Management-aware Inter-Domain Routing for End-to-End Quality of Service

Mark Yampolskiy[1,2,4], Wolfgang Hommel[2,4], Vitalian A. Danciu[3,4], Martin G. Metzker[3,4], Matthias K. Hamm[4]
[1] *German Research Network (DFN), Alexanderplatz 1, 10170 Berlin, Germany*
[2] *Leibniz Supercomputing Centre (LRZ), Boltzmannstr. 1, 85748 Garching, Germany*
[3] *Ludwig-Maximilians-Universität Munich (LMU), Oettingenstr. 67, 80538 Munich, Germany*
[4] *Munich Network Management (MNM) Team, Oettingenstr. 67, 80538 Munich, Germany*
myy@dfn.de, hommel@lrz.de, danciu@nm.ifi.lmu.de, metzker@nm.ifi.lmu.de, hamm@mnm-team.org

***Abstract*—Services sensitive to network quality converge onto general-purpose data networks which, in contrast to special-purpose (e.g., public telephony) networks, lack built-in quality control functions needed by many applications, like Internet telephony or video conferencing. High-volume, high-performance applications such as those in Grid and Cloud computing may be too important for customers to rely on mere promises of network quality, while at the same time requiring connections traversing multiple network operators' domains. Thus, in addition to end-to-end QoS assurances, customers of these applications demand management functionality for those connections made available to them. Traditional routing procedures are insufficient to select paths according to these requirements, as they rely on evaluation of only one parameter (e.g., hop count), while QoS parameters alone will account for multiple independent metrics.**

We present a solution that addresses these issues by combining a routing procedure, a common set of QoS operations, and an information model for the representation of connection properties within and across administrative domains.

***Keywords*-end-to-end; quality of service (QoS); inter-domain routing; network management**

I. Introduction

Today's convergent or converged networks are intended to support a growing number of major services with highly varying requirements on the transport system. Such services include end-user facing services, such as voice and video telephony and their conferencing counterparts, but also high-capacity interconnects between the scientific sites of grid installations or the provisioning of cloud services, co-provided in different administrative domains, as well as connections between parts of virtual private networks (VPN links).

The inter-networking layer (i.e., the Internet protocols) does not support quality management inherently. Instead, many different Quality of Service (QoS) schemes have been implemented by different network operators. They do assure a stated quality of the *transport*, but typically omit customer-facing *management* capabilities. At the same time, the scope of network management is limited to single administrative domains.

Nevertheless, customers of important and expensive applications require network management facilities as part of the service being provided. Their capabilities range from read-only inspection functions (e.g., performance monitoring) to functions that alter the state of the network (e.g., adjustments of communication channel parameters).

Such capabilities are readily provided within single domains, but *inter-domain* communication channels (i.e., connections spanning multiple autonomous administrative domains) will require the co-operation of all domains in order to achieve *end-to-end quality guarantees* as well as an aggregate management function presented to the customer as part of the service. Such communication channels, provided as a service to a (paying) customer, we call *concatenated services* (CS).

A. Concatenated services

Combining the outlined demand and focus, our work specializes on the development of a solution for concatenated services, which are probably the most challenging type of point-to-point connections with respect to planning and operation. The following properties are characteristic for CS [24]:

- **User perspective:** a guarantee for the E2E quality of the connection and its management is required;
- **Service composition:** the E2E service is composed of horizontally (i.e., at the same network layer) concatenated connection segments, which are realized by different SPs;
- **Organizational relationships:** all SPs involved in the service's provisioning are independent organizations and are considered equal partners.

Due to high complexity of such connections, some scenarios exhibit unacceptable connection planning and establishment delays, especially when preparing a connection with non-trivial QoS requirements. In some cases, the planning phase, i.e., the identification of path segments that adhere to such requirements, may take up to several weeks (e.g., [36]). This is due to the lack of standardisation and automation of a planning process spanning multiple administrative domains. Each leg of the route must be negotiated with the owner/operator, including the accessibility of a suitable next hop and the QoS and management requirements for the connection. While this concerns long-term connections (i.e., of longer duration than the planning phase), it is obvious that such planning times cannot be always tolerated.

The algorithms underlying common routing protocols (link state, distance vector) are not applicable, as they rely on fulfilment of the optimality criterion. In contrast, the set of routing metrics dictated by QoS parameter thresholds not necessarily holds this criterion, leading to undecidable choices. For example, assuming the requirements for a connection were minimum bandwidth b and maximum delay d with $(b, d) = (1Mbit, 100ms)$, the choice between two alternatives $(1Mbit, 50ms)$ and $(2Mbit, 75ms)$ cannot be made by "shortest" path semantics alone: the first alternative is better in terms of delay, while the second one is better in terms of bandwidth.

A standardisation solution may equally prove unfeasible: Allowing the requester of the connection (user, customer) to specify the kinds and values of the QoS parameters obviates the use of a linear projection function that might calculate "best" values from values of known, i.e., pre-defined, QoS parameters.

In addition, several connection topologies are conceivable. Even though the necessity for QoS assurance exists for different connection types, in the presented work we focus specifically on dedicated point-to-point connections. A discussion of application areas and aspects of different connection types, e.g., point-to-multipoint, is explicitly omitted. Furthermore, from the customer point of view as well as from the perspective of services built on top of network connections, the quality of the connection is important, but not the technology used for its realization. As the bridging between different network layers and technologies is very well understood and is broadly applied, e.g., in the Internet, we assume in our work that network connections are realized on the same ISO/OSI reference model network layer, and consider in our further discussion only the quality of connections and connection segments.

Finally, when devising a routing procedure for use across multiple independent carriers, acceptance of the procedure is crucial for its success: providers may reside in different legal domains, they may have different levels of interest in providing such services (depending on their business model, the state and load of their network, their management capabilities, etc.), and they may have different views on sharing the network management information required by the service (i.e., the managed connection) that is to be provided.

Consider the example sketched in Figure 1. SP_1's customer requires an end-to-end link between a start end-point within SP_1's domain and another end-point (target) within the domain of a different service provider SP_3, which is not a neighbour of SP_1. The customer specifies certain properties with regard to the quality of the transport (minimum bandwidth, maximum delay) as well as requirements on the management of the link during its operation phase (monitoring, constraints on maintenance windows). Via an agreed-upon customer service management (CSM) interface the customer formulates a request to SP_1 to aggregate and to deliver such an end-to-end link.

The customer's knowledge of the topology shown in the upper part of the figure is limited to the end-points in the domains of SP_1 and SP_3, respectively. Each service providers' knowledge is limited to its own domain, and

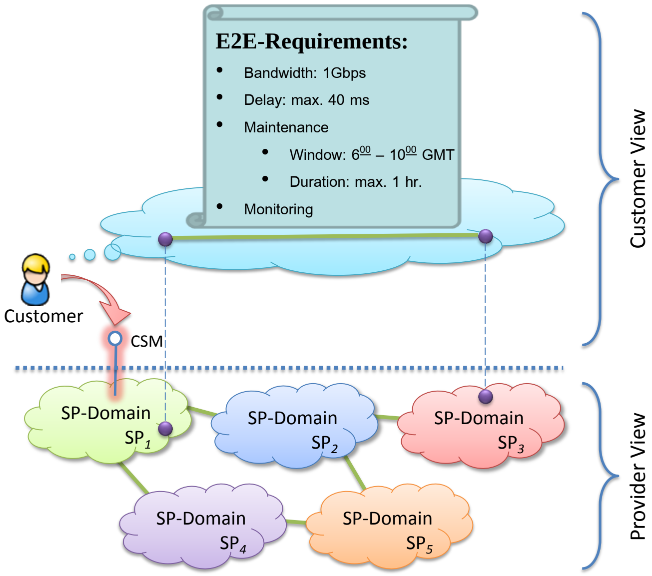

Figure 1. Example request for a concatenated service with QoS and management requirements

the identity of their neighbour SPs. Starting with only this information, the service provider SP_1 needs to find a route to the target end-point in SP_3's domain, such that the aggregate link adheres to the specification of bandwidth and delay and at the same time fulfils the management requirements specified by the customer. Thus, to deliver the link according to specification, SP_1 needs the co-operation of the other SPs in order to select the parts of the route (hops) to the remote end-point and to ascertain the quality and management properties of each part of the link. It is understood that during operation/use of the link each SP will be responsible for the parts of the link that pertains to its own domain. The monitoring has to reflect information from all participating domains in order to provide an end-to-end view to SP_1's customer, and all participating domains need to comply with the customer's requirement on maintenance windows.

In this work, we address the setup of such multi-part connections that involve multiple services providers, each of which is responsible only for parts of the overall network connection. In particular, we focus on the provisioning of high-quality, managed communication channels across multiple autonomous administrative domains; we specifically include operations support in our QoS considerations. The idea underlying this work is to establish transport quality and management properties at the time of routing.

B. Line switching

Experience gained with circuit and packet switched networks has provided deep insights into details about which technical and organizational measures are needed for quality assurance, and about which real-world challenges have to be overcome.

Circuit-switching technology, which is used, e.g., in *Public Switched Telephone Networks* (PSTN), has proven

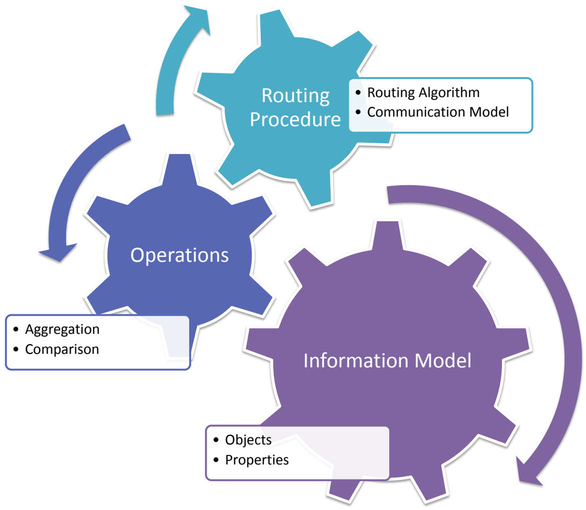

Figure 2. Three inter-operating conceptual building blocks

to be a viable solution for quality assurance, but at the same time it is not truly optimal regarding resource utilization. Packet-switched networks on the contrary have turned out to utilize the available resources much better, but at the same time the parallel communication flows can interfere with each other and consequently affect the quality of each other, for example due to overload packet drops.

Experience gathered in both types of networks types shows that the quality assurance first and foremost requires a thorough planning of the resources needed to meet the quality goals. Furthermore, the desired quality can only be achieved if the necessary amount of resources is allocated to each connection instance. Finally, resource reservation alone does not necessarily guarantee that the goals are met during the service instance operation. Regarding this aspect, current best practices suggest the establishment of management procedures for all life cycle phases of the service instances.

