Bandwidth Allocation and Reservation - End-to-End Specification

Kavoussanakis, K
(EPCC, The University of Edinburgh, UK) *et al*

22 May 2006

EGEE is a project funded by the European Commission
Contract number INFSO-RI-508833

The electronic version of this EGEE Technical Reports is available
on the CERN Document Server at the following URL:
<http://cdsweb.cern.ch/search.py?p=EGEE-TR-2006-004>

EGEE

BANDWIDTH ALLOCATION AND RESERVATION END-TO-END SPECIFICATION

Document identifier:	**EGEE-TR-733663-BAR_EndtoEnd_specification-v1-0.doc**
Date:	**18/05/2006**
Activity:	**JRA4: Development of Network Services**
Document status:	**DRAFT**
Document link:	**https://edms.cern.ch/document/733663/1**

Abstract: The Bandwidth Allocation and Reservation (BAR) activity within JRA4 of the EGEE project specified and implemented the necessary components and interfaces to enable the EGEE Grid middleware to request and use guaranteed bandwidth services. This report describes the components and interfaces required for an end-to-end BAR service and how they interact.

1. INTRODUCTION

As Grid performance is inherently dependent on underlying network performance, over-utilisation of networks due to rerouting, denial of service attack or simply an unusual increase in load may result in poor Grid performance. For certain Grid applications and critical Grid operations, poor performance may be unacceptable. In such cases, enhanced network services can ensure that network performance requirements are met.

These enhanced network services, offered by network providers including the European National Research and Education Networks (NRENs) and the pan-European network GÉANT2, allow the reservation of network paths between a specified set of sites over a period of time, with guaranteed Quality of Service (QoS) characteristics. Initially, these services are implemented by the network providers over IP networks using Premium IP [R1], [R2]. In the future, as European networks evolve, these services may be provided using technologies such as Point-to-Point Ethernet (a guaranteed capacity service provisioned between two Ethernet ports) and Lightpaths (a point-to-point optically-provisioned service with an associated service level specification).

The Bandwidth Allocation and Reservation (BAR) activity within JRA4 of the EGEE project has specified and implemented the necessary components and interfaces to enable the EGEE Grid middleware to request and use these enhanced network services.

[R3] discusses the requirements that a BAR service must meet, based on those specified jointly with the EGEE SA2 activity [R4]. It also includes a state-of-the-art survey of existing resource management systems (extensively discussed in [R5]) and, more importantly, specifies the interface between the four high-level components involved in BAR. These components are as follows:

1. Higher Layer Middleware (HLM), the Grid middleware components that make use of the BAR service.
2. Network Service Access Point (NSAP), a component that resides within the networks and is outside the scope of EGEE. Service Activity 3 (SA3) of the European GÉANT2 project has developed AMPS (Advanced Multi-domain Provisioning System), which is an NSAP implementation that enables inter-domain Premium IP reservations.
3. Local Network Service Access Point (L-NSAP), required to configure the local networks involved in a guaranteed bandwidth reservation.
4. BAR, a Web service developed by JRA4, provides an interface suitable for the HLM to make reservations, coordinating with NSAP and L-NSAP to make lower-level network reservations.

Requirements for four BAR network services are identified. The Guaranteed Deadline File Transfer (GDFT) service requires the BAR component to determine the appropriate capacity and service duration that is required to guarantee that a file with a certain size can be transferred before a given deadline, allowing capacity to be traded off against service duration whilst adhering to the given deadline and overall capacity limits. The Virtual Leased Line (VLL) service provides a guaranteed bandwidth between a certain start and end time. The two other services, Video and Visualisation, cannot be supported by the Premium IP based bandwidth reservation currently available in GÉANT2.

This report should be considered *the definitive BAR architecture document*. It specifies BAR as JRA4 sees it, within the context of EGEE-I. It details the interactions of the BAR software components and lists their functional specification. While the architecture has been influenced by the circumstances at GÉANT2 and also the requirements of the EGEE middleware, we believe it to be autonomous and general enough to be applicable to any network service.

The document is organised as follows. In chapter 2 the BAR architecture is presented in detail. The different components and interfaces are identified. The specification of these components is presented in chapter 3, while the interfaces are detailed in chapter 4. Note that the interfaces, which are covered in [R3], are only summarised in this document. Finally, chapter 5 presents the current status of BAR.

2. ARCHITECTURE

2.1. BAR ARCHITECTURE OUTLINE

Analysis of Grid applications and middleware requirements [R4] leads us to define a single Web service called the "BAR Web service". This acts as a single point of access within a site to allow reservation of various guaranteed bandwidth services. The BAR Web service handles user authentication and authorisation, and provides an interface that is simple to understand and use. This insulates users from the detailed network-level specifics and multiple network domain concerns involved in a guaranteed bandwidth reservation.

Figure 1 shows how the BAR service interacts with HLM and the network services.

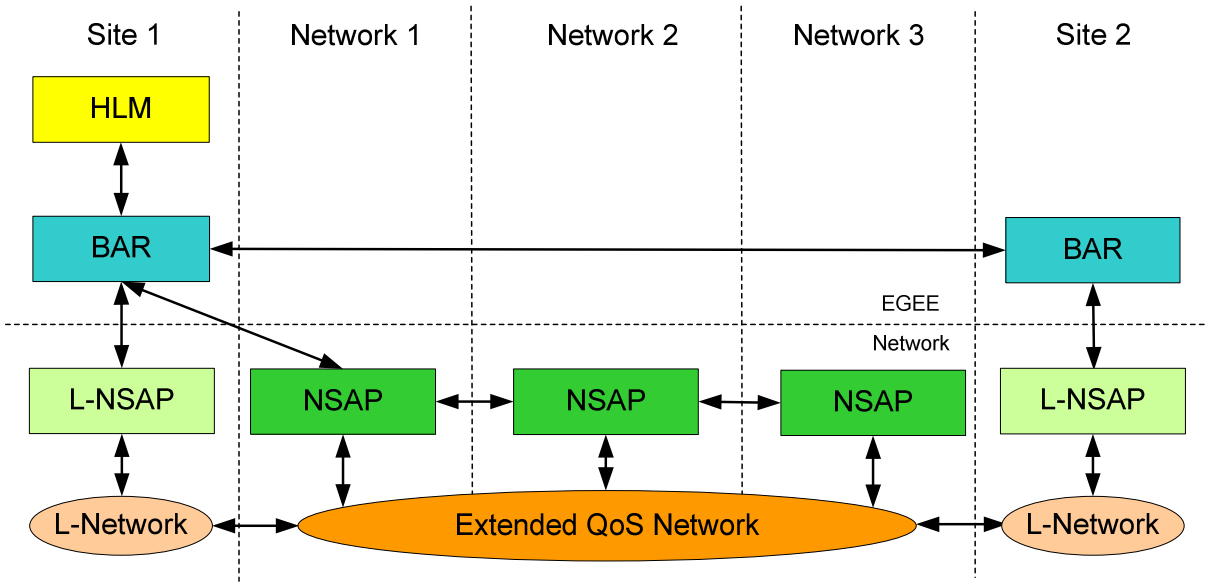

Figure 1: BAR service architecture

Figure 1 shows how the BAR Web service fits into a multi-domain networking environment. We call Grid applications and middleware Higher Layer Middleware (HLM). As shown in the figure, HLM at a site (Site 1) contacts the BAR Web service at that site to reserve guaranteed bandwidth from an end-point at that site to an end-point at another site (Site 2). The BAR Web service then contacts the network services and, if a bidirectional reservation is required, the Remote BAR Web service (i.e., the BAR service at Site 2) to carry out the reservation. Note that the BAR Web service Grid middleware component is owned by EGEE, whereas the network providers' services (on Networks 1-3 on Figure 1) are not.

