Bandwidth on Demand Services for European Research and Education Networks

Mauro Campanella
GARR
Via Celoria 16
20133, Milano, Italy

Afrodite Sevasti
GRNET
56 Mesogion Ave.
11527, Athens, Greece

Radek Krzywania
PSNC
Noskowskiego 12/14
61-704 Poznan, Poland

Kostas Stamos
RACTI
61 Riga Feraiou str.
262 21 Patras, Greece

Victor Reijs, Dave Wilson
HEAnet
5 George's Dock
Dublin 1, Ireland

Chrysostomos Tziouvaras
GRNET
56 Mesogion Ave.
11527, Athens, Greece

***Abstract*- The Joint Research Activity 3 of the European project GN2 aims to specify and develop a prototype for a Bandwidth on Demand service destined to operate in a multi-domain environment using heterogeneous transmission technologies. This paper reviews the project's key achievements (general architecture and early prototyping results) and reports on current research and development areas (abstract network description, technology stitching and path finding) with a particular focus on inter-domain issues.**

Keywords— Bandwidth on Demand, GMPLS, abstract network language, NREN, GÉANT2, inter-domain

I. INTRODUCTION

The GN2 European project [1] encompasses a range of research activities to advance both networking and user services in Europe. Central to this project, is the goal of providing high-quality services from one end user to another over multiple interconnected networks. The GÉANT2 [1] network connects 34 countries via 30 national research and education networks (NRENs), using multiple 10Gbps wavelengths. GÉANT2 also connects to worldwide NRENs and the public Internet to ensure a global Gigabit-per-second connectivity for all users. The GN2 Joint Research Activity 3 (JRA3) has specified and is now prototyping a Bandwidth on Demand (BoD) service intended to operate in a multi-domain environment using heterogeneous transmission technologies. This paper reports on the current achievements of JRA3, focusing on its current design and implementation in our inter-domain service provisioning solution. The work presented here is utilizing and complementing existing research and standards in order to provide a system which is innovative as a whole.

II. RELATED WORK

At the beginning of the JRA3 activity, a questionnaire was distributed to a targeted set of potential BoD users (both science projects and research network operators) to assess user BoD service requirements. These users shared a common requirement for a dedicated high bandwidth connection between remote sites. By analyzing the responses to the questionnaire (the detailed results of which can be found in [2]), the need for BoD-like services was identified for a number of different communities both in Europe and worldwide. BoD users reported an interest in an end-to-end BoD path that could be provisioned using a standardized interface and that would support resource reservation and service monitoring. While it is intended that in early phases of the BoD project, the signaling interfaces would be static, increased intelligence through dynamic signaling will eventually be required. User expectations of the duration of the BoD service proved to be quite diverse. Also, a wide range of bandwidth requirements was reported, ranging from 10Mbps to 10Gbps. Further, a great diversity of underlying network technologies was seen, the most common being 1 and 10 Gigabit Ethernet and SDH. This initial set of requirements, although limited have drived the design of the BoD service as presented here. A second iteration of the requirements' survey is planned .

A survey was also conducted into projects that are active in the same research area as JRA3 and are developing network capacity management middleware. The latest results of this work can be found in [4]. Table 1 provides an overview of some of the projects that were surveyed by JRA3 and compares these to JRA3.

TABLE I. CAPACITY MANAGEMENT PROJECTS SURVEYED BY JRA3

Project	Comparison with the JRA3 activity
BRUW [5]	The proposed service model is similar to JRA3 . BRUW focuses on service provisioning using L3 MPLS QoS in contrast to JRA3 which assumes diverse networking technologies.
DRAC [6]	DRAC is designed as a single domain management tool in contrast to JRA3 which targets a multi domain environment. JRA3 follows a vendor-neutral approach, whereas DRAC targets only Nortel equipment.
DRAGON [7]	DRAGON focuses on intra and inter-domain BoD provisioning and deals mainly with GMPLS enabled domains.
IPSPHERE [8]	The IPsphere Forum's mission is to deliver an enhanced commercial framework - or business layer - for IP services. Although both the IPsphere and the JRA3 approach address multi domain resource access and service provisioning, IPsphere operates at the IP layer and focuses on business level functionality, unlike JRA3
MUPBED [9]	MUPBED is based on the uniform deployment of an ASON/GMPLS control plane using a layered model, which is a different approach to JRA3.
OSCARS [10]	OSCARS currently focuses on single-domain dynamic LSP establishment; inter-domain provisioning is proposed on their roadmap.
UCLP [11]	UCLP is mainly concerned with resource partitioning and resource access and utilization. According to UCLP, users own/control the resources in the network. This concept is not fully in line with the approach followed by JRA3. There are, however, similarities between the two approaches, such as the abstraction of resources, which in UCLP is achieved with the use of lightpath objects.
VIOLA [12]	VIOLA defines and tests end-to-end service models using a bandwidth reservation system. Although the work is more on circuit switching, intra-domain provisioning and integration with Grid middleware, similar concepts with JRA3 are identified in the inter-domain path computation, the abstraction of network topologies and the definition of policy and AA(A) modules.