Based on this experience, we focus on concatenated services realized via *line switching* concepts, because this technique has proven to be a viable solution for quality assurance. As line switching can be emulated in a packet switched network, e.g., applying a combination of MPLS and RSVP in IP networks, our choice of the switching paradigm imposes no practical limitation regarding its application to the infrastructure used by real-world network operators.

C. Contribution

In this article, we explain the inter-operation of three conceptual building blocks that address the problem at hand, as sketched in Figure 2: a procedure relying on a management-aware inter-domain routing algorithm, which has been presented in [1], a dedicated information model (IM) for network connections and their QoS-specific properties. This IM, first described in [3], defines the structure and semantics of the information necessary to the routing algorithm. The last building block is a generic QoS function scheme, i.e., a collection of operations which is needed during the routing in order to operate on customer- and user-specific QoS parameters (cf. [2]). During the defined routing procedure, the connection parts are selected and the required quality of all parts is defined. Further, during routing the management functionality from all involved domains can be integrated into the management functionality of a service instance. As the later aspect of routing as well as the interplay of all solution building blocks has not been published so far, we discuss it in more detail.

We summarise the requirements on such an approach in the Section II and proceed to outline our approach in Section III. Thereafter, we describe each of the three building blocks in Sections V, VI, and VII, respectively. Subsequently, in Section VIII, we illustrate the inter-operation of the building blocks by an example before discussing the properties and limitations of the approach in Section IX. We conclude the article with ideas for further investigations in Section X.

II. Requirements

The challenges which have to be overcome in order to realize CS are manyfold. Partially they are caused by CS characteristic properties.

A. Fundamental assumptions

The case described in the Introduction implies that our approach is supported by a set of assumptions, as follows:

a) SPs are independent organisations; they are not obliged to participate in the provisioning of a service/connection.
b) A service instance (connection) cannot be realised by one SP alone.
c) Information about the network topology within an SP's domain is not available.
d) Reliable information about an SP's network capabilities is not available.
e) Management functions and management information are not offered by SPs without prior agreement.
f) Multiple independent QoS parameters will be specified for a connection according to the user/customer demands.
g) There is no fixed set of QoS parameters that will be requested by customers; and there is no fixed combination of parameters.
h) There is no fixed set of management operations that will be requested by customers.
i) Requirements on an existing service instance may change during its life-time.

An important challenge for our approach was to avoid a violation of any one of these assumptions. For this reason, the approach must fulfil the requirements formulated in the following.

B. Collection of requirements

The foremost requirements are those pertaining to the aim of our work:

1) The approach shall determine a path across multiple autonomous SPs' domains.
2) The approach shall ensure that QoS parameter thresholds are respected for the end-to-end service.
3) The approach shall ensure that management capabilities are provided for the end-to-end service.
4) Where necessary, management information shall be aggregated in a manner opaque to the customer.
5) Multiple independent QoS parameters shall be specifiable for a service instance.
6) Multiple independent management functions shall be specifiable for a service instance.

In addition, several requirements originate from the settings of concatenated services.

Participation of multiple SPs: Globalization of business and research collaborations has the consequence that the communication partners, which are using the same network connection, can be spread over the entire world. Due to economic and often also legal reasons, such connections usually cannot be realized by only a single *Service Provider* (SP). Similar to services in the Internet and in PSTN areas, multiple SPs have to be involved in the realization of a single network connection, leading to the following requirements:

7) No requirements shall be imposed on items within SPs' domains.
8) The approach shall not be dependent on the number of SPs that co-provide the service.
9) The approach shall not be dependent on global up-to-date information or on a central instance.

Customer-specific QoS-combinations: The quality of the customer-faced services in general depends on the different quality parameters of the underlying connections. Furthermore, in general it does not only depend on a single QoS parameter, but rather on the service-specific combination of multiple independent QoS parameters. For example, a video streaming service might be insensitive to delay and jitter as long as the connection bandwidth enables the pre-buffering of content that is not yet displayed to the user. However, for an internet-telephony service, the delay and jitter of the underlying connection might lead to a negative user experience. During video-telephony and -conferencing, also the synchronization between video and audio signals becomes important. In summary:

10) The customer shall have only one point of contact, her original SP.
11) The customer shall be given means for managing the service, once established.
12) The customer shall be allowed to specify QoS parameters and management capabilities.

Highly dynamic environment: The rapid development of new services as well as the very high dynamics of changes is characteristic for modern IT services. As the quality requirements on the underlying network connections might differ between services, also the high adaptivity and extensibility of solutions becomes a critical success factor. Especially the extensibility of support for new QoS parameters is needed, which have not been considered before:

13) The approach shall be extensible regarding the QoS parameters supported.
14) The approach shall be extensible regarding the management capabilities provided to customers.
15) Service planning shall be automatable.

III. Approach Outline

We argue that user- and customer-tailored requirements for connection services can be only truly fulfilled, when already considered during the ordering process. Our approach is a routing process, consisting of three conceptual parts: a *routing procedure*, a *generic function scheme* and an *information model* (see section I-C). A routing procedure satisfying the requirements we laid out in Section II takes into account all known existing connections with their properties and end-to-end requirements when finding a path between two end-points.

By employing line switching to realize concatenated services, our algorithm may regard the path through an SP's domain as an atomic link, defined by its ingress and egress points. Based on this decision, our algorithm does not require knowledge of every SP's network topology. This reduces the problem-space and leaves it entirely to the SP to realize the link. As a result, our information model may describe links, properties and functions in a technology agnostic way. Abstracting from components and technologies, it is easier to model inter-domain links, where the end-points are in domains of different SPs.

Connection requirements directly affect the tasks of the routing procedure, which in the case of line switching are: the selection of the path between the two end-points, the designation of all connection segments and interconnecting all points along the chosen path [35]. *QoS routing* (also referred to as *QoS-aware routing*) extends this definition by taking into account end-to-end user requirements. QoS requirements are defined for all connection parts and have to be guaranteed by all SPs, in order to meet the given end-to-end goals.

As discussed in the requirements section, for true quality assurance the management of service instances by the customer has to be considered. Therefore in our work we introduce *management-aware routing* as a QoS-aware routing with additional tasks, namely:

- the definition of management functionality, which have to be provided for each involved connection segment and
- the integration of specified functionalities into multi-domain management functionality of the whole service instance.

As management processes are specified as an interaction between roles with specific responsibilities, the assignment of roles to the involved SPs must be considered as an integral part of the overall management-aware routing procedure. The mentioned routing types and their tasks are depicted in Figure 3.

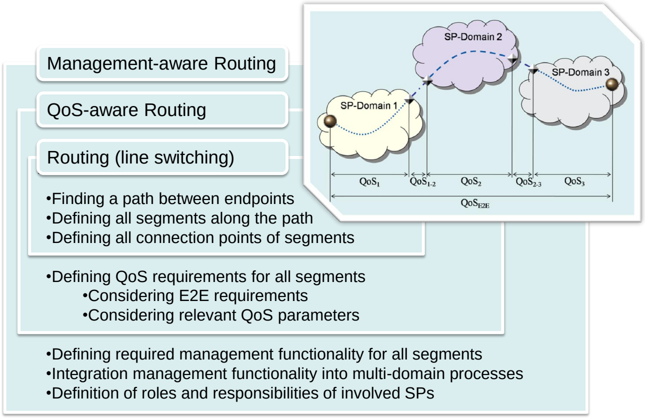

Figure 3. Routing types and corresponding tasks

In our approach the algorithm is designed to split the planning and evaluation of (sub-)paths into multiple smaller tasks, which we see as the key to having a generic set of functions. By breaking down these two problems into smaller parts, our algorithm can be seen as a framework where each metric may provide its own functions and rule sets. Our goal is to describe this generic approach, not to provide an explicit formulation of QoS-requirements and metrics. An example of how our concept may be instantiated will be given in Section VIII.

A. Routing procedure

The routing procedure consists of a *routing algorithm*, which is broadly seen as the very core of every routing approach, and *communication patterns* in the form of a communication model. The purpose of a routing algorithm is to choose between different available connection segments and select those, which will form a path between two endpoints. The details of our algorithm can be found in [1]. In this paper our focus lies on the interaction between components of proposed solution.

Our routing procedure is dedicated to find and establish a path in advance, with an explicit planning phase, before user data can be sent.

As our approach shall not require SPs to share their entire topology knowledge with each other, there needs to be interaction between SPs so that viable paths can be found, requested and managed. Communication patterns describe how routing instances interact with each other and when information is to be requested from another SP, as we aim to keep information exchange to a minimum. Communication relationships also describes how management functions of involved connection segments will be connected and consolidated in order to be offered to the customer.

Basically, we distinguish between *source routing* and *routing by delegation*. While source routing in its extreme form means that the entire routing is performed by the customer's service provider, routing by delegation allows an SP to ask another SP to find a (sub-) path for the remaining part of the route. Through the combination of these two techniques the routing procedure can adapt to policy constraints, concerning the information exchange of SPs.

B. Operations on properties

A generic function scheme is required to specify and evaluate boundaries and optimality criteria for different kinds of quality requirements. The example in Section I-A includes *bandwidth* and *maintenance window* as quality requirements on the requested connection. Bandwidth is limited by the capacities of involved physical links. The link with the *lowest* capacity determines the *maximum* bandwidth achievable on a path. Requirements on maintenance windows are less straight forward than on bandwidth. In the case of maintenance windows *upper and lower boundaries* for maintenance planning as well as the duration of maintenance work are specified. Clearly, an algorithm needs a different set of operations to reason on maintenance windows in contrast to bandwidth. The generic function scheme provides a common interface, so that specialized operations for each parameter may be accessed by the algorithm in a unified manner.

Thus the main goal of our function scheme is to enable the algorithm to take all possible kinds of metrics into account. The most important part of this task is defining the semantics of our proposed generic functions to ensure compatibility of metrics and interoperability between communication partners. Another part of this task is providing a method for representing QoS-requirements in our information model.

C. Information model

Existing information models (IM) often neglect the importance of inter-domain connections. Even though inter-domain connections are comparatively small parts of the overall path, their quality is indispensable for the computation of the overall end-to-end quality. Due to very restrictive SP policies, each SP usually has access only to its end of an inter-domain connections. There are cases in which it is possible to derive the properties of inter-domain links from two partial views onto the same physical link, like available resources on that link. For such cases our IM provides the possibility to describe partial views of different SPs and means to correlate these views.