There are two types of network service agents, the NSAP and the L-NSAP. The Network Service Access Point (NSAP) is a Web service that assesses and carries out the reservation request. We assume NSAPs to be present in transport networks, such as GÉANT2 and NRENs, thus forming an *extended QoS-enabled network*[1]. NSAPs are concerned with the configuration of network equipment on the backbone. As explained in [R3], the NSAP abstracts the network-specific services but still speaks a language that is network-oriented. The Local NSAP (L-NSAP) is of equivalent functionality

[1] The term backbone in the rest of this document is assumed to mean "extended QoS-enabled network".

to the NSAP but is concerned with the configuration of equipment on the local network of the source and destination sites. This component is used to configure the local network. Further details of how a reservation is established and the request flow between the components in Figure 1 are provided in section 3.

The architecture above adheres to the basic assumption in our work, that EGEE is an end-user for the network providers. As such, it can drive the agreement of interfaces and act as early adopter of the services, but it cannot define the components that will run on the network; thus the horizontal line separating EGEE and the network domain. The HLM and BAR entities belong to the EGEE administrative domain, and as such they are developed by EGEE. Network service providers are free to implement NSAPs as they see fit. The only constraint is that the **standard** HLM-BAR, BAR-NSAP and BAR-L-NSAP interfaces defined by EGEE and GÉANT2 **must be** honoured.

JRA4 uses the GÉANT2-NSAP (AMPS) in our pilot service. L-NSAPs belong to the administrative domain of the end-sites and the resulting diversity of local equipment and policies complicates matters. JRA4 only investigated reference implementations of L-NSAP [R7].

Throughout this document we may use the qualification *Local* for the BAR and L-NSAP instances on Site 1; we always use the qualification *Remote* for the BAR and L-NSAP instances on Site 2 (the differentiator being that the HLM contacts the Site 1 BAR).

2.2. NETWORK SERVICES EXPOSED BY BAR

Based on the user requirements and the underlying service available by GÉANT2, BAR exposes the following two network services:

1. Guaranteed Deadline File Transfer (GDFT) Service

   The GDFT service provides network transport between two servers guaranteeing the transfer of a given volume of data within a given period of time. This implies that only an **average**[2] bandwidth is guaranteed while the actual bandwidth may vary over time. The GDFT service may be used for instance as part of a Grid data replication service.

2. Virtual Leased Line (VLL) Service

   The VLL service provides a dedicated amount of capacity (i.e., a **sustained** bandwidth) in the network between two specified set of machines over a period of time. Note that the actual bandwidth will remain constant. This service can be used for instance to guarantee the timely delivery of service messages.

2.3. TWO STAGE PROVISIONING PROCESS

Networks impose a minimum reservation period (in the case of GÉANT2 this is two weeks) as a result of the fact that configuration is currently manual. Even if the reservation process is fully automated, a minimum reservation period is reasonable to impose so as to minimise the amount of configuration needed in the extended QoS network. This reservation period may be significantly larger than the requirement of the applications and it is desirable to allow subdivision of such reservations among nodes and users in a subnet, for more efficient and effective use of the service. In addition, typical Grid jobs can only specify the IP flow parameters (exact source and IP addresses and optionally port numbers, bespoke DSCPs etc.) just before a job starts, while the point of BAR is to reserve the traffic in advance.

Therefore, EGEE SA2 and JRA4 have come to the conclusion that a two stage (Service Reservation/Service Activation) provisioning model is required (see Figure 2). Another constraint

[2] Although this is not currently implemented, BAR could replace the request with a sequence of different bandwidth requests so as to accommodate the file transfer.

addressed by this model is that of the lead time between receipt of the request and effect of the reservation, which again is likely to not fully disappear even if the process gets automated. Assuming that the Service Activation has less of a lead time than Service Reservation, the model is more user-friendly.

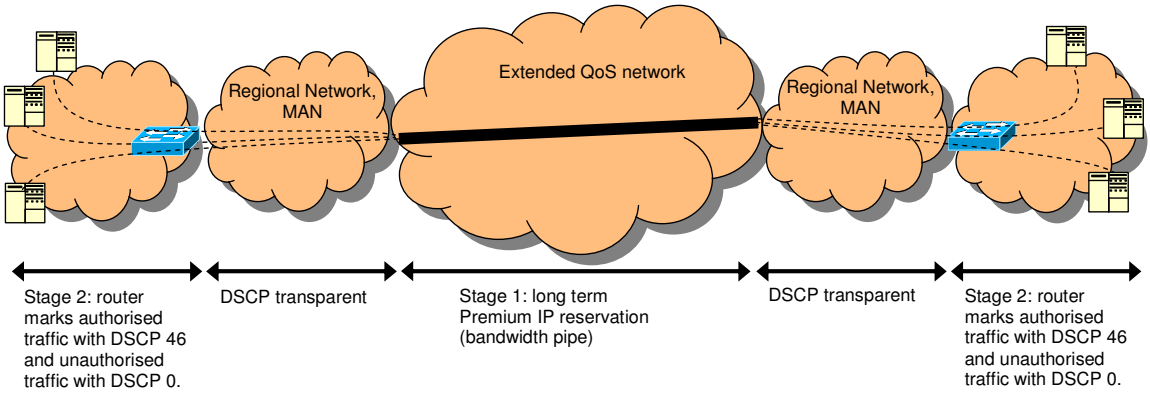

Figure 2: Two stage provisioning process

A detailed explanation of this two stage network service provision is available from [R7] and [R8]. The two stage process can be summarised as follows:

1 Service Reservation (SR) Stage: The HLM or more likely an administrator makes a *long term* guaranteed bandwidth (i.e., Premium IP) reservation between two sites (e.g. subnets) using the BAR Web service. Since Premium IP reservations can only be done in the "Extended QoS network", the BAR Web service forwards this request to the NSAP. The NSAP then makes the appropriate reservation in the extended QoS domain. It should be noted that, to satisfy the current data-replication workflow, this long term reservation is made between two sites (e.g., between a subnet in CERN and a subnet in another institute) and not between two machines.

2 Service Activation (SA) stage: When a Grid job is about to start, the HLM contacts the BAR Web service and requests a (potentially short term) Service Activation, within the context of a Service Reservation made earlier. As the reservation made in the backbone is still effective, the BAR Web service only contacts the L-NSAP services (at the two end sites) in order to configure the local networks. Note that, for this to work, any network between the local network and the backbone would have to be *DSCP transparent* (i.e., it must not alter the value of the DSCP field).