III. ARCHITECTURE AND MAIN PRINCIPLES

The architecture of the JRA3 BoD service has been designed to meet the fundamental requirement for operation in a multi-domain, multi-technology environment. The research activity has produced an overall framework and basic architecture specification for the BoD service [3].

The JRA3 BoD service has the following characteristics:

- Multi-domain: the end-users may be located in different administrative domains.
- Capacity: the BoD service will provide end-to-end capacity guarantees. The capacity is not restricted to a predefined set. The minimum amount of capacity that can be requested will depend on local domain policies and restrictions imposed by the technology used (e.g. SDH has a 155 Mbps granularity).
- Point-to-point: the BoD service provides Point-to-Point links. Point-to-Multipoint may be realized as a set of point-to-point services.
- Symmetric capacity: the same capacity is provisioned in both directions between end-users. Although application requirements might be asymmetric, for the BoD system prototype this approach was adopted due to the symmetric properties of the technologies initially used
- Symmetric paths: the BoD service provides identical forward and return path routing.
- Scheduled reservations: A BoD service instance may be requested in advance. A reservation is expected to last from days to years. A minimum time period is required between the request and the actual provisioning of the service - this time will be a function of the level of automation in the process, which will gradually increased throughout the phased BoD system development.
- Protection: the system is capable of providing either no protection, partial protection or full protection. All services with full protection require the set-up of two completely separate paths from source to destination, including the physical layer, so as to survive failures e.g. from fiber cuts.

The architecture aims to achieve some key concepts:

1. Provisioning of the end-to-end service is implemented over a set of administrative domains which collaborate on a peer-to-peer basis. This is a distributed approach, which allows independence of policies and implementation in each domain. This requires an Inter-domain Manager in each domain for the exchange of network information and user requests, using a common abstract network description language with the peering domains.

2. The solution should be both modular and extensible to support external services such as the authorization and authentication infrastructure (AAI) or existing functionalities for example the Intra-domain Manager, where it is present. The system leverages a set of defined interfaces and supports unique implementations in individual domains.

3. An abstract network resource representation is used to describe the technology specific details in a common format. This provides a non-ambiguous, simple and common language to all the BoD system components and, in particular, for negotiations between domains. It became clear early on in the engineering process that the whole system and the inter-domain communication would be simplified with the use of a common representation of the network and its service components. The abstract representation adopted has been designed specifically for the needs of the JRA3 BoD system, as no existing standard is available to meet the needs of a multi-technology, multi-domain environment.

4. The local domain management function is split from the inter-domain external communication and management. This separation allows R&D on the less standardized multi-domain sector to proceed autonomously. At the same time, it supports the leveraging of existing inter-domain managers through wrappers and interfaces, exploiting a modular approach.

A. The JRA3 BoD service architecture

The GN2 JRA3 BoD system is composed of the following modules:

- Domain Manager (DM)
- Inter-domain Manager (IDM)
- Technology Proxy (and Resource Modeling) module
- AAI module
- Policy module
- Pathfinder (inter-domain and intra-domain)

and two ancillary modules:

- Information Storage System
- Location service

Amongst these modules, the IDM is a key component and is discussed in the next section. Currently, JRA3 development efforts are focused on the implementation of this module.

The main function of the Domain Manager (DM) is to instantiate BoD instances within a single domain. The DM has a detailed knowledge of the topology of its domain. It participates in the inter-domain pathfinding process by examining the feasibility of providing an end-to-end path within its local domain. It contacts the Technology Proxy module to request the configuration of the BoD service instance.

The Technology Proxy module performs the translation of requests received by the DM (in abstract network language) into vendor- or equipment-specific configurations. The BoD architecture foresees set of proxy modules, each one independent of the others, which can be modularly attached to the BoD system. The proxy module may configure the network via an existing Network Management System (NMS) or act as a GMPLS agent; consequently it does not act directly upon the network, but relies on an intermediate control layer.