We designed our information model to describe link properties similarly to connection requirements, which allows us to use the same data model for both. This eases interaction between peer entities as a conversion from requirements to resources is not required, thus leaving less room for misunderstandings and different interpretations. Also, to uniform information exchange, our information model includes management functions so that they can be treated like any other QoS-parameter in the routing procedure.

IV. Related Work

Our presented work covers many aspects in order to provide end-to-end links with user-defined quality-requirements and management-functions. Even though most of the problems outlined in Section II have been addressed in previous work, none of them covers the topic in a full extend.

Originally, physical line switching technologies were used in PSTN networks. In ATM, logical line switching in combination with resource reservation has been implemented. As these technologies were tailored to fit the needs of the voice-only telephony, they do not offer the required flexibility of novel services provided over, for example, the Internet. In this area of (virtual) circuit switching, the Dynamic Circuit Network (DCN) cooperation is probably the most advanced research project. The main focus of this project is on the dynamic provisioning of dedicated paths, literally guaranteeing bandwidth to customers [4], [5]. Among others, projects like OSCARS [6], DRAGON [7], Phosphorus [8], and the Géant-developed AutoBAHN [9] are involved in this cooperation and several successful demonstrations have been presented to the research community [10]. Another approach to on-demand circuit provisioning is the DICE (DANTE–Internet2–CANARIE–ESnet) [11] collaboration of European and North American research networks, where an architectural concept, based on the experience gained in previous projects has already been elaborated and published [12].

The mentioned projects and collaborations are focused primarily on two aspects:

1) various techniques and technologies for dynamic circuit switching and resource reservation within a single administrative domain and

2) interoperability between the developed management systems as well as between networking technologies used in these domains in order to automatically switch multi-domain network connections.

Plans for the consideration of QoS parameters are limited to sole connection propert – its bandwidth. Management aspects in dynamically provisioned circuits are limited to circuit monitoring, mostly reusing praxis approved concepts of the Géant E2E Link Monitoring System (E2Emon) [13]. DICE plans to extend this monitoring concept with a combination of measurements at different network layers and with different monitoring techniques.

An advanced example of network connections with customer-tailored properties is the Géant E2E Links service (also referred to as Géant Lambda) [16], [17]. Among other scenarios, Géant E2E Links have been used to realize challenging connections for the international research project Large Hadron Collider (LHC) [14] and for the Grid cooperation DEISA [15]. These links were established considering multiple QoS parameters as well as management aspects, like inter-domain trouble-shooting procedures or the coordination of maintenance windows among multiple SPs. The biggest drawback of this approach is very long connection establishment time, as these links are planned and set up manually. As planning of Géant E2E Links might consider potential possible connections for which infrastructure still have to be procured, installed, and configured, estimating the time necessary to find and set up a path between two endpoints is very difficult. In the case of Géant E2E Links, the time between circuit ordering and service production start – influenced by different factors – can vary between a few weeks and several months [36].

Management of end-to-end multi-domain network connections is mostly limited to monitoring in the form of technology specific solutions, e.g., the well-established monitoring in SDH networks [21]. Another, still emerging technique is OAM (Operations, Administration, and Management) for Ethernet [22], which aims at a wider scope than merely end-to-end monitoring. For example, failure signaling has already become part of OAM for Ethernet, which allows for quicker detection of and reaction to component failures along the entire path. The more generic multi-domain solutions, like the E2E Monitoring System (E2Emon) for Géant E2E Links, are currently limited to a project-specific combination of technology-agnostic QoS parameters [13].

The routing is mostly covered in the graph theory. In a graph that models a network as vertices and edges, QoS-parameters on path segments are usually represented as edge-weights. Established routing algorithms are based on Bellman's optimality principle and find paths between two vertices. Bellman's optimality principle is valid for the class of functional equations of finding the maximum, the minimum or the k-th element [25]. As this class requires (at least) a partial ordering of the domain, these algorithms are not applicable for multi-dimensional search-spaces and especially multi-weighted graphs. The different weights on a single edge cannot always be merged to a single one. The example in Section I-A includes requirements on bandwidth and maintenance windows. As bandwidth requirements cannot be converted to maintenance window requirements and vice versa, a partial ordering over both weights cannot be found and Bellman's optimality principle cannot be applied here.

Proposed algorithms that are capable of searching high-dimensional spaces, e.g., SAMCRA [27] or H_MCOP [28], require full information about network topology, link properties and connection properties. This make them inapplicable for multi-domain routing, as it collides with the restrictive information policies of most SPs and maximises search-space.

Routing algorithms operate on information about available connections or connectivity possibilities. One of the biggest problems of information models (IM) for network description is data-hiding. For instance, the Common Information Model (CIM) [31] focuses on the description of relations between services and underlying network technologies. While the CIM is a very good basis for managing single networks it cannot be used for our purposes. Modelling a connection segment with QoS-parameters requires modelling of the entire network the segment resides in. In order to comply with the strict information policies in multi-domain environments, networks need to be modelled abstractly with hardly any information. This cannot be accomplished using the CIM and other information models.

Furthermore, almost none of the IMs used for network descriptions associates multiple abstract properties with connections. One noticeable exception is the ITU-T recommendation G.805 for the description of optical transport networks [30], where multiple technical connection properties can be associated with a single connection. In the recommendation these are only technical parameters that are needed to interconnect the segments, e.g., multiplexing of different channels. Especially this recommendation does not consider connection's quality and management functionality properties.

V. Inter-Domain Management-aware Routing

The main task of routing is always path selection. As mentioned at the beginning of this paper, in order to guarantee end-to-end requirements, the required properties of all segments have to be planned. This includes considering the management functionality needed in further phases of the service instance's life cycle. Our proposal for routing and defining of quality targets for the chosen segments is presented in Section V-A. As has already been pointed out, an integral part of *management-aware routing* is the integration of management functionality provided by involved domains into integrated overall management functionality for the whole service instance. We analyse these integration aspects in Section V-B.

A. Connection planning

As the routing procedure operates on knowledge about available potential connection segments, we introduce our information representation to the extent we deem needed to understand our routing proposal. An elaborated discussion of our information model and an elaboration of our decisions will be given in Section VII.

The routing procedure we propose operates on semi-global knowledge about available connections. All this information is represented at the abstraction level of an SP's organizational domain. This means that we look only at connections between network equipment installed at administrative edges of provider networks. In general, all connections which can be realized either within a single SP's domain or interconnecting two neighboring SPs are potential *connection segments* of an end-to-end connection. As from an SP's point of view every realized connection segment is a provided service, we refer to an endpoint of a segment as *Service Connection Point* (SCP). We choose this abstraction in order to be independent of network technologies and of application areas in which our routing concept can be applied. A SCP can be mapped to various "edge-components", from a logical UNI/NNI interface when looking at paths through the Internet, to a Point of Presence (POP) when planning backbone connections.

Semi-global knowledge means that for routing we need information over multiple, but not necessarily all, SP-domains. This knowledge can be extended on demand, by requesting additional information from other SPs. Consequently, during routing we distinguish between already known connection segments and existing but not (yet) known connection segments. The same applies to SCPs of those segments. Further, some of the known connection segments can be considered by the routing algorithm as a part of the planned route.

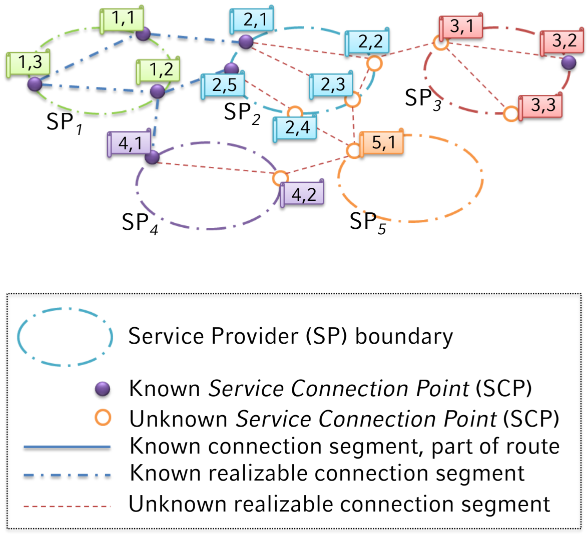

Figure 4. Semi-global knowledge about available connection segments at the beginning of routing [1]

Corresponding to the example presented at the begin of this paper, Figure 4 presents a possible initial semi-global knowledge of SP_1. In this figure, SCPs are labeled with IDs. These IDs consist of two numbers, where the first number identifies the SP to which a SCP belongs and the second number is a sequential numbering of SCPs within an SP. Please note that we have chosen this ID representation only for the sake of better readability and reference in this paper.

Due to various factors, e.g., steady ordering of new and decommissioning of existing end-to-end connections, the resources available to possible end-to-end connections changes constantly. The consequence is that an SP cannot expect to have up-to-date knowledge about all available connection segments. At the same time every SP always has exact knowledge about its own network equipment and available capacities. We further assume that every SP can obtain information about capacities available to interconnections with neighboring SPs. We see this as a common situation for most SPs, as such information is required for proper management of the own network infrastructure. Further, we assume that every SP can determine the quality properties that can be guaranteed over all available own links.

For instance, in Figure 4 SP_1 is in charge of routing between SCPs with IDs "1,3" and "3,2", it can choose which connection segments should be considered as a part of the path. When selecting these segments SP_1 relies on knowledge about available capacities and own preferences regarding which neighbouring SP to use. At the same time SP_1 has to consider customer requirements, i.e., for all considered paths the realizable connection properties must be computed and compared with the user-provided end-to-end requirements. If, at least, one of the end-to-end requirements cannot be fulfilled, another (for the SP lesser preferable) alternative must be considered in the same way.

In the case when SP_1 selects the way through SCPs ”1,1” and ”2,1”, it also implies that the next segment is realized by SP_2. In [1] we discuss advantages and drawbacks of different routing strategies. Based on this discussion, we recommend the *source routing* strategy as long as possible. This means that, in this particular case, despite the fact that the next segment is realized by SP_2, the selection of this segment for the end-to-end connection should be done by SP_1. This in turn means that SP_1 has to obtain actual information from SP_2, containing the “remote” SCPs through which SP_2 is willing to realize connection segments and which properties can be guaranteed for these segments.

We see two main advantages in using source routing strategies: 1) avoiding (or reduction of) nested communication relationships speeds up the overall communication processes and decision propagation, and 2) retrieved information about available connection segments and their properties can be reused, if alternatives have to be examined. Nevertheless, employing source routing requires high trust relationships between communicating SPs. Especially in big open provider collaborations, e.g., the Internet or the PSTN-network, this is not always the case. To meet this problem we propose *on-demand delegation* of the routing task for the remaining part of a path. This is based on the praxis approved theory, that SPs are more likely to trust topologically closer SPs and especially all neighbouring SPs.