2.4. BAR FLOW

2.4.1. Services Reservation (SR) Stage

The GÉANT2 NSAP service uses an asynchronous communication model. This means a NSAP user has to first submit a request and later on query the NSAP for the outcome of the request. For this reason, the BAR service also implements an asynchronous communication model. That is, a user first submits a Service Reservation, then later queries whether it is successful or not. Let's examine the request flow for this asynchronous communication model.

2.4.1.1. Service Reservation: Submit Service Reservation

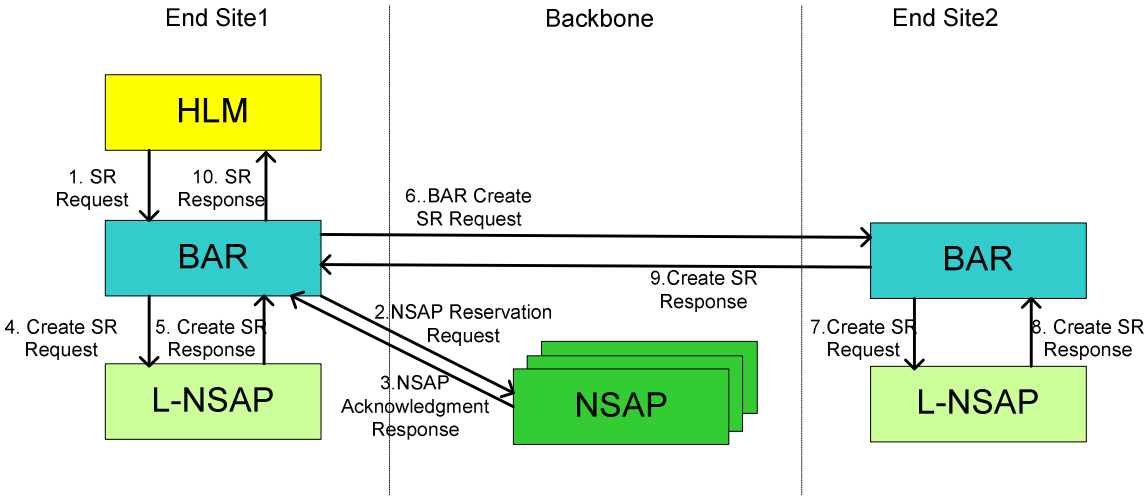

Figure 3: BAR request flow at SR stage; part 1 (request for a reservation)

1. The HLM or an administrator submits a Service Reservation request, for a long term reservation, to its designated BAR service using the specified interfaces.
2. The BAR service translates this application-oriented request to a network-oriented request and delegates it to the nearest designated NSAP at the edge of the backbone.
3. The NSAP acknowledges the receipt of the request immediately to the BAR service. In the case of the GÉANT2 AMPS, behind the scenes the designated NSAP coordinates with other NSAPs along the path to the destination, and a collective decision is made as to whether the reservation is in principle approved. However, as this is a time consuming activity the decision has to be queried later on through the BAR service.
4. The BAR service asks the L-NSAP at the source of the traffic to create an entry for the Service Reservation, so that is knows about this and can manage it during Service Activation.
5. The L-NSAP responds stating whether it created the entry.
6. The BAR service then asks the Remote BAR service to create an entry for the Service Reservation in its corresponding L-NSAP (this is necessary so that the Remote L-NSAP can manage the SR during Service Activations).
7. The Remote BAR service forwards this message to the Remote L-NSAP service.
8. The Remote L-NSAP responds stating whether it created the entry.
9. The Remote BAR service passes this message back to the original BAR service.
10. The BAR service returns an appropriate response back to the HLM (possibly stating that the reservation request is being processed and the result of the request should be queried later on).

A deficiency with the above request flow is that a BAR service waits until the HLM initiates a request to query the NSAP service about progress with the reservation. An improved method would either:

1. Get the NSAP to notify BAR service once a decision is made whether to approve or deny a reservation (“push”).

or

2. Have the BAR service (independently of requests from HLM), at regular intervals, keep asking NSAP until a decision is made (“pull”).

Notification is not supported by the current GÉANT2 NSAP service. Therefore, the only viable option that a BAR service could use currently is use the "pull" mechanism. The BAR prototype [R9] does not implement the "pull" feature.

2.4.1.2. Service Reservation: Query Service Reservation

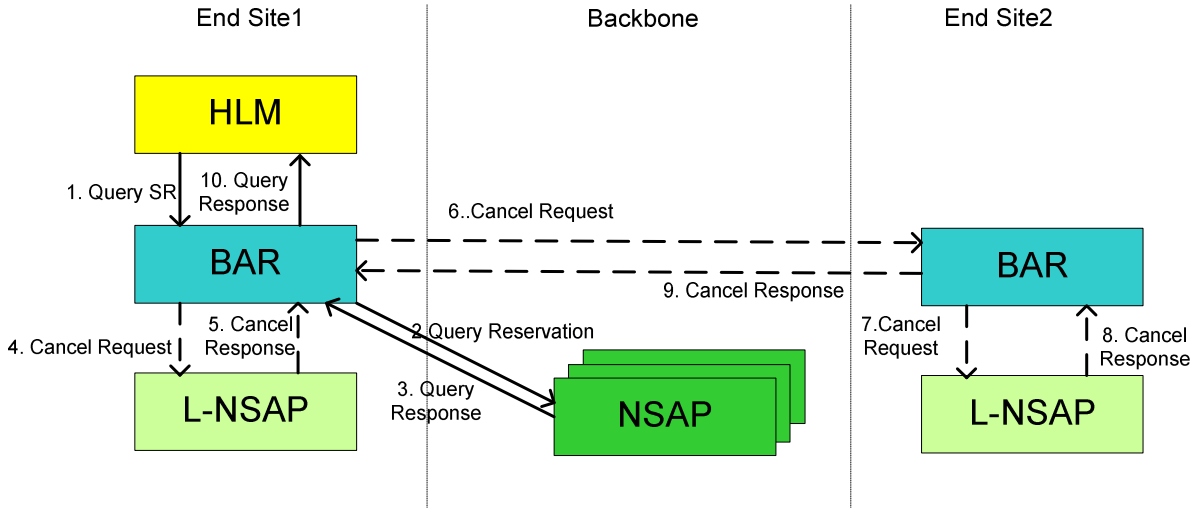

Figure 4: BAR request flow at SR stage; part 2 (query reservation)

1. The HLM or an administrator queries the Service Reservation it made earlier to the BAR service.
2. The BAR service makes a corresponding query to its designated NSAP service.
3. The NSAP sends a response with the status of the request.
4. If the NSAP cannot make the reservation:
    a. the BAR service cancels the "create" request it made to the Local L-NSAP service
    b. The Local L-NSAP responds to the cancel request.
    c. The BAR service sends a cancel request to the Remote BAR so that the "create" request made to the Remote L-NSAP can be cancelled.
    d. The Remote BAR service forwards this request to the Remote L-NSAP.
    e. The Remote L-NSAP responds to the cancel request.
    f. The Remote BAR service forwards the cancel response to the original BAR service.
5. The BAR service sends an appropriate response to the query request. The response should include the status of the Service Reservation (i.e., Pending, Successful or Unsuccessful).