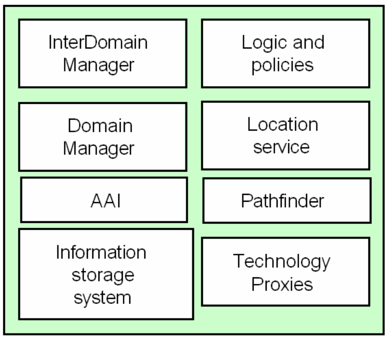

Figure 1. JRA3 BoD system Modules

The Policy module contains all the rules and policies which are available for use by other modules when inspecting and elaborating on a request. The rules are collected in a single module for easy maintenance, modification and to enforce coherence.

The Pathfinder module contains the algorithms and the logic to search for a path that satisfies each BoD reservation request according to specific sets of constraints, algorithms and policies. The module is composed of two main blocks devoted to inter-domain and intra-domain enquiries. The search returns a list of candidate paths using constraint-based shortest path algorithms.

The Information Storage System and the Location Service function mainly offer support to the other modules. The Information Storage System is responsible for providing storage, archival and database functionalities for data explicitly relevant to the BoD system, while the Location Service locates the addresses of all type of services and modules.

If the local administrative domain does not provide basic network management services, such as a Network Management System (NMS), monitoring services and an Authorization and Authentication Infrastructure (AAI), the BoD architecture may implement the minimum required level of these functionalities.

The architecture recognizes the importance of an addressing and labeling scheme for all the components of the BoD system. The GN2 JRA3 activity plans to use IPv6 addresses both as labels and addresses at the control plane.

IV. The Inter-Domain Manager

The IDM is the BoD module responsible for receiving BoD service requests from a user, application or another domain and is responsible for approval and instantiation of these requests. The functional specification of the IDM is reported in [13]. Engineering the inter-domain BoD provisioning is the 'glue' that allows streamlining of the reservation process. This requires definition of the information exchange needed to set up inter-domain BoD service instances. The IDM functionality is independent of the underlying individual per-domain implementations and applies also to manual per-domain BoD provisioning. The GN2 JRA3 research effort has pragmatically focused initial development on the inter-domain layer of the BoD system as it will be common to all domains and must be available from day one.

When service end points are in different domains, the architecture requires the IDMs involved to cooperate on a peer-to-peer basis to create the requested end-to-end path. Such a distributed model allows a great level of independence within domains. Each domain is free to choose the policies and technologies it wishes to use to provide the BoD service. The peering model also allows scaling to a large number of domains and imposes little or no synchronization constraints on the implementation.

The high degree of independence of domains, from the technologies used in the data plane, is achieved by adopting the service and resource abstraction network language.

The IDM module has to:

- Serve as the one and only ingress point to the BoD system. The IDM receives and processes multi or single domain BoD reservation requests. The reservation process takes into account preexisting resource utilization.

- Select the next domain to contact using the Pathfinder module and establish an end-to-end path for serving reservation requests. It produces a list of feasible end-to-end routes/paths through the BoD domains.
- Participate in a commit process between all IDMs along the end-to-end path used to serve a BoD reservation request. The current approach is to use a chain-model commit procedure, similar to the one used by RSVP (Resource ReSerVation Protocol).
- Interact with the AAI service, to authenticate the identity of the BoD service requestor and his authorization privileges for the BoD service. The IDM always applies the local domain BoD rules and policies. It uses a credit management system for the controlled allocation of bandwidth resources among the BoD users of the domain.
- Operate an accounting and logging sub-system that keeps, processes and presents accounting data of the BoD service usage and availability per BoD session and in general within the domain.
- Implement the IDM domain specific policies in each domain, using the information provided by the local DM, for allocation of BoD resources and for management of the BoD service within the domain.

The IDM relies on the local DM to implement the service requested in the form of a provisioned circuit. It is the DM that deals, through the Technology Proxy, with the physical details of the particular network domains and the different technologies used to implement the BoD circuit. For more details refer to [3] and [13].

The IDM is divided into the blocks shown in Fig. 2.

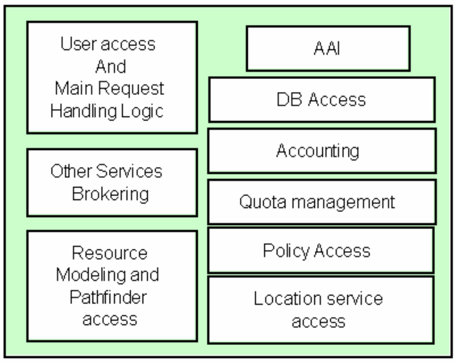

Figure 2. IDM internal blocks

The JRA3 activity has implemented a first, simplified IDM prototype, the key modules of which are detailed in the following paragraphs. Security issues at the IDM level will be addressed in the next phases of the IDM development.