Similar to a CSM interface (see Section I), which is commonly referred to as the management interface between provider and customer, we introduce the *Domain Service Management* (DSM) interface as a means for communication between independent SPs interested in a collaborative realization of end-to-end network connections. Communication between SPs should always be performed through the DSM interface.

As mentioned in Section III, obtaining information about SPs’ available capacities, as well as making decisions about segments from outside an SP will most probably violate domain policies. In order to avoid or at least minimize violations of these policies, we propose to specify always strict boundaries for requests. When requesting a segment, the requesting SP provides information about the remaining part of the end-to-end path, i.e., the two determining SCPs between which the path still has to be established, and about the connection properties the end-to-end path has to comply to. Such procedures allow queried SPs to provide only few alternatives matching the specified restrictions, as opposed to providing full and unreflected information about an SPs infrastructure. Furthermore, we propose that the queried SPs provide these alternatives in the most-to-least preferable order. Even though misuse cannot be fully avoided, we propose that the SP responsible for routing (in the example SP_1) tries the alternatives in the specified order.

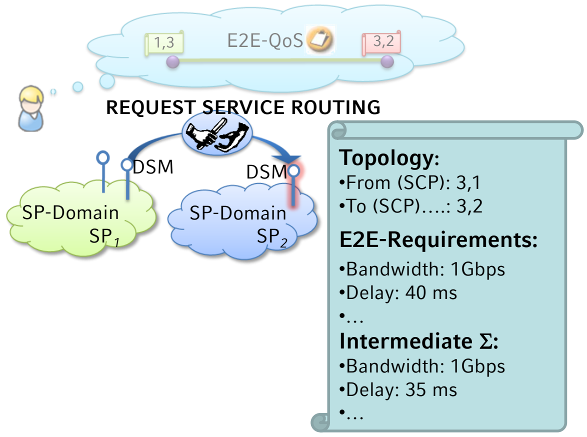

Figure 5. On-demand delegation of the routing task [1]

In our example, if SP_3 refuses to provide information about available connection segments, SP_1 can delegate the routing task to SP_2 (see Figure 5). In this case topology restrictions, customer specific end-to-end requirements as well as properties realizable by the found partial path are passed to SP_2. Advantages and disadvantages of *routing-by-delegation* are opposite to those of source routing: information about SPs’ available capacities are hidden and the selection of the most preferable route is guaranteed. On the other hand, communication becomes to be nested and information about previously checked alternate routes cannot be re-used, which might lead to redundant checks.

Our proposed routing procedure incorporates not only the principles for the selection of the next connection segment, but also communication between SPs needed for information requests and delegation of the routing task. These advanced considerations make it impossible to describe our procedure as a pure pseudo code alone. Instead we define a state-machine (see Figure 6), showing planning-phases as states and outcomes as transitions. Beginning with an intermediate SCP that is at the end of the path considered so far, the next connection segment is determined. If required information is missing, a planner enters an information-retrieval phase (A2) where the missing information is requested from a SP. After calculating the properties of the entire path, including the new next segment, these properties are evaluated against end-to-end requirements. If at least one of specified requirements is violated, an alternative segment has to be considered. If requirements are fulfilled, the distant SCP of the new segment is considered as a new intermediate SCP and the planner returns to the starting state (A1). We need to support the case that an information request is rejected. In this case, the routing task is delegated by a *FINDROUTE*-request to the last SP in the considered path (A5). Our procedure terminates, i.e., reaches an accepted end-state, either when the desired endpoint is reached or no more alternative connection segments are available.

As it is very common for the description of an algorithm to discuss its runtime analysis, please note that the main

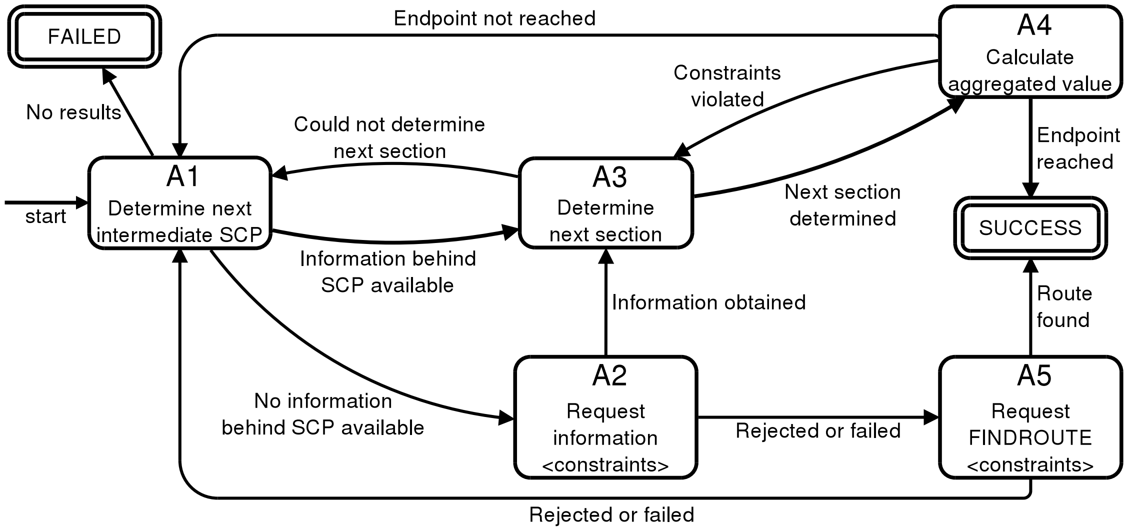

Figure 6. Routing processes

purpose of any algorithm is to achieve its goals. Therefore the quality of an algorithm should not only be evaluated through the performance of an implementation, but primarily through the quality and the properties of the yielded result. In order to provide such an evaluation of our proposal, we point at the order in which connection segments are considered for a new route. Regardless whether source routing or routing by delegation is used, consideration of available segments is identical to the Depth First Search procedure. In this procedure, alternatives are considered in an SP-specific order of preference. A criterion for considering an alternative is not fulfillment of at least one specified end-to-end requirement. Consequently, the proposed procedure is nothing else but an inter-domain policy-based routing procedure considering end-to-end requirements.

B. Tying management functionality together

Experience shows that quality guarantees are only possible to hold if the planned quality targets are leveraged by management procedures. The management of a service instance is in turn only possible if the management of all its integral parts is possible. Therefore two of the central purposes of management-aware routing are the definition of management functionality of each involved connection segment and its integration into the overall management of new multi-domain service instances.

In a multi-domain environment, direct access to the network infrastructure of any SP is in general not possible, as this is very likely to violate provider policies. Therefore the integration of management functionality means negotiation of *communication interfaces* and *responsibility areas* as well as *communication relationships* with all involved SPs. The possibility to negotiate communication relationships and responsibility areas for each connection should influence the acceptance of our approach positively, as we give SPs a framework through which they can control information- and command-flows. Negotiating DSM interfaces is an essential part of integrating management functionality and should be performed during the negotiation of other relevant connection segment properties like QoS parameters and management functionality.

Very important for tying together management functionality is the possibility to specify one or more DSM addresses by information requests. Representation of information by inter-domain communication will be discussed in Section VII. DSM address is presented in class COMMUNICATIONDSM (see Figure 11). This enables SPs to specify multiple communication interfaces for different purposes, e.g., regular information requests, monitoring or event-triggered notifications. Concerning the confirmation of requests, SPs providing connection segments should also specify communication interfaces for their management functionalities. Furthermore, this will guarantee flexibility to SPs to separate interfaces for service instance negotiation and operation, as well as to specify different interfaces for different customers and/or service instances.

Every SP providing a connection segment as a part of overall end-to-end service, is responsible to ensure the agreed quality targets are met and to provide management functionality through the negotiated interfaces. For the integration of management functionality we define that each SP is in charge, i.e., responsible, for the area for which it has performed the routing task. To illustrate this we refer to the example in Figure 4, outlined before. In the example the final path is going through the SCPs with IDs ”1,3”, ”1,1”, ”2,1”, ”2,2”, ”3,1”, and ”3,2”. Since routing has been delegated to SP_2, the example has multiple responsibility areas, depicted in Figure 7.

We propose to distinguish between two types of responsibility delegation. Following the first option, which we call *FullProxy*, in our example SP_2 would build an abstraction layer, hiding the entire remaining part of the path from SP_1. This means that both, build-up and tying of management functionality, is hidden by the proxy-SP and SP_1 has no direct control over them. The main advantage

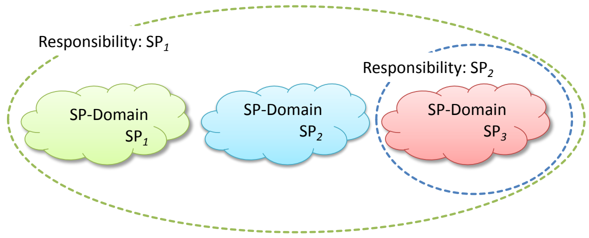

Figure 7. Responsibilities for function integration in example

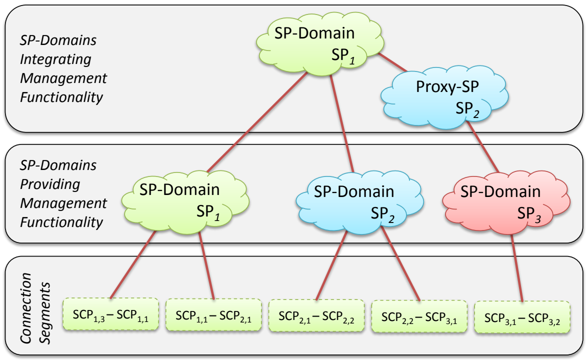

Figure 8. Communication relationships by *FullProxy* option

of this option is the best possible compliance with restrictive SP policies. Therefore we see this option (similar to the applicability of the *routing-by-delegation*) as an always feasible one. The disadvantages of this option are illustrated in Figure 8. The most apparent disadvantage of this option is the increase of communication hops between SPs, due to nested communication relationships, which will lead to increased reaction times. Another significant disadvantage is the increase of intermediate processing by proxy-SPs. Finally we want to point out the inevitable deviation of reaction time by connection parts with direct connection of management functionality, e.g., negotiated using *source routing*, and connection parts with nested integration of management functionality, e.g., negotiated using *routing-by-delegation* with *FullProxy* option. Such deviation not only decreases the overall end-to-end performance, but also might require special treatment with respect to the realization of such a solution.