2.4.2. Service Activation (SA) Stage

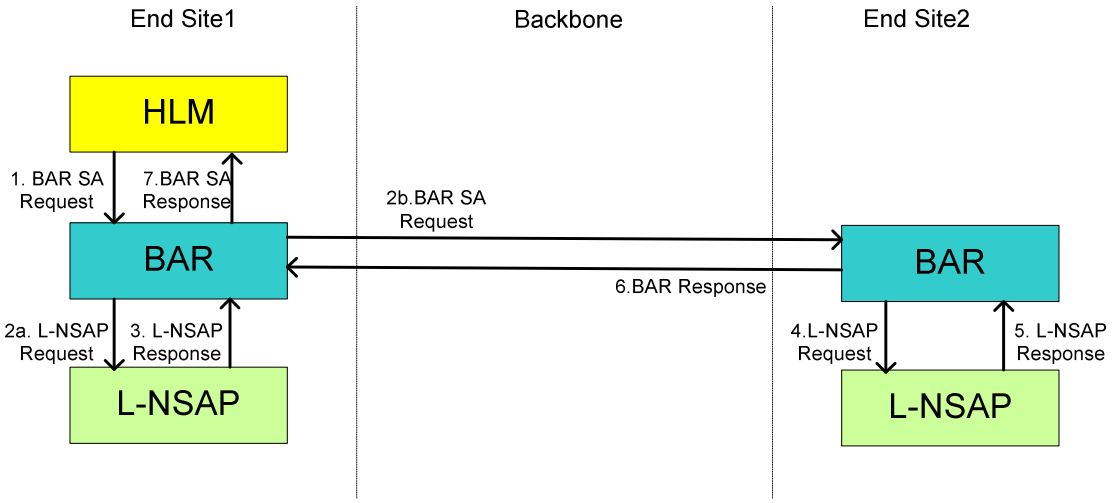

Figure 5: BAR request flow at SA stage

1. The HLM, acting as a client, makes a Service Activation request to its designated BAR service using the specified interfaces.
2. The BAR service translates this application-oriented request to a network-oriented request and delegates it to
    a. the L-NSAP at its end-site;
    b. and the Remote BAR service at the other end-site
3. The L-NSAP responds to its BAR service if it can effect the reservation.
4. The Remote BAR service delegates to the Remote L-NSAP.
5. The Remote L-NSAP service responds to the Remote BAR service stating whether it can effect the reservation.
6. The Remote BAR service sends a reply to the original BAR service with the decision on the request that the Remote L-NSAP made.
7. BAR responds to the client to indicate whether the activation is successful. The Service Activation request is successful if and only if both the local L-NSAP and the Remote BAR service responses are positive. If a local network rejects the request, BAR may need to cancel a successful activation at the other L-NSAP.

2.5. SECURITY

Security is required at the following levels:

- Between the HLM or other clients and the BAR service, to ensure that no-one can alter or intercept the requests and to ensure that only authorised users can successfully make requests to the BAR service via either their Client or the HLM.
- Between the BAR service and the L-NSAPs/NSAPs, to ensure that no-one can alter or intercept the requests and to ensure only authorised BAR services can successfully make bandwidth reservations.
- Between BAR services communicating with each other (to set up the L-NSAPs at each end-site involved in a bandwidth reservation), to ensure that no-one can alter or intercept the

requests and to ensure only BAR services that have been delegated a Proxy by an authorised user can successfully make requests to another BAR service.

- Between NSAP services; this is outside the scope of JRA4 and will not be considered further.

The definition and interaction of the entities involved in the above is contained in [R10].

2.6. BAR AND SERVICE LEVEL AGREEMENT

EGEE JRA1 has proposed a resource-independent architecture for Service Level Agreement management called *Agreement Service* [R11]. This is based on the Grid Resource Allocation and Agreement Protocol working group (GRAAP-WG) [R12] of the Global Grid Forum (GGF). It proposes an architecture for Grid SLA signalling comprising two functional levels: the agreement layer and the service provider layer [R13]. The agreement layer implements the communication protocol used to exchange information about SLAs between the customer and the service provider, and it is responsible for ensuring that the SLA guarantees are enforced by a suitable service provider. An agreement is successfully created if the corresponding guarantees can be enforced by one or more service providers, which are also responsible for supervising the status of their resource pools.

In our common publication with EGEE-JRA1 [R14] we have substantiated the compatibility of the BAR architecture with the Agreement Service. This compatibility, depicted in Figure 6, ensures that BAR can be adopted by the EGEE Workload Management System as an Agreement Service, while the NSAP and L-NSAP are types of *Reservation and Allocation Service Provider* (RASP).

Figure 6: BAR as an Agreement Service

EGEE-SA2 has proposed another approach for end-to-end SLA establishment, which is targeted at providing an operational SLA service to EGEE RCs. According to the EGEE-SA2 approach [R15], the SLA establishment procedure is synchronised with the two stage provisioning procedure described in section 3.2; by the end of the Service Reservation stage a border-to-border SLA is established. Moreover, by the end of the Service Activation stage an end-to-end SLA is established, as depicted in Figure 7.

The EGEE-SA2 proposal mainly involves manual procedures for establishing end-to-end SLAs. However the level of automation could be highly increased as the GÉANT2 NSAP SLA processing capabilities improve.

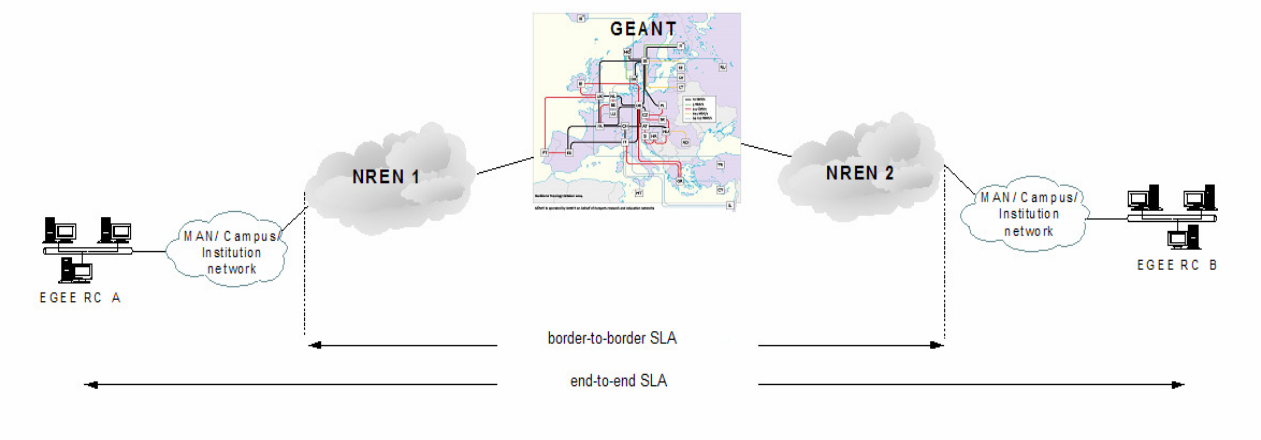

Figure 7: BAR and SLA establishment

3. SPECIFICATION OF COMPONENTS

The components related to a BAR Web service are shown in Figure 1. This section specifies what is expected from these components.