B. The Inter-domain Pathfinder

The inter-domain Pathfinder module is responsible for producing a list of paths that satisfy the reservation requests. It receives a set of parameters from each reservation request, takes into account the local policies and the information received by other IDMs and computes a list of candidate end-to-end paths over which the request can potentially be implemented.

The IETF PCE [21] working group developments are relevant to this area of the JRA3 work, although the PCE group is focused on MPLS and GMPLS technologies while the JRA3 BoD system has a broader scope. The PCE work is based on the requirements for inter-area and inter-AS MPLS Traffic Engineering developed by by the Traffic Engineering Working Group, stated in [22] and [23] respectively. Those requirements explicitly mandate that no topology or resource information is distributed between domains, so as to preserve IGP/BGP scalability and confidentiality. This assumption is considered as an option in the JRA3 Pathfinder.

In the first prototype of the BoD system, the Pathfinder is a simple static module that returns a set of pre-defined paths based on fixed topologies. For the next phase, a more complex pathfinder module is being implemented based on the principles described in [13]. The design effort is focused on the extension and adaptation of the standard implementation of the intra-domain OSPF-TE protocol to an inter-domain pathfinding routing protocol suited for the BoD service needs.

While in a standard routing protocol, signaling, topology discovery and update, routing policies and route computation are tightly coupled, in the BoD architecture, these functionalities are decoupled and placed in different modules.

- Inter-domain topology signaling is placed in the IDM.
- Inter-domain topology discovery and update is performed at the IDM, which receives updates from all other IDMs and creates an abstracted inter-domain topology.
- The routing algorithms operate on the abstracted topology produced by the Pathfinder.
- Policies are applied both in the IDM and in the Pathfinder. Various set of policies exist, such as signaling policies (filtering) and request handling policies (which includes user access policies) in the IDM and path computation policies (algorithm parameter values) in the Pathfinder.

The Pathfinder module receives various input parameters from the Reservation module for each reservation request such as end-point identifiers, the requested capacity and resiliency. The above list of parameters can be enhanced with additional parameters and constraints, such as the cost of the path.

The Pathfinder module, located in the source domain, computes the complete path from the source domain to the destination domain. Then, the source domain IDM contacts the next domain in the path, which in its turn contacts the next one using the announced path (or using its own routing computation in case of a mismatch) in a chained model.

The source domain's view of the abstracted inter-domain BoD topology is built from the announcements from other IDMs. For each specific request, the Pathfinder routing algorithm uses the abstracted topology to return a list of feasible paths based on the parameters of the reservation request. In order to return multiple paths, k-shortest path algorithms ([14], [15],) are used.

Since the Pathfinder uses a chained model, it is preferable to have the source domain compute the path and announce it to the subsequent domains. Even if all domains may re-compute the same end-to-end path, announcement of the path is useful in order to reduce load and ensure consistency. In case the chained reservation process fails at some point, the source domain will use a different path (either from the set already computed by the Pathfinder or by calling the Pathfinder module again) and start again.

The Pathfinder implementation is planed to use the OSPF protocol to distribute information about the links and build the database necessary for path computation. In order to carry Traffic Engineering information using OSPF, the Pathfinder module will use Opaque LSAs of type 9, 10, 11 according to RFC 2370 [16]. There is no plan to partition the BoD service space (which means that all IDMs should see the same abstract inter-domain topology), this means in OSPF terminology that the BoD service will reside in a single area. So for the purposes of the Pathfinder module, the preferred Opaque LSA type is Type 10, which has an area flooding scope [17].