Therefore we see the necessity to introduce a second delegation option, which we call *TransparentProxy*. Using this option it is possible to signal the proxy-SP that the communication interfaces for management functionality specified by the requester should be re-used during reservation and ordering of the delegated part of the path. The biggest advantage of this option is the possibility to keep communication relationships as flat as possible. As SP-policies might restrict acceptable communication relationships, e.g., only to neighboring SPs, this option may not always be feasible. Therefore, we see the *FullProxy* option as a fallback solution, similar to *routing-by-delegation* being a fall-back solution for the case when *source routing* is not applicable.

VI. Generic QoS-Operations

During the routing two operations are needed: 1) the values of connection segments considered as a part of the route should be aggregated; and 2) the aggregated value should be compared with the E2E requirements. All these operations have to be performed on the properties (and property combinations), which are relevant for the particular customer request. In Section VII we will show, how various connection properties (and their combinations) can be described in a similar fashioned way. Unfortunately, the similar fashioned description alone is not enough for the similar fashioned operations on those properties. The reason is the various semantics of the values related to various connection properties. For example, operations needed for aggregation and comparison of *delay* values are not applicable for other very important QoS parameter – *bandwidth*. The same can be said about operations on other properties, e.g., on *maintenance window*.

In order to deal with this problem, in [2] we have defined a generic function schema, which enables treatment of various connection properties and their combinations in a similar fashioned way. The basic idea is that all parameters are distinguished bases on global unique IDs. Further, all IDs for every supported qualitative and quantitative QoS parameter as well as for management functionality and its property should be defined in a registration tree. This opens possibilities to specify for every property ID a set of functions needed for operations on it (see Figure 9). The functions _AGGREGATE_LINKS and _ORDER_COMPARE are used to aggregate and compare values of a single property. These functions operate on properties of connection segments, and are therefore relevant for the routing procedure described in Section V. The reasons for the remaining three functions are slightly more complex. Due to restrictive SP policies, access to the network equipment is generally impossible from outside of one's own SP. Function _AGGREGATE_LINKPARTS is needed for computation of properties by segments interconnecting two neighbor SPs (so called *Inter-Domain Links*). This function is needed because in general no direct access to the network infrastructure of the neighbor SP is possible, and consequently both SPs have only own (partial) knowledge about properties, which can be guaranteed by a particular Inter-Domain Link. The remaining two functions (_SELECT_BEST and _SELECT_WORST) are needed for operation on value ranges, which can be specified by SP as realizable connection property. Value ranges can be used instead of specifying multiple connection segments between same two SCPs with associated fixed values. For a more elaborated discussion about the needs and application areas of these functions please see [2] and [3].

The association of these functions with property IDs enables the similar fashioned operations on various properties. This means that during the routing for every relevant property the corresponding functions can be obtained based on property ID. Subsequently, these functions can be used to operate on the values of the properties, which have been specified for connection segments.

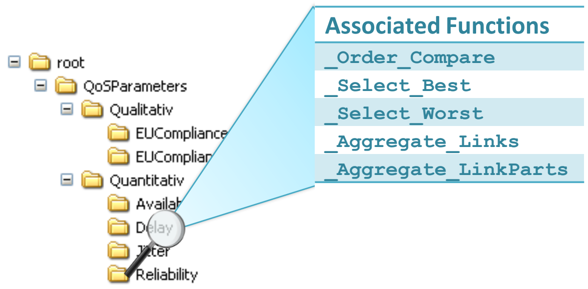

Figure 9. Association of functions with a property ID in the registration tree [2]

The described functions are always associated with a single property. In order to operate on the customer specific property combinations, we derive operations on multiple properties based on single-property operations. The aggregation of multiple properties is straight forward and is defined as a combination of per-property aggregations. The comparisons of property combinations is, however, more complicated. The reason is the possibility that by comparison of two value-combinations some of the properties in the first and some other of the second combination are better. This situation is reflected in the definition for comparison of value combinations (see Figure 10). The "≺" symbol is used here to denote which value is better.

$$\overrightarrow{Compare}(\vec{U},\vec{V}) ::= \begin{cases} =, \text{ if } & \forall 1 \le i \le m : u_i = v_i \\ \prec, \text{ if } & \forall 1 \le i \le m : (u_i \prec v_i \\ & \quad \vee\ u_i = v_i) \wedge \\ & \exists 1 \le j \le m : u_j \prec v_j \\ \succ, \text{ if } & \forall 1 \le i \le m : (u_i \succ v_i \\ & \quad \vee\ u_i = v_i) \wedge \\ & \exists 1 \le j \le m : u_j \succ v_j \\ \neq, \text{ if } & \exists 1 \le i \le m : u_i \prec v_i \wedge \\ & \exists 1 \le j \le m : u_j \succ v_j \end{cases}$$

Figure 10. Comparison of property-combinations [2]

Even though in the presented paper these functions have been motivated by their necessity during the routing, their application area is much broader. For instance, they can be used for the monitoring of already established connections and the comparison of monitored values against designated targets.

The definition of similar fashioned operations on connection properties has several advantages. Among the most important is the easiness of extensibility for the support of new connection properties. Furthermore, by extending property support no changes or adaptation in implemented routing algorithm will be needed. And finally, if a global registration tree is used among all SPs for property IDs and function definition, operations on property values and their combinations will be identically among all SPs.

VII. Information Model

The basis for every routing is the knowledge about available connection segments and their properties. As described in Section V, during the routing missing information has to be requested from the other SPs. In order to comply with SP's information policies, the requested information has to be restricted strictly to the needed information. Figure 11 describes how the requested properties of the segment can be specified. In order to support user-specific E2E requirements, various combinations of relevant properties can be used within the RequestedProperties class. We distinguish between tree kinds of properties: Qualitative QoS parameters, quantitative QoS parameters, and management functionality (reflected correspondingly in classes QualitativeQoS, QuantitativeQoS, ManagementFunctionality). The distinction between various properties is made based on globally-unique IDs. This is a basis for a similar-fashioned treatment of various parameters, as it is described in Section VI. With management functionality class combination of properties (class Property) can be associated. This class is used for specification of different management function specific properties, e.g., polling interval for monitoring. Similarly to QoS parameters, the distinction between properties is made based on its IDs. Finally, values (class AssociatedValue) can be associated with quantitative QoS parameters and with properties of management functionality.

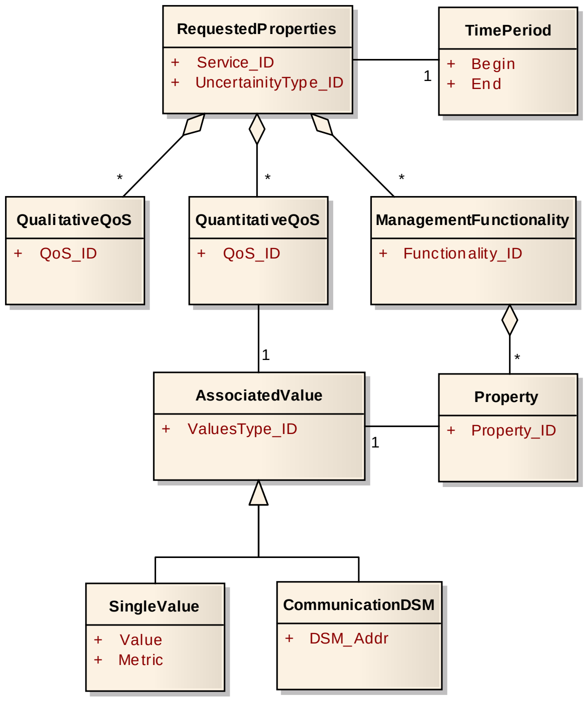

Figure 11. Specifying properties by service instance requests

If SP which can provide connection segments accept information request, it has to provide own up-to-date information. In our work we assume that every SP has exact knowledge about its own network equipment and available capacity. We further assume that every SP can obtain information about capacity available for interconnection with the neighboring SPs. We see this as a very realistically assumption for most SPs, as such information – among other – is required for proper management of the own network infrastructure. Further we assume that every SP can determine the quality properties which can be guaranteed over available own connections.

The description of information is significantly more complex then information request, as it also should describe various topological properties (see Figure 12). As has been discussed at the beginning of this article, an expectation that SPs will share the information needed for management of the own infrastructure (i.e., detailed network topology representation as well as capacity and usage of particular physical network connections) can be seen as a very unrealistic one. One of the most critical aspects in a multi-domain environment is very restrictive information policies. The omission of the collision with those policies is the main reason, why the proposed routing procedure operates on information at the abstraction level of an SP's organizational domain. Such a representation hides most realization aspects, like network topology, which strongly complies with the mentioned information policies.

For the description of available connections three classes are used: CompoundLink, ComponentLink, and ComponentLinkPart. The class CompoundLink is used as a wrapper for multiple alternative connection segments connecting the same two SCPs. Such a situation can occur, for instance, due to alternative physical connections available between these SCPs. Even though such details are hidden from SP's outsiders, they can result in connection segments with different properties. The class ComponentLink, which represents the connection segments, is associated with the class LinkProperties, which specify segment's properties. The structuring of connection segment properties is identical to the one of information request (see Figure 11). Finally, the class ComponentLinkPart represents the SP's view at an Inter-Domain Link. As discussed in Section VI, every SP has only information about quality guarantees realizable at the SP-facing side of an inter-domain connection. Therefore, the whole quality of such connections has to be calculated from two views on it of involved SPs. For this purpose function _Aggregate_LinkParts has to be associated with connection properties (see Section VI).

Every of these three connection classes interconnect two SCPs, which contain their globally-unique IDs needed for segment stitching as described above. Further, every SCP is associated with a SP Domain in which it is located. The reason is the necessity of SP responsible for routing to communicate with those SPs, which can provide connection segments. During the routing, considering a new connection segment as a part of the route implies that the next connection segment should be adjacent to the distant SCP of the considered route. Information about such an adjacent segment can only be obtained from an SP owning the mentioned SCP. In close collaborations of very limited amount of SPs it can be assumed that every SP knows the interface to communicate with every other SP. In open and/or highly dynamic collaborations this is not the case. We however assume that every SP can maintain up-to-date information about communication interfaces of all its neighbors. Therefore by specifying the domain owning the SCP, a communication interface DSM_Addr of this SP should always be provided.

