve the campus AA infrastructures to make it use of campus ID Architecture in the eScience context?

6.5 Evaluation from the JRA5 perspective

SLP could be useful for guests visiting a remote institution. It can be made available after authentication of the user to offer services beyond mere network access. It is recommended to follow that item on a federation level, but not highly prioritised.

VoIP offers a great benefit for eduroam users. As eduroam reaches a mature stage it is only a matter of time when the challenges mentioned in the IP telephony chapter will be solved so that the eduroam users can make use of mobile telephony without having to rely on commercial cell phone carriers.

The VO concept seems to offer some interesting aspects for the NREN community, as mentioned above a lot of questions still need to be answered before knowing how to integrate it with eduGAIN.

7 Conclusions

This deliverable outlines a number of recent developments in the wider research area of JRA5.

For the roaming work item, RadSec will contribute to the improvement of the current eduroam infrastructure. This is already covered in the work plan (see Transition to eduroam service [DJ5.0.1]). Research and collection of practical experience is going on in parallel regarding Diameter. WiMAX is outlined only, as it does not pose any special integration problem.

The interoperation with mobile networks is an interesting area, open for further investigation. The EAP-SIM method used in that context can open up a way for integration with eduroam. In the AAI area, the provision of SAML 2.0 will be of severe importance. The transition to the SAML 2.0 protocol set at the core of the eduGAIN infrastructure is intended, without requiring any change from the participating federations or the relaying parties. This transition will provide almost automatic compatibility with Shibboleth 2.

The further development and implementation of Liberty Alliance's ID-FF shall be closely monitored by JRA5, with the potential of integrating new features into eduGAIN in mind. The newly developed Microsoft CardSpace represents a proprietary AAI solution, which should also be taken into account as it is incorporated into the new operating system Windows VISTA. CardSpace and semantic web are technologies that might influence the eduGAIN development and should be kept within view.

Supporting IPv6 in the eduroam and eduGAIN infrastructures remains an important topic. Provision of IPv6 to the end user devices including DNS is still not very common, which makes applicability in the near future unlikely. IPv6 should most likely stay on the agenda, but without immediate need for operation. Mobile IPv6 (MIPv6) poses further challenges because of insufficient implementation level. Following the experimental work in Ukerna on MIPv6 is the only feasible recommendation at this point.

IP Telephony is seen as a very attractive application based on eduroam. It is assumed that this will work without complications via the eduroam infrastructure. This needs to be checked in a practical set-up.

All other technologies may have an impact in the long term and should be monitored accordingly.

8 References

[6net]	http://www.6net.org/publications/deliverables/D3.2.3v3.pdf
[BSTR]	http://www.ietf.org/internet-drafts/draft-ietf-mip6-bootstrap-ps-05.txt
[COMP]	http://www.terena.org/activities/compendium/2006/compendium06_web.pdf
[DELPG]	http://www.6diss.org/publications/manuals/deployment-guide.pdf
[DIB]	http://klub.com.pl/dhcpv6/
[IOSIPSEC]	http://www.cisco.com/en/US/products/sw/iosswrel/ps5187/ products_configuration_guide_chapter09186a00804160bf.html#wp1161180
[IOSMv6]	http://www.cisco.com/en/US/products/sw/iosswrel/ps5187/ products_configuration_guide_chapter09186a00804160bf.html
[ISC]	The Internet Systems Consortium. http://www.isc.org/
[MDNS]	http://www.multicastdns.org/
[MIP6C]	http://www.ietf.org/html.charters/mip6-charter.html
[MIPv6H]	http://www.tldp.org/HOWTO/Mobile-IPv6-HOWTO/
[MSMIPv6]	http://www.microsoft.com/technet/community/columns/cableguy/cg0904.mspx#EPH
[MIAO]	http://www.ietf.org/internet-drafts/draft-miao-mip6-ft-02.txt
[NECINE]	IPv6 Resource. http://internecine.eu/
[NOKMv6]	http://press.nokia.com/PR/200411/969183_5.html
[RFC2581]	ftp://ftp.rfc-editor.org/in-notes/rfc2581.txt
[RFC3775]	http://www.ietf.org/rfc/rfc3775.txt
[RFC3776]	http://www.ietf.org/rfc/rfc3776.txt
[WLSWP]	http://www.cisco.com/en/US/products/hw/wireless/ps430/ products_white_paper09186a00800b469f.shtml

9 Acronyms

In JRA5 used acronyms can be found in the JRA5 Glossary of Terms [GN2DJ511]. Additional terms are listed below.