C. *Abstract network representation*

JRA3 has defined a schema that will be used for providing a abstract view of the BoD topology at the inter-domain level in a technology neutral way. Detailed results of this work can be found at [3]. A separate instance of the schema will be installed at every administrative domain that provides the BoD service. This schema contains a specific description of the local domain and summarizes information about each domain that provides the BoD service. An indicative list of the main entities of the schema is provided here:

- An *adminDomain* is a domain under the control of a single administrative authority.
- A *provDomain* is a provisioning domain which is a subset of the administrative domain and uses a common provisioning method (or technology).
- A *node* is a network element that is part of a provisioning domain.
- A *port* is a physical or virtual interface on a node.
- A *link* is a connection among two ports that can belong to i) neighboring nodes of the local administrative domain, ii) neighboring nodes of different administrative domains where one of the nodes reside at the local administrative domain (inter-domain links), iii) non-neighboring nodes where the link entity can either represent a connection among two edge ports of a single domain (virtual links) or the union of contiguous intra-domain links (composite link).
- A *path* is a BoD path that can be i) currently used by a BoD user, ii) a pre-reserved BoD path that will be activated in the future iii) a candidate path produced by the pathfinding module.
- A *linkToPath* is an intermediate entity for representing the links that belong to a single path.
- A *cost* is the cost of a link or a path.

D. *Reservations*

A service request is submitted at the source domain (Home Domain) and is examined using a chain communication model, so that each domain is only in contact with its direct neighbor. Fig. 3 presents a reference topology, with three domains managed by the JRA3 BoD system (domain A, B, and C) and two client domains (Source and Destination). The user requests a BoD reservation from the Source domain to the Destination domain, asking for 1Gbps of capacity and a transmission delay not higher than 50 ms.

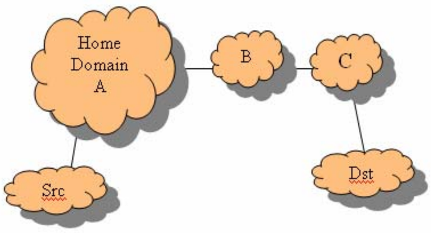

Figure 3. Reference network topology for reservations

First, Home Domain A initiates the reservation process by local validation and executing inter-domain pathfinding identifying a list of feasible paths, e.g. path A-B-C. This includes preparation of the local domain constraints that must be agreed by all domains along the path.

If Home Domain A has sufficient resources available to perform the reservation, the request is forwarded to the next domain on the reserved path (Fig. 4).

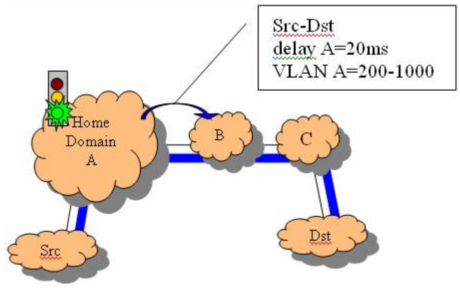

Figure 4. The request is forwarded to the chosen neighbor

The constraints are attached to this request, and are propagated all the way to the last domain on the reservation path. Lastly, Domain B performs local resources' checks and defines its own constraints for the reservation. If the reservation is realizable in Domain B, the request and constraints are forwarded to Domain C (Fig. 5).

E. *IDM-to-IDM communication*

The IDM-to-IDM communication is designed along the Web Services technology principles. This approach allows to easily exchange data between IDMs, relying on IP network delivery service. The messages exchange process is compatible with the RSVP and RSVP-TE protocols as defined in [24] and [25], although the messages are not in RSVP format. Instead, more complex data structures are transferred between domains, using Web Services. Reservation information and states are stored in each domain's local database.

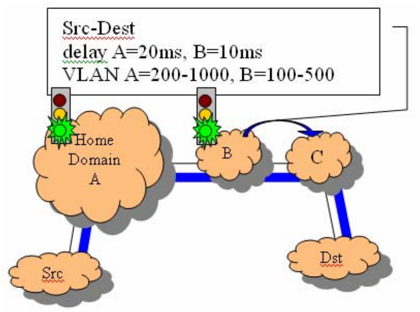

Figure 5. Request forwarding along the reservation path

After a failure, the information recovered from the database is used to recover the local domain state, which is validated and harmonized with neighboring domains' state before becoming active. In case of such recovery, communication or domain failure at any reservation step, procedures will keep consistent the state of reservations within the BoD systems along all domains. This makes the system less vulnerable to errors in resource management, overbooking or incomplete reservations. So although the responsibility for reservation progress belongs to the Home Domain, the process is distributed and involved domains collaborate in emergency situations.

IV. MONITORING OF INTERDOMAIN CIRCUITS

Inter-domain monitoring at layers below IP is being investigated now by the GÉANT2 project. At Layer 3, active and passive monitoring can be implemented in a consistent method across multiple domains (by using a common set of metrics, such as one-way packet delays). In such scenatio, it is relatively simple to inject end-to-end test traffic - e.g. ICMP Echo - in a non-intrusive fashion.