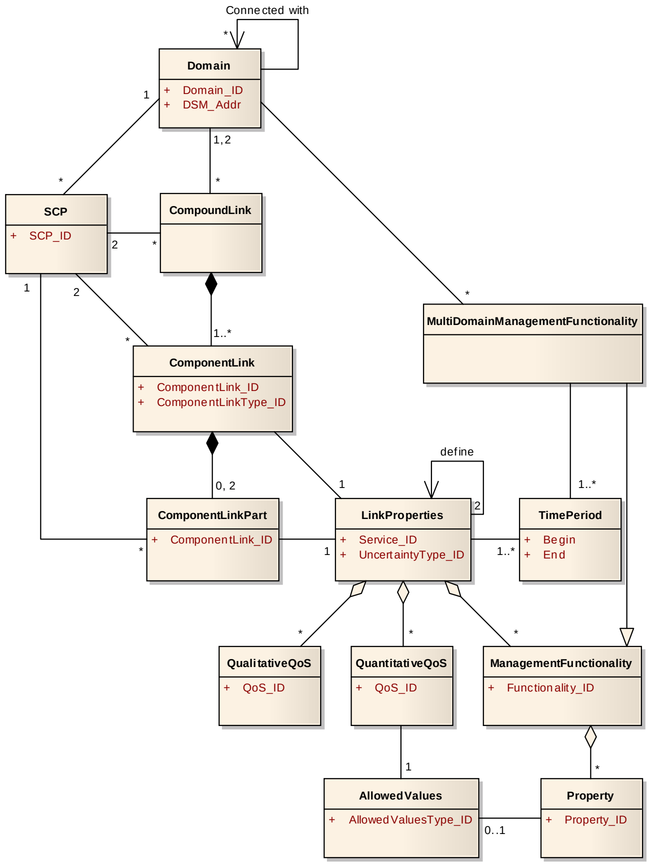

Figure 12. IM for available connections and properties of a single SP [3]

Finally, the MultiDomainManagementFunctionality class has to be mentioned. This class is derived from ManagementFunctionality, which enables the specification of multi-domain management functionality in a similar fashioned way as it is done by the management functionality of connection segments. This description is needed by the delegation of multi-domain management functionality, as it will be described in Section IX-A.

The inter-domain routing procedure, which we have described in Section V, operates on a semi-global knowledge, i.e., the knowledge spanning over multiple – but not necessarily all – SP domains. At the same time, each SP is only able to provide information about connection segments in which realization it is involved. This raises the question, how such single-domain views of different SPs can be combined to semi-global multi-domain view?

If every SP can provide its view at the connection segments it is involved in, deriving a multi-domain view from a multiple single-domain views can be realized relatively easily. We propose that every SP, when specifying segments, always provides the IDs of two SCPs at its ends. It is necessary that these IDs are globally unique. Only then the SP responsible for routing can "stitch" these segments at the SCPs with identical IDs together (see Figure 13).

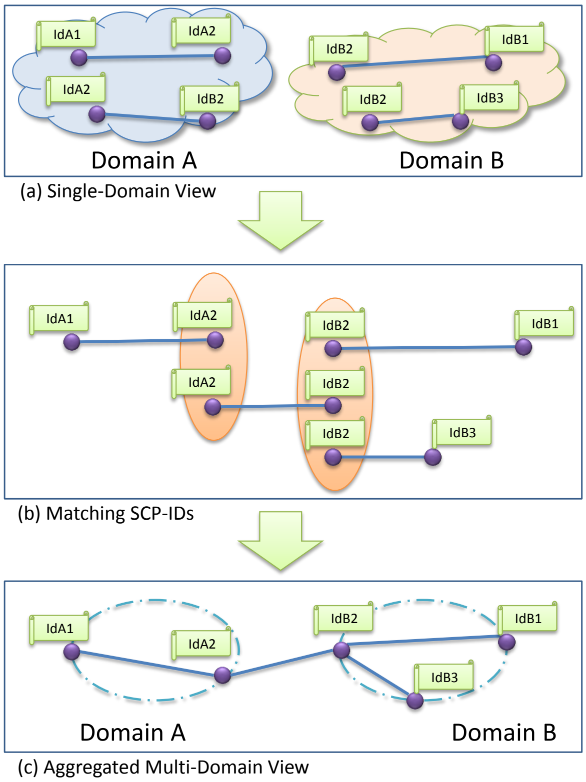

Figure 13. Deriving a multi-domain view from single-domain views [3]

A more elaborated discussion about details of information model as well as about reasons for the selected representation can be found in [3].

VIII. Putting it all together

In in Sections V, VI, and VII we have described the three building blocks of our proposal. In order to illustrate how elaborated concepts can be applied, we reuse the example outlined at the beginning of this article. We recall that in the example the customer has contacted SP_1 and has requested a connection between SCPs with IDs "1,3" and "3,2" with following properties:

- Bandwidth: 1 Gbps
- Delay: max 40 ms
- Maintenance
  - Window: 6^{00} – 10^{00} GMT
  - Duration: max 1 hr.
- Monitoring

From the customer's point of view, SP_1 is the sole provider of the E2E connection with the mentioned properties. For the planning of a corresponding multi-domain path SP_1 needs routing functionality. Furthermore, in order to comply with the outlined parameters also multi-domain *Monitoring* functionality is needed. As SP_1 can provide both functionalities, it can start the process of finding the route fulfilling all end-to-end customer requirements.

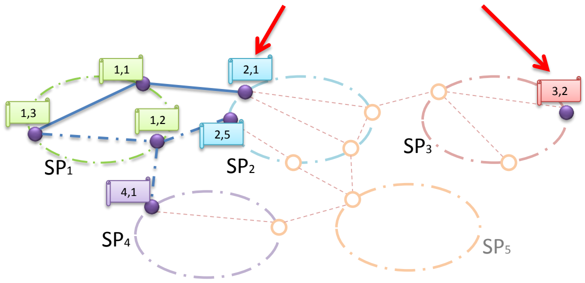

Figure 14. Global view at the available services

Figure 14 presents the situation after SP_1 chose the path through its own network. It still has to find the route between SCPs "2,1" and "3,2", but has no up-to-date knowledge about connections available in SP_2. Consequently a request for information should be send to SP_2. As proposed in Section V-A, this information request should be accompanied by the topology- and property-restriction. We recall that the main motivation of restricting requested information was the necessity to comply with very restrictive information policies. An additional positive effect is the significant reduction of connection segments, which have to be considered for the route. This in turn speeds up the overall routing process.

```
<RequestedProperties Service_ID="CS" UncertainityType_ID="Guaranteed">
  <TimePeriod Begin="Now" End="OpenEnd"/>

  <QuantitativeQoS QoS_ID="Bandwidth">
    <AssociatedValue ValueType_ID="SingleValue" Value="1" Metric="Gbps"/>
  </QuantitativeQoS>

  <QuantitativeQoS QoS_ID="Delay">
    <AssociatedValue ValueType_ID="SingleValue" Value="140" Metric="ms"/>
  </QuantitativeQoS>

  <ManagementFunctionality Funcitonality_ID="Maintenance">
    <Property Property_ID="Window">
      <AssociatedValue ValueType_ID="ValueRange" Min="0300" Max="0700" Metric="GMT"/>
    </Property>
    <Property Property_ID="Duration">
      <AssociatedValue ValueType_ID="SingleValue" Value="1" Metric="hr"/>
    </Property>
  </ManagementFunctionality>

  <ManagementFunctionality Funcitonality_ID="Monitoring">
    <Property Property_ID="CommunicationDSM">
      <CommunicationDSM DSM_Addr="dsm.domainSP1.com/mon/instance932/"/>
    </Property>
  </ManagementFunctionality>

</RequestedProperties>
```

Figure 15. Specifying needed properties in information requests

In Section VII we have presented how requested properties can be structured (see Figure 11). For communication between SPs, this structure has to be encoded. Currently, application-layer protocols are often built on top of SOAP communication. As the information in such communication is encoded in a human-readable XML format, this suits very well for illustration of concepts. Figure 15 presents how the example specific information restrictions can be represented in XML according to the proposed structure. Please note that the proposed property structure can be encoded differently.

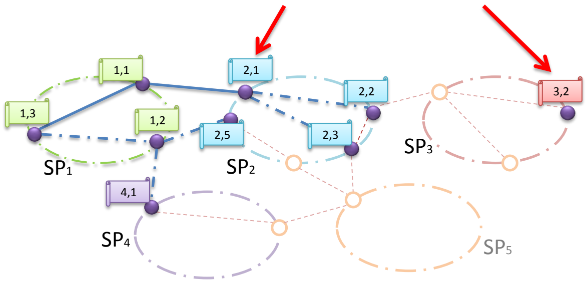

Figure 16. Global view at the available services

The structuring of the requested information has been described in Section VII (see especially Figure 13). By information request SP_1 also specify its DSM address for monitoring functionality. This is a crucial aspect in order to tie functionality together, as it has been described in Section V-B. The representation in XML format can be done similarly to the presented request. Please note that regarding the answer, SP_2 should provide own DSM-address(es) through which the monitoring functionality of available connection segments can be reached during instance operation. These addresses might vary between defferent service instances due to various SP-internal reasons.

From	To	Route-Part?	Property	Value
$SCP_{1,3}$	$SCP_{1,1}$	✓	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 19 ms 06:00-10:00 GMT 1 hr
$SCP_{1,3}$	$SCP_{1,2}$	-	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 15 ms 06:00-10:00 GMT 1 hr
$SCP_{1,1}$	$SCP_{2,1}$	✓	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 4 ms 06:00-10:00 GMT 1 hr
$SCP_{1,2}$	$SCP_{2,5}$	-	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 3 ms 07:00-09:00 GMT 1 hr
$SCP_{1,2}$	$SCP_{4,1}$	-	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 6 ms 06:00-07:00 GMT 1 hr
$SCP_{2,1}$	$SCP_{2,3}$	-	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 20 ms 06:00-08:00 GMT 40 min
$SCP_{2,1}$	$SCP_{2,2}$	-	Bandwidth Delay Maintenance Window Maintenance Duration	1 Gbps 10 ms 07:00-09:00 GMT 1 hr

Table I
KNOWLEDGE TABLE OF SP_1

According to the routing process specified in Figure 6, up to now we have selected SCP "2,1" as the next intermediate SCP (action A1), have recognized that some information is missing, and therefore requested this information (action A2). Now we have to determine, which connection segment should be considered as next section in the path (action A3). Figure 16 depicts graphically the knowledge of SP_1 after SP_2 has provided the requested information. Please note that the information about irrelevant connection segments has not been provided at all. In order to illustrate the follow-up actions of the defined routing procedure, in Table I we present the knowledge base available to SP_1 at this point of procedure. It contains imaginary values associated with properties of available segments as well as the flag whether the particular segment is considered as a part of the route or not. Please note that for SP_2 there is no need to provide information which is not fulfilling the restriction.