3.1. HLM

HLM is the end-user of the BAR Web service. These middleware applications interact with a BAR service programmatically using the BAR interface specified in Section 4.1. As a reservation of guaranteed bandwidth would only be a **small** part of the business logic of such an application, the BAR Web service should provide a simple interface and should respond to any reservation requests quickly. What is meant by a simple interface is that the service should only provide a set of operations that are necessary to request, query and cancel a reservation. A simple interface also means a small set of parameters per operation: the parameters in an operation should be organised as a small set of mandatory parameters (ones that are absolutely necessary so as to process a request) and extra, optional parameters that the HLM may want to use. What is meant by quickly is that the response will be programmatic, even if involving human intervention; see also section 2.4.1.

An unresolved issue that HLM faces in using the BAR Web service is locating the correct service instance based on the "source" host-name(s) in the reservation. The current assumption is that one BAR Web service per Resource Centre (RC) may be adequate. A solution based on this assumption is to have a registry holding a set of RCs and the location URIs of their corresponding BAR Web services. Then the HLM would have to first deduce the RC based on the "source" host name(s) in the reservation, then look up the registry to obtain the URI of the correct BAR Web service based on the RC. Once the URI to the correct BAR Web service is obtained, then the HLM can use the BAR Web service available at that location to submit the reservation. If gLite already provides such a registry for other services (e.g. the Logging and Bookkeeping Service or GLUE Schema) then a straightforward and uniform way to locate a BAR service can be devised and may be supported in the BAR client API.

3.2. BAR

The functional specification of the BAR service is available from [R16]. That document lists the functional and non-functional requirements on the service, using MoSCoW notation [R6].

The following functional requirements MUST be satisfied by the BAR service (further referred to as the service):

1) The service MUST support the BAR *Service Reservation* function described in [R3].
2) The service MUST support the BAR *Service Activation* function for the Guaranteed Deadline File Transfer (GDFT) Service Type described in [R3].
3) The service MUST support the BAR *Service Activation* function for the Virtual Leased Line (VLL) Service Type described in [R3].
4) The service MUST support the BAR *Cancel Reservation* function described in [R3].
5) The service MUST direct the *Service Reservation* function to the *Reserve Network Service* function described in [R3].
6) The service MUST direct the *Cancel Reservation* function to the *Cancel Network Service* function described in [R3].
7) The service MUST pick up an instance-specific configuration from its local BAR configuration file.
8) The service MUST support BAR *Query* functions described in [R3].
9) The service MUST direct the *Query* functions to appropriate *Query Network Service* functions described in [R3].

The non-functional requirements are of SHOULD priority and concern the areas of: reliability and availability; configuration; installation; usability; use of databases; and administration. The document also defines 15 functions to satisfy all functional and non-functional requirements.

Figure 8 shows the components that make up a BAR Web service. The "Request Processor" component in a BAR Web service authenticates requestors (HLM or BAR) and authorises operations using the configuration and user certificates. It then validates the request parameters and passes them to the "Network Request Handler" to generate a network-oriented request. The "Network Request Handler" forwards the request it has received to the "Request Mapper" in order to obtain network-oriented values when building a request for a network service (for NSAP or L-NSAP). The "Request Mapper" looks up any extra parameters required by a "Network Request Handler" that are missing from the original request or calculates them. The "Network Request Handler" builds relevant network-oriented requests and submits them to the necessary network services (NSAP, L-NSAP or BAR). It logs items such as Service IDs. The "Network Request Handler" processes the responses received from network services, builds an appropriate BAR response and passes it to the "Request Processor". The "Request Processor" returns this BAR response to the requestor (HLM or BAR).

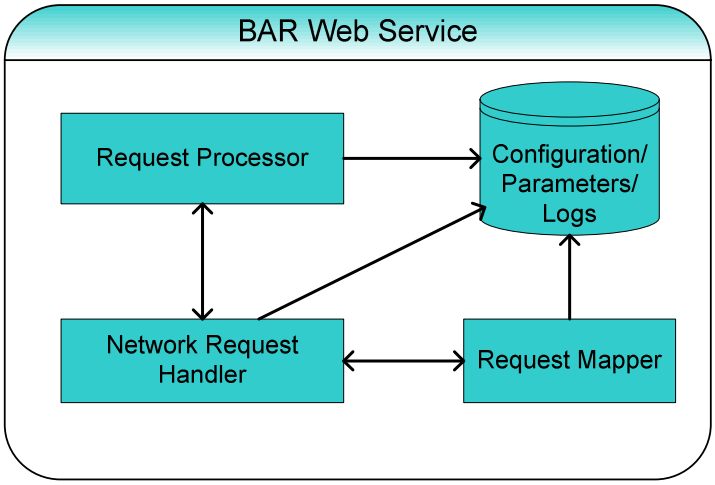

Figure 8: Components in a BAR Web Service

The design based on these components and their interactions is available from [R17].

3.3. NSAP

Although the specification of the NSAP is outside the scope of EGEE, an overview of its functionality is provided in this section. The NSAP specification, implementation and deployment are currently ongoing work within the SA3 activity of the GÉANT2 project. Their first NSAP implementation, AMPS, was released in December 2005. In February 2006 BAR and AMPS were integrated for a series of inter-domain test reservations in the GÉANT2 sandbox environment also involving GRNet and GARR. The following discussion refers to Premium IP and describes the GÉANT2 situation, although our general architecture is not explicitly tied to this service.

Each participating administrative domain (NREN or core backbone) needs to deploy an NSAP. NSAPs of adjacent domains interact with each other in such a way that a Premium IP reservation can be requested over multiple domains. The process of several NSAPs coordinating an inter-domain reservation setup is depicted in Figure 9.

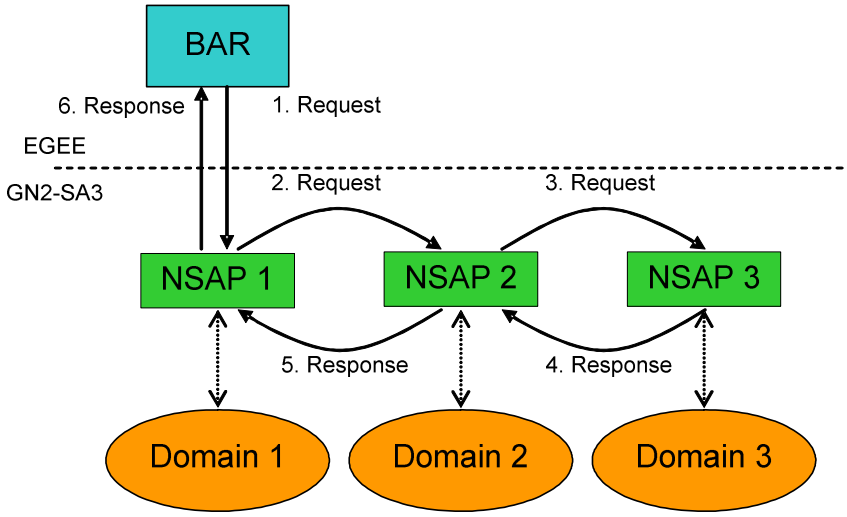

Figure 9: Inter-domain reservation setup

The Premium IP service provided by GÉANT2-SA3 has the following characteristics:

- The maximum capacity that can be allocated for Premium IP is limited by 10% of the smallest link that is being used to implement the service. In the current GÉANT2 network, PIP individual or aggregated reservations can never exceed 1 Gbps.
- Both uni- and bi-directional reservations are supported. Hence, bi-directional reservations can be requested by means of a single request for a bi-directional reservation or by means of two requests for uni-directional reservations.
- The capacity for each direction can be specified separately (*asymmetric capacity*).
- The end-points of the service on the backbone may be located in different administrative domains. However, all domains in the backbone part of the reservation path should be part of what GÉANT2-SA3 defines as the extended QoS network.
- Reservations that will start in the future can be requested (*advance reservations*).
- A unified Authentication, Authorisation and Accounting (AAA) system will be deployed in all GÉANT2-SA3 domains. The authentication and authorisation mechanisms are currently discussed in GÉANT2-JRA5. [R1] lays down the policy for the networks that wish to take part in the Premium IP service. The most relevant policy in the context of EGEE is related to accounting. In order to control the Premium IP usage, the users/projects will be given a credit quota. When a Premium IP service is requested, an appropriate amount of credits is deducted from their account. EGEE is expected to be provided with the necessary credits to be able to use the GÉANT2-SA3 service. However, the quota system will not be implemented in the first phase of the service deployment.

The successor of the GÉANT network, GÉANT2, is being deployed during 2005-2006. In addition to the existing IP services, this network will also offer point-to-point (P2P) services with dedicated capacity. These P2P services will offer higher capacity services (up to 10 Gbps) than Premium IP can offer. The JRA3 activity of the GÉANT2 project has the task to develop an inter-domain system to automate the provisioning of these P2P services. The concept will be similar to what SA3 has defined for Premium IP, namely an NSAP in each domain. Again services are established by adjacent NSAPs communicating to each other. As a result, it should be straightforward to make use of these new P2P services in the context of EGEE. The interface between the NSAP and BAR is defined with this in mind.

3.4. L-NSAP

A Local-NSAP (L-NSAP) is located in the local network of an EGEE end-site. The functionality of an L-NSAP is probably a bespoke implementation that is tailored to the technology and equipment deployed at a specific end-site. As a result, although the Web service must implement the interface defined in [R3], the implementation of the L-NSAP may differ per site. For instance, the AAA mechanisms used may vary and should be catered for by the BAR service. In addition, the technology within the local network may be Ethernet, IP or any other appropriate technology.

The following functionality is considered for the L-NSAP:

- The ability to receive and process a bandwidth allocation request from the BAR software. This interaction should be automated to ensure a timely response back to the BAR software.
- Provisioning the local network. This consists of the automatic or manual configuration of the egress router. In addition, it may be required to configure network one or more switches or routers within the local network. The local network may rely on layer 2 QoS technologies such as Ethernet class of service or DiffServ in case of a routed IP network [R18], [R19]. However, it is the first QoS-capable node (workstation, switches in the local network or edge router) that should mark and/or rate-limit the traffic. This, of course, strongly depends on the capabilities of the network equipment deployed in the local network.
- Check whether the access link has sufficient capacity available. In other words, the sum of all reservations should not exceed the access link capacity (or a predefined percentage thereof). This should be done for the outgoing traffic. In future, a similar check could be implemented to check the available capacity on the access link in the incoming direction, i.e. towards the organisation. The latter requires that the Remote L-NSAP be contacted even in the case when a uni-directional reservation is set up.

Ideally, the edge router of the end-site should do the following:

- For each reservation, a filter or classifier has to be configured at the output interface of the egress router of the end site. This filter will select a certain traffic flow or aggregate based on certain conditions e.g. source and destination IP prefixes, protocol port numbers etc. Rate-limiting and marking then have to be applied to this set of packets. It is the job of the L-NSAP to (automatically) take care of the addition and removal of the configuration.
- Traffic that matches the filter (i.e. is part of a reservation) should be passed transparently when the end-host has already set the DSCP [R20] to 46. Alternatively, the edge router should mark the traffic with DSCP 46. In addition, the traffic flow has to be rate-limited according to the agreed reservation capacity.
- As a default action remark all packets with DSCP 46 to DSCP 0.

Note that the design and implementation of the L-NSAP is neither in the scope of EGEE nor of the GÉANT2 project. As part of our work for EGEE have documented the functional specification of the L-NSAP and reviewed existing implementations for it [R7]. For the duration of EGEE, configuration of local equipment is manual, although the L-NSAP Web service interface for the request response is implemented.

4. INTERFACES

To support the two-stage service provisioning described in section 2.3, the BAR Web service provides two categories of operations: service reservation and service activation operations. As mentioned above, separate types of users are likely to use each category. The long term service reservations are likely to be made by a human user such as an EGEE site administrator whereas service activations are likely to be used by software just prior to transmitting application data (i.e. making use of the network service).

The interfaces are defined fully in section 7 of [R3].

4.1. HLM-BAR INTERFACE

4.1.1. Service Reservation

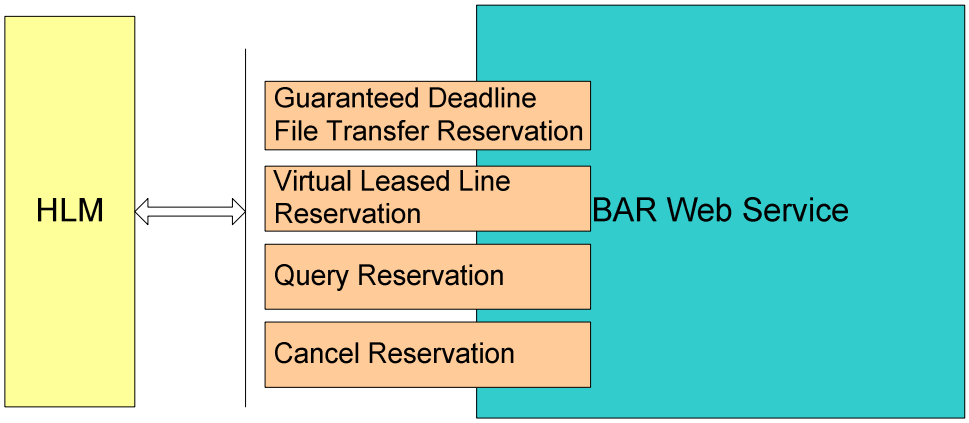

Figure 10: BAR Service Reservation

Figure 10 shows the BAR Web service Reservation interface that is exposed to the HLM. Based on the Service Types defined in [R3], the BAR web service provides operations to support the GDFT and VLL reservations. The "Query Reservation" and "Cancel Reservation" operations are provided to query and/or cancel the existing reservations.

Note that the BAR prototype client only exposes a single reservation type, as the pre-emptive bandwidth allocation mode of use discussed in section 2.3 does not lend itself to a GDFT reservation.

4.1.2. Service Activation

As Figure 11 shows, two Service Activation operations are provided to allow the HLM to make use of a guaranteed bandwidth reservation. Operations are also provided to query or cancel existing activations.