This technique is no longer valid where L1 and L2 services are provided. While IP traffic tests can be performed on a L1 or L2 circuit, such tests will not validate either the BoD circuit delivered to the user, or the range of Layer 3 protocols the customer may choose to run over it. One must instead look to Layer 2 and below for live, non-intrusive monitoring information on a end-to-end BoD service instance

Although the interfaces between Layer 2 domains are well understood, the technologies used to implement the BoD service may vary widely. While each domain may have enough information to troubleshoot its own service, there may be no clear method to correlate that information across domains.

The solution to this problem may be found in the principle of subsidiarity. Each domain should be responsible for monitoring its own network and ultimately, must be responsible for troubleshooting its own technology. Once the relevant domain in which the fault lies can be identified, any further actions can and should be performed by the operators at that domain. This leaves the coordination between domains to be performed by the inter-domain system, based on input from each of the domains.

In order to provide that input, a set of standard metrics must be agreed between domains. Such metrics must be able to be fulfilled regardless of the underlying technology. This requirement lends itself to a phased approach to the coordination task, beginning with metrics that are most critical and may be easily correlated among different technologies, before moving on to metrics that require further research. JRA3, together with other activities in the GN2 project, has selected a basic set of metrics that can be extracted per domain and concatenated on an end-to-end basis for the purposes of the BoD service monitoring. These metrics are:

1. Up/down status of the service
2. Error thresholds exceeded on the service
3. Any other more complex characteristics (e.g. capacity or delay)

Each domain, therefore, performs its own technology specific monitoring (e.g. Errored Seconds in SDH and lost frames in Ethernet) and 'translates' the data into a format common to the inter-domain monitoring system (e.g. UP/DOWN status of a BoD circuit). A visualization tool may then be used to retrieve, concatenate and interpret the per-domain data and display the end-to-end status to the user. The identified fault may then, where necessary, be tracked to the relevant domain for more detailed fault finding.

JRA3 is currently working on the implementation of monitoring tools based on these principles; these tools can be readily integrated with the BoD provisioning system, in close collaboration with the JRA1 and JRA4 activities of the GN2 project [1].

V. TECHNOLOGY STITCHING

In a multi-domain BoD service, many different technologies will be present and a single end-to-end service may need to transit all of these technology domains. Each administrative domain should be able to use their preferred transport technologies. A large number of candidate technologies are available, these include Ethernet over WDM, SONET/SDH, MPLS or native Ethernet and L2 tunnels over IP networks. In order to support automated provisioning in a multi-technology environment, a technology stitching framework is being defined by JRA3.

Four domain types are defined: aggregated, administrative, provisioning and technology.

- An aggregate domain is an aggregation of multiple domains using the same provisioning method (tool or technology) and it can cover multiple administrative domains.
- A single provisioning system (using CLI, NMS, (G)MPLS etc) will interface between the Network Elements and the user in a provisioning domain.

- A provisioning domain can handle multiple technologies (e.g. IP and Ethernet Network Elements).
- A provisioning domain can also be implemented by human intervention (for example a Network Operations Centre: NOC)

The purpose of the stitching framework is to allow all these domains work together to provide an end-to-end service. Static technology stitching has been a networking reality ever since the wide-area IP network came into existence. To get two LANs attached to the edge of an IP network to interwork, their IP addresses and sub-networks have to be decided and communicated to all users. Defining such processes in a standardized way and automating it is the job of the technology stitching framework. This is also an important part of other projects [18] and [19].

Stitching ensures that technologies, implemented in different layers, are able to support a path. This requires that each technology domain be able to perform layer adaptation. For instance an SDH based administrative domain will be able to transport Ethernet using an appropriate adaptation layer such as GFP. The same principle applies to Ethernet over WDM where adaptation may be performed using Generic Framing Procedure (GFP) [20], L2 MPLS (such as Virtual Leased Line-VLL) or L3 MPLS QoS.

Fig. 6 shows the stitching framework in detail. Two planes, important to the communication channel between domains, may be identified, these are:

1. The configuration plane

Normally configuration may be performed using a provisioning system, directly on the network elements or using an automated process. Three different levels can be distinguished:

a. An interface between two peering domains

In this case each interface needs to be configured at each specific layer; for instance in an IP system this involves configuration of the IP addresses and the subnetwork. For an Ethernet layer is could be a VLAN number and for an SDH/SONET layer this could be the Virtual Container Group (VCG). Sometimes an automated process between adjacent domains may be used to provide this, such as auto-negotiating speed (L1), DHCP (L3), route propagation (L3; static, EGP, etc.).

b. An adaptation level within a domain

Some domains may use SDH/SONET as their transport method instead of native Ethernet; Ethernet thus needs to be encapsulated within SDH/SONET containers. These adaptation layers have a close relation with the interface level.

c. A switching/routing level within a domain

Within a domain one needs to connect the ingress and the egress interfaces to realize a connection (can be done for instance by an IGP routing protocol within L2 MPLS VLL or IP network or by a patchcord in an WDM domain).