According to the information provided by SP_2, two connection segments are available. The most preferable one is the connection between SCPs "2,1" and "2,3". According to the defined routing procedure, the SP responsible for routing has to consider available connection possibilities in the most-to-less preferable order of the SP providing them. Therefore SP_1 has first to consider this connection possibility and therefore calculate the aggregated value of the partial route (action A4 in Figure 6).

In order to calculate the path properties, aggregation functions for the relevant properties have to be accessed first. As described in Section VI, these functions are associated with the property IDs. These functions can be used for semantic-aware operations on the values of particular connection properties. This means that aggregation function is:

- *min* for the Bandwidth
- *addition* for Delay
- *intersection* for Maintenance window
- *max* for Maintenance duration
- Logical *AND* for availability of Monitoring

Using these functions and values from the knowledge table we can calculate the properties of the intermediate path going through SCPs "1,3", "1,1", "2,1", and "2,3" (Action A4 in Figure 6). Now we have to compare the calculated multi-property value with the end-to-end customer requirements. This is done again based on the functions defined for the particular properties (see Section V). The result of the comparison will be $\neq$, as the all properties except delay are better than the requirements, but the requirement for delay cannot be fulfilled by this path. Consequently, the next alternative connection segment, e.g., between SCPs "2,1" and "2,2" has to be considered according. As the path going through SCPs "1,3", "1,1", "2,1", and "2,2" hold the requirements, its last SCP "2,2" is chosen as next intermediate SCP (action A1) and the outlined procedure can be repeated.

Let us assume that with the above outline procedure the path fulfilling all end-to-end requirements and going through SCPs "1,3", "1,1", "2,1", "2,2", and "3,1" have been found (see Figure 17). This means that the following segment(s) can be only provided by SP_3. As the information about available connection segments is missing, SP_1 has to request this information from SP_3 (action A2 in Figure 6). Let us further assume that SP_3 rejects this information request due to some internal reasons. For this case our proposal foresees a delegation (action A5) of a routing task to the last SP in the already found path, i.e., in this case to SP_2. Graphically this situation is depicted in the Figure 5.

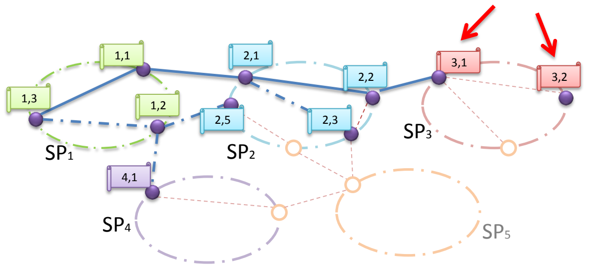

Figure 17. Global view at the available services

If SP_2 accepts the request for routing delegation, it takes over the routing for the remaining part. It receives information about end-to-end requirements, values which can be guaranteed by the intermediate route and two SCPs between which it has to perform routing. SP_2 choses SCP ”3,1” as its intermediate SCP (action A1) and proceeds as outlined before. If the end point is reached, SP_2 reports the properties of the whole found path to the requester of the routing task.

Please note that in order to tie management functionality together, SP_2 has to specify to SP_3 the communication DSM (see Section V-B). In the case of FULLPROXY delegation, SP_2 specifies its own address. In this case, if the route could be found, it also has to report its own DSM to SP_1. The provider SP_2 should further aggregate monitoring information of all following segments in its own responsibility area. By this delegation option, SP_2 provides SP_1 exactly with the aggregated view.

Regarding the delegation by using the TRANSPARENT-PROXY option, during the routing in its own responsibility area, SP_2 should specify for monitoring purposes the DSM address of SP_1. Also, concerning the acceptance of SP_3 to provide the last segment, SP_2 should report the DSM of SP_3's monitoring instance to SP_1. In this case SP_1 is only responsible for routing but not for aggregation of monitoring information.

IX. DISCUSSION

In this section we will outline our thoughts about aspects going beyond the scope of the presented paper, but highly related to the presented solution. At first we will discuss the delegation of management tasks to the trusted third party SPs. Then in Section IX-B we will present our thoughts about reservation and ordering of planned route. Section IX-C is dedicated to outline our view at the acceptance of the presented solution by SPs. Finally, in Section IX-D we will present positioning of our solution among other existing alternatives.

A. Delegating Multi-Domain Management Tasks

The discussion up to now has implicitly presumed that every SP-domain approached by the customer as well as every proxy-SP can take over the whole functionality, like routing and integration of needed management functionality. The practicability of such solution depends on different factors. Most important are the complexity and costs for the development and maintenance of such solution by every SP. To a large extent this might be influenced by the utilization of the multi-domain functionality through requests for new and operation of established E2E connections.

In practice, two concepts have been often successfully combined in order to cope with high complexity and costs: specialization and sharing. A per se example is the DNS service, which is realized by specialized SPs and shared among SPs specializing on, e.g., content provisioning. Therefore we propose to support the delegation of multi-domain management tasks. As most common multi-domain management tasks, the following should be mentioned: routing, monitoring, accounting, or outage localization. Figure 18 depicts the delegation of the routing task. Instead of performing routing by themselves, as this was defined in our original proposal, provider SP_1 can delegate this task to SP_R which should follow the management-aware routing procedure. Please note that even if the trustworthiness of such specialized multi-domain management service providers might be very high, especially in large collaborations it does not necessary eliminate the necessity of task-delegation to proxy-SPs.

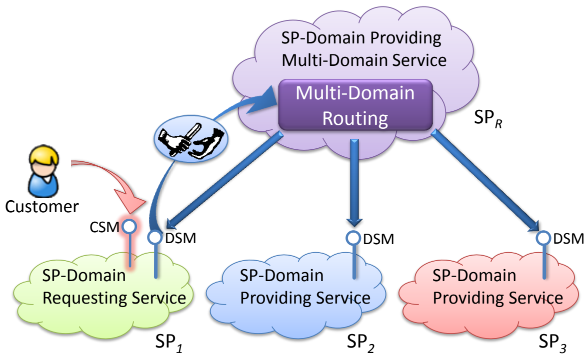

Figure 18. Delegating multi-domain functionality to a third party

B. Route Reservation and Ordering

In packet switching, like in IP-networks, the routing is an all-including procedure, as it is combined with the packet forwarding to the next hop which will be then in charge for further routing and forwarding. In line switching, the situation is more complex and routing can be seen as a general plan which still has to be implemented. If the plan is not sufficient, than even perfect implementations of it will not lead to the desired result; but even if the plan is perfect, it can be undermined by its bad implementation. Therefore we see the described management-aware routing as the most critical – but not sole – prerequisite for provisioning of multi-domain network connections compliant with user-specific E2E requirements.

As we consider a multi-domain environment, the information which is used during the routing is not only semi-global, but also can be considered as not 100% up-to-date. This is mainly caused by concurrent requests, which can come from (in general various) SPs responsible for routing of different connections. On the one hand, various connection segments considered for these connections might be realized upon the same physical infrastructure. On the other hand, as a means for reduction of capital expenditure for infrastructure, every SP is interested in maximizing resource usage. Consequently, in the time between the information request and the actual ordering of the selected connection segments, the needed physical resources might be assigned to other service instances. In other words, we face the multi-domain concurrency for the same limited resources.

Various elaborated techniques are available for dealing with the concurrency issue. The most common way to avoid conflicts and/or minimize deadlocks is the reservation of the resources before their ordering. This technique can be also applied to concatenated services at the SP abstraction layer (see Figure 19). Every node in Figure 19 represents the state of one service part, e.g., a connection segment, from the perspective of SPs responsible for the provisioning of service instance. As the communication between different SPs can only be performed via some communication protocol, the transitions arrows between states are described with different kinds of requests. Please note that these names only reflect the semantics of requests and do not imply any suggestions for protocol implementation.

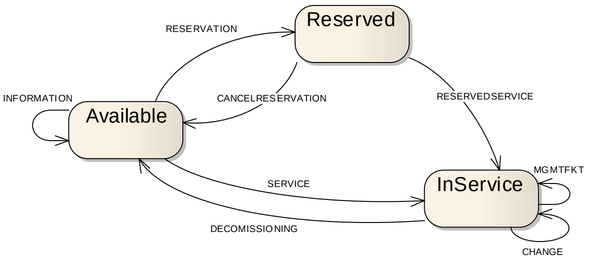

Figure 19. State transition at multi-domain abstraction level

Even though the state transition by itself is quite common, two additional aspects have to be specified: 1) behavior of SPs regarding requests for state transition of a single service part; and 2) the way to apply such requests to multiple service parts.

Concerning the definition of SP's behavior we propose to reuse elements of bilateral and trilateral negotiation models, as they are described in [34]. More precisely, if SP_{RS} (RS for "Requesting Service") requests from SP_{PS} (PS for "Providing Service") the reservation of connection segment with a certain quality, SP_{PS} should be able to confirm the reservation with equal or smaller quality as requested. This should cover the case if SP_{PS} in the time between information- and reservation-requests has already reserved some needed resources for a different service instance. After the reservation is confirmed, for each reserved service part, SP_{RS} should be able to request its ordering. In this case, however, SP_{PS} should be either able to provide the quality equal or better than the one confirmed by reservation, or report the failure of ordering.

The outcome of the management-aware routing procedure is a combination of connection segments along the route associated with the apparently feasible QoS thresholds. As all SPs providing those segments are also known, simultaneous communication with all those SPs would be possible. Concerning the reservation of service segments we however argue against the simultaneous communication with all SPs. In this case, reduced quality by reservation of a sole segment would automatically lead to a violation of the E2E user requirements. Instead, we propose to reserve segments sequential one-by-one from one endpoint to another one. In the case that SP_{PS} confirms reduced qualities, SP_{RS} can try to balance the performance degradation in one segment by increasing the planned thresholds of remaining segments. For this purpose information obtained during the routing by information requests can be reused. If the balancing is not possible, a re-routing for the remaining part of the path can be performed. If even this fails, SP_{RS} should cancel reservations of all connection segments which have been already successfully reserved. In order to minimize reservation time, messages about cancelling reservation can be sent simultaneously to all relevant SPs.

If reservation of all segments has been successful and E2E requirements are met, requests for ordering of these segments can be sent to corresponding SPs. As – corresponding to the proposed behavior – SPs may only provide better quality than the confirmed by the reservation, requests for ordering can be sent simultaneously to all SPs. If at least one of SPs cannot fulfill its commitment, it will report failure to SP_{RS}. In this case SP_{RS} has to cancel reservations of all remaining segments immediately.

Leveraging the proposed procedure for reservation and ordering of connection segments along the found route we score two goals: 1) first of all the fulfillment of E2E requirements; and 2) minimizing the time needed for resource reservation before a new service instance can become operational.