[4Suite] [4Suite is a platform for XML processing and knowledge-management, consisting of a library of integrated tools for XML processing, and an XML data repository and server with a rules-based engine]

[ANSI] [The American National Standards Institute or ANSI is a private nonprofit organization that oversees the development of voluntary consensus standards for products, services, processes, systems, and personnel in the United States.]

[AVP] [Attribute Value Pairs]

[CGI] [The Common Gateway Interface, see RFC 3875, ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc3875.txt.pdf]

[CS] [Circuit Switched]

[DNS SRV] [A DNS RR for specifying the location of services, see RFC 2782, http://tools.ietf.org/html/rfc2782]

[HSDPA] [High-Speed Downlink Packet Access (HSDPA) (Sometimes known as High-Speed Downlink Protocol Access) is a 3G mobile telephony protocol in the HSPA family, which provides a roadmap for UMTS-based networks to increase their data transfer speeds and capacity.]

[ID-WSF WSC] [ID-WSF Web Services Consumer libraries that enable developers to incorporate SAML 2.0 functionality into a variety of Web services and client-based applications.]

[IIS] [Microsoft Internet Information Services (IIS; formerly called Server) is a set of Internet-based services for servers using Microsoft Windows.]

[ISIM] [An IP Multimedia Services Identity Module (ISIM) is an application running on a UICC smart card in a 3G mobile telephone in the IP Multimedia Subsystem (IMS).]

[OASIS] [a global consortium that develops data representation standards for use in computer software.]

[ODBC] [Open Database Connectivity (ODBC) provides a standard software API method for using database management systems (DBMS). The designers of ODBC aimed to make it independent of programming languages, database systems, and operating systems.]

[OWL] [The Web Ontology Language (OWL) is a language for defining and instantiating Web ontology. An OWL ontology may include descriptions of classes, along with their related properties and instances. OWL is designed for use by applications that need to process the content of information instead of just presenting information to humans.]

[PANA] [Protocol for Carrying Authentication for Network Access. PANA is an IP-based protocol that allows a device to authenticate itself with the network to be granted network access.]

[PHP]	[PHP is a reflective programming language originally designed for producing dynamic Web pages and remote application software. PHP is used mainly in server-side scripting, but can be used from a command line interface or in standalone graphical applications.]
[PS]	[Packet Switched]
[PSTN]	[The public switched telephone network.]
[RBAC]	[Role Based Access Control]
[RDFS]	[RDF Schema is an extensible knowledge representation language, providing basic elements for the description of ontologies, otherwise called RDF vocabularies, intended to structure RDF resources.]
[RDQL]	[RDF Query Language, which has been implemented in a number of RDF systems for extracting information form RDF graphs.]
[RSS]	[RSS is a family of web feed formats used to publish frequently updated digital content, such as blogs, news feeds or podcasts.]
[SLPv2]	[Service Location Protocol, Version 2, see RFC 3224, ftp://ftp.rfc-editor.org/in-notes/rfc3224.txt]
[SPARQL]	[SPARQL Protocol and RDF Query Language, is an RDF query language that allows for a query to consist of triple patterns, conjunctions, disjunctions, and optional patterns.]
[SPI Firewall]	[A stateful firewall (any firewall that performs **s**tateful **p**acket **i**nspection or stateful inspection) is a firewall that keeps track of the state of network connections (such as TCP streams, UDP communication) travelling across it. The firewall is programmed to distinguish legitimate packets for different types of connections. Only packets matching a known connection state will be allowed by the firewall; others will be rejected]
[SS7]	[Signalling System #7 (SS7) is a set of telephony signalling protocols which are used to set up the vast majority of the world's public switched telephone network telephone calls.]
[USIM]	[A Universal Subscriber Identity Module is an application for UMTS mobile telephony running on a UICC smart card which is inserted in a 3G mobile phone.]
[Versa]	[Versa is a RDF query language, used to query RDF data. Versa differs from most other RDF query languages, which are typically based on SQL or specialised XML vocabularies.]
[WS-*]	[Web Services Standards]
[WSRP]	[Web Services for Remote Portlets (WSRP) is a technology agnostic protocol designed for accessing remote Portlets in a standard manner. The WSRP specification defines a web-service interface for interacting with interactive presentation-oriented web services]
[XUL]	[The XML User Interface Language, is an XML user interface markup language developed by the Mozilla project.]