2. Stitching plane

Besides the above configuration plane, one needs to transport certain parameters within the same peering layer (e.g. IP addresses/subnetworks of end-user domains). Because these peering domains don't have to be adjacent, the parameters need to be transported as part of the communication channel of the IDMs. The aim of the JRA3 BoD service is that an automated stitching process will replace the human communication process (e-mails and/or phone calls), which takes place in most present-day cases.

At the time of writing this, JRA3 is defining the stitching parameters to make sure that the above planes are defined at a technology agnostic level (IDMs) so that it can be implemented at the technology specific level (DMs). The technology agnostic stitching parameters will be part of the earlier described abstract representation model.

The GN2-JRA3-defined IDM-IDM interface provides the possibilities to transport the information needed to provision part of the path in each domain through the domain's DM (by using the JRA3-defined IDM-DM interface). A working path can be implemented if all the parameters are negotiated beforehand between the peering domains (which don't have to be adjacent). The Pathfinder makes sure that this is possible as part of its constraints and path feasibility examination processes. If the involved domains indeed have reserved and negotiated the needed parameters as part of the Pathfinder operations, it is likely that stitching can be implemented in each domain at the moment of provisioning the path.

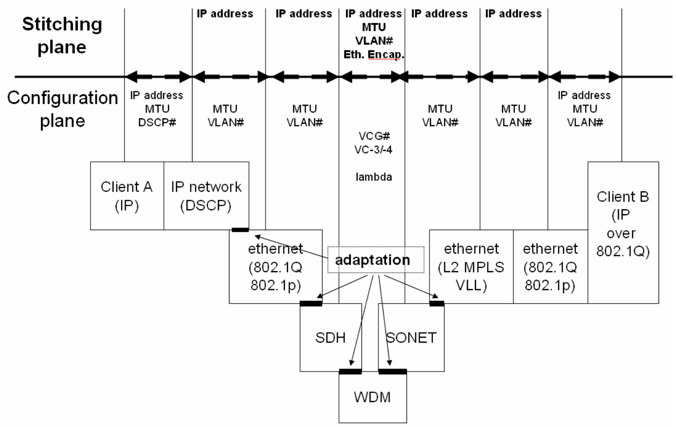

Figure 6. Stitching and configuration planes

After the overview of pathfinding, inter-domain signaling and technology stitching processes made in the last two sections, it has to be pointed out that no single existing, standard protocol would meet the needs of the environment for which the JRA3 BoD system is designed. Although mechanisms and modules of widely accepted protocols are adopted or reused (RSVP, OSPF-TE), the implementation of the main BoD system funtions (such as exchanging abstracted topology information between domains, examining the technical feasibility of an end-to-end path setup over multiple domains, catering for advance reservations, allowing independence of intra-domain provisioning) has only been made possible with the development of an initially proprietary set of protocols, interfaces and modules.

VI. Early Prototyping and testing

A major goal for the JRA3 IDM initial prototype has been to validate key elements of the IDM functionality (such as the IDM communication schema), as defined by the BoD service framework. Testing of the prototype was also planned to allow potential bottlenecks to be identified.

Some components of the IDM have been implemented as static data sources. The Pathfinder module has been equipped with XML configuration files which include pre-defined reservation paths. Also the operation of the Domain Manager module has been emulated with the use of XML files containing information about local domain constraints, such as available VLANs or delays.

The prototype has been examined against six test cases of varying complexity and connectedness. The system has been tested for its ability to handle request concurrency, overlapping resources and overlapping reservation times. Test scenarios included prototype IDMs deployed physically in different domains (in NRENs in Poland, UK and Greece) thus using IP services for IDM to IDM communication and involving delays of a real-world environment. In this way, the effects of the distributed architecture in the case of competing BoD requests were studied. Several tests were tried from simple BoD topologies (a BoD request with no expected feasibility check problems) to complex topologies (with many domains, many different capacities and paths possible).