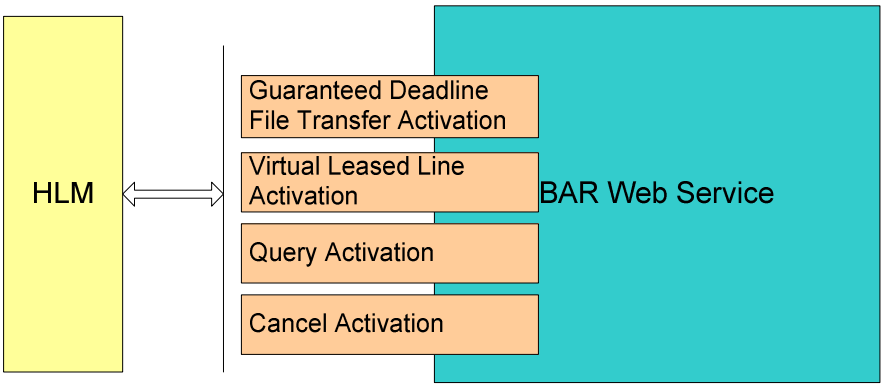

Figure 11: BAR Service Activation interface

4.2. BAR-BAR INTERFACE

The main purpose of the BAR-BAR communication is to pass requests to the Remote L-NSAP service (see Figure 5). To keep the BAR service interface and implementations simple, the BAR-BAR operations are defined to be similar to the HLM-BAR operations (see section 4.1, and also Figure 10 and Figure 11).

A BAR service needs to distinguish whether the requestor is the HLM or another BAR service. This is because a BAR service performs a different set of communications based on the requestor. If the requestor is the HLM then the BAR service talks to its nearest NSAP, the Local L-NSAP and the Remote BAR service; steps 2-5 and 8 in Figure 5. If the requestor is another BAR service then the BAR service **only** needs to talks to the requestor BAR and to its corresponding L-NSAP service (avoiding any NSAP communication); steps 4-6 in Figure 5.

To avoid unintentional usage of the BAR-BAR communication by HLM, BAR has a separate Remote BAR interface, which is only used for the communication between two BARs. This allows BAR to have a "public" interface for HLM with meaningful parameters, and a "private" interface with additional parameters for internal use.

4.3. BAR-NSAP INTERFACE

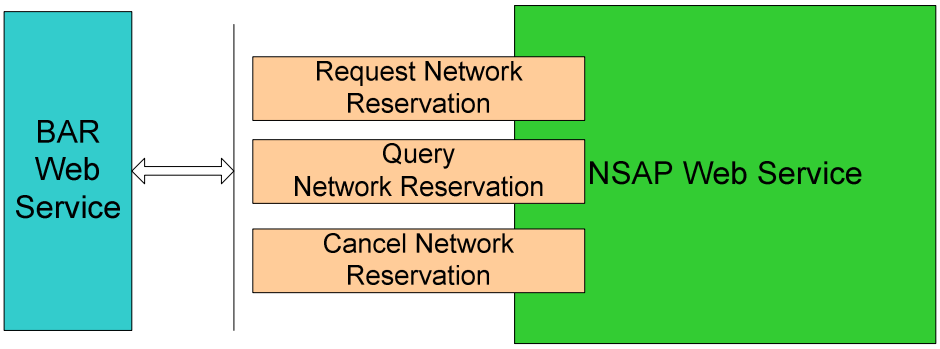

Figure 12: NSAP Web Service

Figure 12 shows the interface exposed by the NSAP Web service. Its interface is used by a BAR service to access network-oriented services in order to fulfil HLM Service Reservation requests. The parts of the NSAP interface and the parameters they accept are described in [R3]. The operation "Request Network Reservation" requests a network service by specifying values for parameters understood by network services. That is, BAR has to translate a high-level Service Type like VLL to

parameters required by the "Request Network Reservation" operation. The operations "Query Network Reservation" and "Cancel Network Reservation" query and cancel an existing network reservation.

4.4. BAR-L-NSAP INTERFACE

The L-NSAP is the Web service that handles requests to configure the router of the local network and also performs admission control. In other words, it checks whether a Service Activation, which is a real-time request for a bandwidth service, can be accommodated by an existing Service Reservation, which represents a bandwidth pipe in the extended QoS network.

After a Service Reservation request has been accepted by the extended QoS network, the BAR component will communicate the details that it received in the response from the NSAP to the L-NSAP. The functions of the L-NSAP interface to install, remove and query Service Reservations are shown in Figure 13.

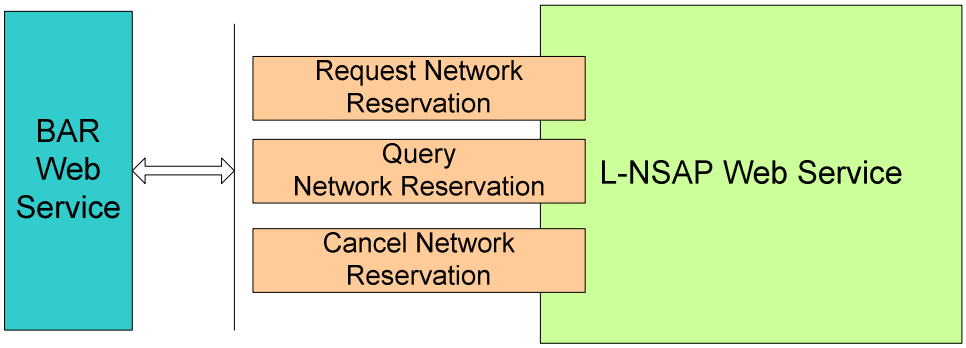

Figure 13: L-NSAP interface for service requests

The L-NSAP is also contacted during a Service Activation. Because the bandwidth in the extended QoS network has already been reserved during the Service Reservation, the Service Activation is only required to check the available capacity of the Service Reservation and to configure a classifier, policer and marker (as defined by [R21]) at the router in the local network. Three operations are supported for the Service Activations. The operation "Request Network Service Activation" requests that traffic defined in the associated Request be authorised to pass through the local network of an EGEE end-site towards the backbone and, if required, that the traffic be marked and policed. The operations "Query Network Service Activation Status" and "Cancel Network Service Activation" query and cancel an existing network reservation. The L-NSAP interface for the Service Activation functions is shown in Figure 14.