Testing has successfully proven the design architecture and the processing flow at the inter-domain level in a pan-European scale. The results drawn from the tests will be used as feedback for the next phase of development.

VII. Future work

The design and implementation of the BoD system are rapidly evolving. It is expected that the Pathfinder module will develop to incorporate new algorithms and a number of different Technology Proxies will be implemented to support operation in real technologically diverse networks. The activity plans to continue its collaboration with other projects and standardization bodies to advance the definition of a network abstraction language and standardize the inter-domain communications for the purpose of BoD service provisioning.

VIII. Conclusions

The design and implementation a multi-domain Bandwidth on Demand service is proving to be a complex task which cannot yet rely on a standardized approach. The GN2-JRA3 activity, due to the requirements of its user base, is tackling the less standardized area of inter-domain service provisioning. Network abstraction and a network description language, technology stitching and inter-domain pathfinding are the areas where research and development effort is fundamental. GN2-JRA3, after having defined a framework and architecture for the inter-domain BoD service, is now in the design and implementation phase of developing a fully-functional version of the proposed IDM, following an instructive prototype implementation and experimentation.

Acknowledgment

The authors acknowledge the fundamental contributions of all the GN2-JRA3 participants to the activity

References

[1] GÉANT2: The pan-European R&E network - http://www.geant2.net/

[2] GÉANT2 Bandwidth on Demand (BoD) User and Application Survey (DJ.3.2.1), URL: http://www.geant2.net/upload/pdf/GN2-05-086v11.pdf

[3] Definition of Bandwidth on Demand Framework and General Architecture (DJ3.3.1), http://www.geant2.net/upload/pdf/GN2-05-208v7_DJ3-3-1_GEANT2_Initial_Bandwidth_on_Demand_Framework_and_Architecture.pdf

[4] Second Review of Bandwidth on Demand Technologies (DJ3.2.2.2), http://www.geant2.net/upload/pdf/GN2-06-206v2-DJ3-2-2-2_Second_review_of_BoD_related_Technologies.pdf

[5] Internet2 BRUW Project - Bandwidth Reservation for User Work, URL: https://discvenue.internet2.edu/athena/

[6] DRAC - Dynamic Resource Allocation Controller, URL: http://www.nortel.com/drac.

[7] DRAGON - Dynamic Resource Allocation via GMPLS Optical Networks), URL http://dragon.east.isi.edu.

[8] IPshere - Tom Nolle, 'A New Business Layer For IP Networks', July 2005 Issue of Business Comunications Review.

[9] MUPBED - Multi-partner European Testbeds for Research networking, URL: http://www.ist-mupbed.org

[10] OSCARS - ESnet On-demand Secure Circuits and Advance Reservation System, URL: http://www.es.net/oscars/

[11] UCLP - User-controlled LightPaths, URL: http://www.canarie.ca/canet4/uclp/index.html

[12] VIOLA - Vertically Integrated Optical Testbed for Large Applications in DFN, http://www.viola-testbed.de/content/index.php?id=goalsengl.

[13] Functional Specification of GÉANT2 Inter-domain Manager (IDM) Prototype, URL: http://www.geant2.net/upload/pdf/GN2-06-091v3-DJ3.3.2_FunctionalSpecificationGEANT2_IDM_Prototype.pdf.

[14] D. Shier, "On algorithms for finding the k shortest paths in a network", *Networks*, Vol. 9, No. 3, 1979. pp. 195-214.

[15] D. Eppstein, "Finding the k shortest paths", *35th IEEE Symp. Foundations of Comp. Science*, Santa Fe, 1994, pp. 154-165.

[16] R. Coltun , "The OSPF Opaque LSA Option", RFC 2370, July 1998

[17] D. Katz, K. Kompella, D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003

[18] T. Lehman, "InterDomain Peering and Provisioning via GMPLS and Web Services", *6th Global Lambda Workshop*, Tokyo 1-13 Sep. 2006

[19] GLIF Control Plane and Grid Integration Middleware Working Group: http://www.glif.is/working-groups/controlplane/

[20] S. Gorshe (technical editor), ITU-T Recommendation G.7041/Y.1303 Generic Framing Procedure

[21] Path Computation Element IETF Working Group http://www.ietf.org/html.charters/pce-charter.html

[22] Le Roux, J., Vasseur, J., and J. Boyle, "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, June 2005.

[23] Zhang, R. and J. Vasseur, "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, November 2005.

[24] L. Berger, (Ed), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003

[25] R. Braden, (Ed), "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification ", RFC 2205, September 1997