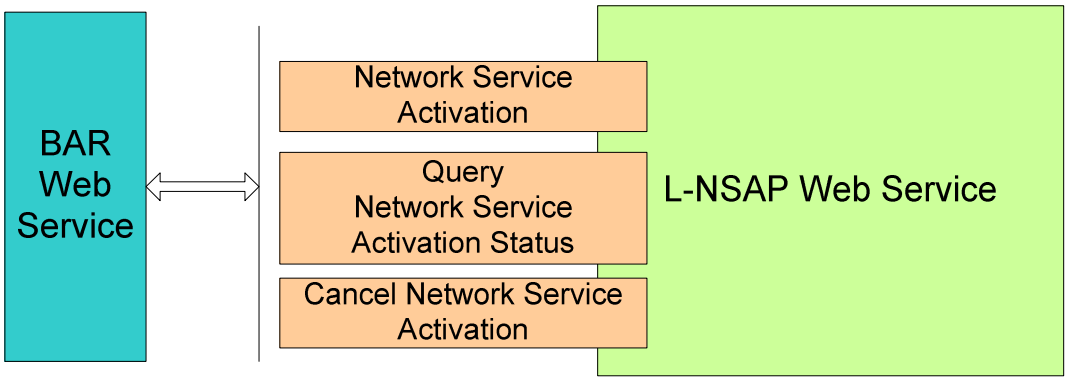

Figure 14: L-NSAP Web service interface for service activations

5. BAR STATUS

The latest and last EGEE BAR prototype is available from [R9]. The prototype implements the single BAR Service Reservation, Query Service Reservation and Cancel Service Reservation functions mentioned in this document. It also implements GDFT and VLL Service Activation functions, and also Query and Cancel Activation functions. Details of differences between the specification and the final BAR prototype implementation are available [R22].

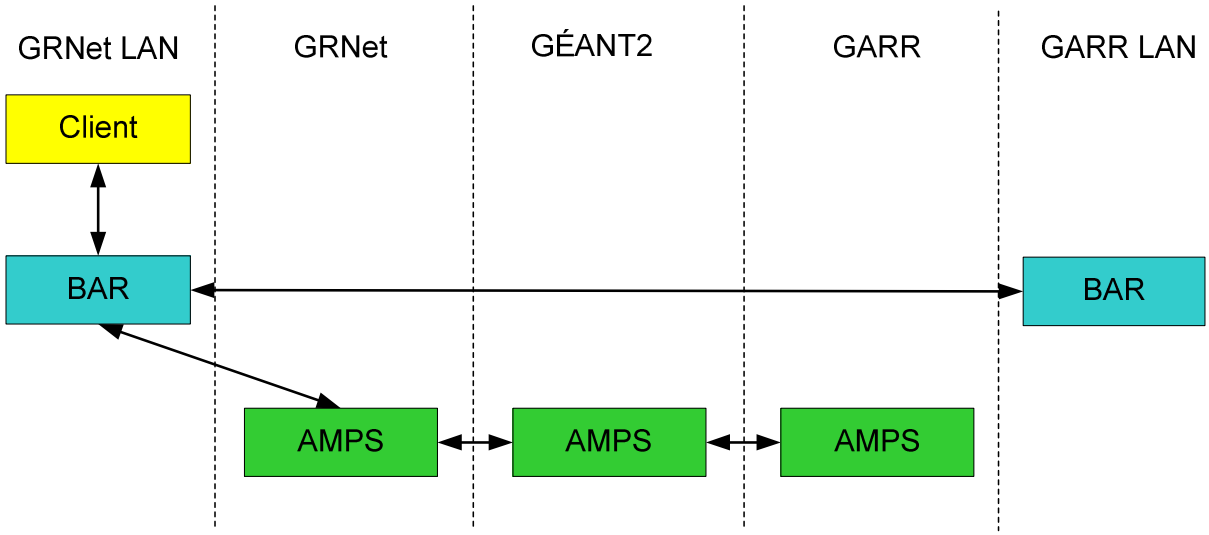

Figure 15: Deployment of the BAR prototype.

The prototype has been integrated with the GÉANT2 AMPS service, as seen in Figure 15. In collaboration with the SA3 activity of GÉANT2, we deployed AMPS on GARR, GRNet and GÉANT2 and used the EGEE BAR Web service and GUI client to request reservations between the GRNet and GARR LANs. These were the world's first software-based, inter-domain BAR requests, validating our architecture and interfaces.

6. REFERENCES

[R1] GÉANT2-DS3.9.1 Policy for Allocation of Premium IP (restricted access)	http://intranet.geant2.net/upload/pdf/GÉANT2-05-017v7.pdf
[R2] "The SEQUIN project"	http://archive.dante.net/sequin/
[R3] DJRA4.1: Specification Of Interfaces For Bandwidth Reservation Service	https://edms.cern.ch/document/501154/1
[R4] EGEE SA2: Requirements and Uses cases from network applications	https://edms.cern.ch/document/476742/1
[R5] DJRA4.1 Addendum: Network Resource Reservation Software And Interfaces - A State Of The Art Survey	https://edms.cern.ch/file/501154/1/EGEE-DJRA41-Addendum-501154-v1.1.pdf
[R6] Key words for use in RFCs to Indicate Requirement Levels, RFC 2119	http://www.ietf.org/rfc/rfc2119.txt
[R7] L-NSAP functional specification	https://edms.cern.ch/document/635851/
[R8] Network Service Provisioning Model	https://edms.cern.ch/document/570622/
[R9] BAR Prototype Final Distribution	https://edms.cern.ch/document/714438/1
[R10] EGEE JRA4: Security Architecture For Bandwidth Allocation and Reservation	https://edms.cern.ch/document/571891/1
[R11] Agreement-based Workload and Resource Management; Andreozzi, S.; Ferrari, T.; Monforte, S.; Ronchieri, E.	In Proc. of the 1st IEEE International Conference on e-Science and Grid Computing, Melbourne, Australia, Dec 2005; IEEE Computer Society, ISBN 0-7695-2448-6, pp. 181-188.
[R12] The Grid Resource Allocation and Agreement Protocol Working Group; Global Grid Forum	https://forge.gridforum.org/projects/graap-wg
[R13] Andrieux, A.; Czajkowski, Dan, A. et al.; Web Services Agreement Specification (WS-Agreement); GGF Grid Resource Allocation Agreement Protocol draft, work in progress	https://forge.gridforum.org/projects/graap-wg/
[R14] "Agreement Signalling and Network Service Provisioning for Grids", T. Ferrari, E. Ronchieri, K. Kavoussanakis, C. Palansuriya A. Patil. In Proceedings of IEEE/Create-Net GRIDNETS 2005 (Oct 6-7) at Boston, MA, USA.	http://egee-jra4.web.cern.ch/EGEE-JRA4/gridnets2005.pdf
[R15] EGEE SA2: Revised SLAs and Policies	https://edms.cern.ch/document/691426/1.1/
[R16] EGEE JRA4: A Functional Specification for Bandwidth Allocation and Reservation	https://edms.cern.ch/document/589518/1
[R17] EGEE JRA4: A Design For Bandwidth Allocation And Reservation	https://edms.cern.ch/document/591621/1
[R18] IEEE 802.3 Standard	http://www.ieee.org/portal/site/mainsite/menuitem.818c0c39e85ef176fb2275875bac26c8/index.jsp?&pName=corp_level1&path=about/802std&file=index.xml&xsl=generic.xsl#802_3

[R19] Application-Driven Networking: Class of Service in IP, Ethernet, and ATM Networks	http://www.redbooks.ibm.com/redbooks/pdfs/sg245384.pdf
[R20] Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers	ftp://ftp.rfc-editor.org/in-notes/rfc2474.txt
[R21] An Architecture for Differentiated Services, RFC 2475	http://www.ietf.org/rfc/rfc2475.txt
[R22] MJRA4.5: End-to-End Specification for Bandwidth Allocation and Reservation	https://edms.cern.ch/document/593453/2/