C. Examine SP-Acceptance for Elaborated Solution

As discussed in Sections I and II, the most critical requirement for an inter-domain routing approach is its acceptance by SPs. In our proposal, this real world requirement has been reflected by considering SP interests and restrictions.

As a proof of concept we refer to our experience within the context of the ***I**nformation **S**haring across **H**eterogeneous **A**dministrative **Re**gions* (I-SHARe) activity [32]. This activity has been established in Géant in order to foster information exchange during the manual planning and operation of E2E Links. During the first phase of this activity, requirements and constraints for the (back then) upcoming tool had to be captured and categorized. This has provided us with a deep insight into demands and concerns of network service providers.

First of all the sharing of detailed network information (such as network topology or overall available capacity) is commonly seen as unacceptable. At the same time it is acceptable for SPs to exchange information about connection segments at SP abstraction level, which are possible for a particular planned E2E Link. Also exchange of more detailed information with neighboring SPs is very common by connection planning. The information model described in Section VII directly reflect these aspects. As our proposal strives for automatic route planning, we expect that multiple alternative service parts should be specified in the order of the SPs' preferences. Primarily, it should simplify information management and also reduce the necessary number of communication steps. Several interviews with operators involved in the planning of Géant E2E Links have shown the acceptance of SPs to provide information about multiple service parts during the routing process, even it is not used in the established manual processes.

By gathering of established in Géant manual route planning processes one further very important constraint has been identified. The only true concern that has been repeatedly mentioned by operations of different SPs is the compliance of the service part choice to the SP preferences. The enforcement of this aspect has been reflected in both routing modes outlined in Section V.

As our proposal can provide routing planning under consideration of SP constraints, we are confident that our solution can be broadly accepted by SPs. As it also considers the customer-specific E2E requirements and management functionality, our solution opens for SPs possibility for delivering of high-quality network services to the customers.

D. Application areas and possible adaptations

Compared to the established connectivity-oriented routing approaches, the proposed solution requires significantly more information, computations, and – which is the most time-consuming part – inter-domain communication. This statement is applicable not only to routing, but also to the subsequent reservation and ordering processes. Therefore, it cannot be considered as a routing procedure for a mass service, as it is the case by IP routing in Internet. Instead we suppose the application area of the developed algorithm to be in the middle-scale niche between mass services, which are focused on a pure connectivity with best-effort quality, and carrier-grade connections, which are mainly manually planned very-long-term connections with dedicated quality specified in contracts (see Figure 20).

The proposed solution can provide near-real-time planning and ordering of a route with customer-tailored E2E connection properties. This solution is applicable to scenarios like video-on-demand, video-conferencing, on demand connectivity for Grid or cloud collaborations, and other areas in which customers demand (and often are willing to pay) not only for the pure connectivity but also for the connection quality. Especially this is applicable for scenarios like video-conferencing, where bad quality

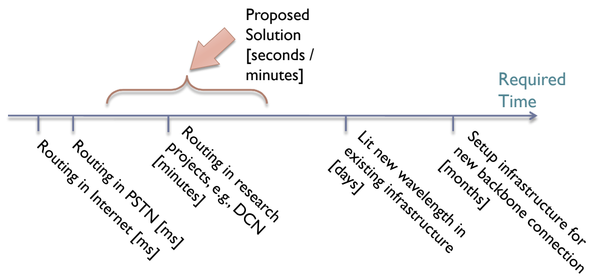

Figure 20. Positioning of proposed solution

of one parameter (e.g., jitter) cannot be covered by another one (e.g., bandwidth).

Depending on the offered service and/or the specifics of the SP cooperation, the proposed algorithm can be modified in order to improve its scalability. Particularly, we would like to outline the following two cases:

- If the connection service is only offered with QoS parameters that do not require E2E consideration of all involved parts, e.g., bandwidth or data encryption, a *routing-by-delegation* approach can generally be used. Especially in combination with the simultaneous resource reservation, it can prove to be very scalable. A very good example for this strategy can be seen in telephone connections, which offer constant bit rate and low jitter.
- In the case of a small and especially very tight SP cooperation, the routing instance can be centralized. In this case only this instance performs the whole E2E routing for all new connection requests, which neglects the concurrence between simultaneously ordered service instances. Consequently, the service part reservation can be omitted or performed simultaneously. A similar strategy with a two-level routing instances is used, e.g., in the DCN cooperation. This approach corresponds to the *source routing* approach (see Section V), in which the routing task is always delegated to a central instance. The applicability of this approach, however, depends directly on the willingness of SPs to provide complete information about all available service parts to the routing instance and to accept its inter-domain manager role.

A further simplification can be applied for operations on supported properties. If only few properties have to be supported and no extension of support is planned for the future, the usage of a registration tree for the definition of operations on properties can be omitted. This should significantly improve the performance of the implemented solution. On the other hand, support of not yet considered properties becomes an issue again.

As a conclusion, the proposed solution is applicable in the most challenging case of an open SP collaboration as well as with a variety of customer-specific QoS parameters.

X. Future work

As a mean to quality assurance for Concatenated Services, we have elaborated a novel conceptual solution consisting of three technology-independent building blocks. As the next logical step we see the evaluation of existing technologies, whether and how they can be used to implement the proposed solution. The desired outcome of this evaluation is a proposal for implementation of the presented solution and its building blocks. In the case if not all concepts can be implemented, the missing functionality can be seen as requirements for future technology development.

The proposed solution covers only the connection planning aspect during the ordering phase of service instance life cycle. In order to support SP collaboration during further phases, we plan to elaborate a proposal for inter-domain management processes. For instance, the methodologies for E2E monitoring have to be defined, which is needed for both monitoring of commitments' fulfillment and for problem localization. We see such processes as an essential part of the overall solution for CSs.

As the occurrence of outages cannot be fully excluded, strategies for handling them have to be elaborated. Therefore we plan to investigate the applicability of self-adaptation techniques as a strategy for the multi-domain compensation of a single-domain quality reduction or outages of service parts.

An additional aspect, which should not be neglected, is the consideration of security in multi-domain environments. Therefore the integration of federated identity and trust management into the management of CSs will be a part of our future research.

As stated at the beginning of this article, our work has focused on dedicated point-to-point connections. For the future, we also plan to evaluate the applicability of the developed solutions to the establishment and management of point-to-multipoint as well as multipoint-to-multipoint network connections with quality assurance.

Acknowledgment

The authors wish to thank the members of the Munich Network Management Team (MNM Team) for helpful discussions and valuable comments on previous versions of this paper. The MNM Team directed by Prof. Dr. Dieter Kranzlmüller and Prof. Dr. Heinz-Gerd Hegering is a group of researchers at Ludwig-Maximilians-Universität München, Technische Universität München, the University of the Federal Armed Forces and the Leibniz Supercomputing Centre of the Bavarian Academy of Science. See http://www.mnm-team.org.

References

[1] Yampolskiy, M., Hommel, W., Lichtinger, B., Fritz, W., Hamm, M.K. Multi-Domain End-to-End (E2E) Routing with multiple QoS Parameters. Considering Real World User Requirements and Service Provider Constraints. In *Proceedings The Second International Conference on Evolving Internet, 2010. INTERNET 2010,* Valencia, 2010.

[2] Yampolskiy, M., Hommel, W., Schmitz, D., Hamm, M.K. Generic Function Schema for Operation on multiple Network QoS-Parameters. Submitted to *The Second International Conference on Advanced Service Computing, 2010. SERVICE COMPUTATION 2010,* Lisbon, 2010.

[3] Yampolskiy, M., Hommel, W., Marcu, P., Hamm, M.K. An information model for the provisioning of network connections enabling customer-specific End-to-End QoS guarantees. In *Proceedings 7th IFIP/IEEE International Conference on Services Computing, 2010. SCC 2010,* Miami, 2010.

[4] Dynamic Circuit Network, Publications. http://www.internet2.edu/network/dc/publications.html [10 May 2010].

[5] Internet2 How to Connect: Internet2's Dynamic Circuit Network. http://www.internet2.edu/pubs/DCN-howto.pdf White paper 2008.

[6] ESnet On-demand Secure Circuits and Advance Reservation System (OSCARS), Publications. http://www.es.net/OSCARS/ [10 May 2010].

[7] Dynamic Resource Allocation via GMPLS optical Networks (DRAGON), Publications. http://dragon.maxgigapop.net/twiki/bin/view/DRAGON/Publications [10 May 2010].

[8] Phosphorus, Publications. http://www.ist-phosphorus.eu/publications.php [10 May 2010].

[9] Automated Bandwidth Allocation across Heterogeneous Networks (AutoBAHN), Bandwidth on Demand Deliverables. http://www.geant2.net/server/show/nav.1914 [10 May 2010].

[10] Joint Techs Winter 2009 TAMU DCN Demo. https://spaces.internet2.edu/display/DCN/Joint+Techs+Winter+2009++TAMU+DCN+Demo [28 May 2010].

[11] DICE (DANTE-Internet2-CANARIE-ESnet) Collaboration. http://www.geant2.net/server/show/conWebDoc.1308 [28 May 2010].

[12] Proposed architecture for inter-domain circuit monitoring. http://code.google.com/p/perfsonar-ps/wiki/CircuitMonitoring [28 May 2010].

[13] Yampolskiy, M., Hamm, M.K. Management of Multidomain End-to-End Links; A Federated Approach for the Pan-European Research Network Géant 2. In *Proceedings 10th IFIP/IEEE International Symposium on Integrated Network Management, 2007. IM'07,* Munich, 2007; 189-198.

[14] LHC - The Large Hadron Collider Homepage. http://lhc.web.cern.ch/lhc/ [10 May 2010].

[15] DEISA - Distributed European Infrastructure for Supercomputing Applications Homepage. http://www.deisa.eu/ [10 May 2010].

[16] Aptet, E., Chevers, J, Garcia Vidondo, M., Tyley, S. *Géant Deliverable DS2.0.3,3: Report on GÉANT2 Ad-vanced Services - Lambdas and Switched Optical* Géant2 technical paper 2009.

[17] Schauerhammer, K., Ullmann, K. Operational Model for E2E links in the NREN/GÉANT2 and NREN/Cross-Border-Fibre supplied optical platform. www.geant2.net/upload/pdf/GN2-06-119-OPConcept.pdf Géant2 technical paper 2006.

[18] Office of Governance Commerce (OGC) *IT Infrastructure Library (ITIL)*. London, 2007.

[19] TeleManagement Forum (TMF) *New Generation Operations Systems and Software (NGOSS)*. 2004.

[20] Hamm, M.K. *Eine Methode zur Spezifikation der IT-Service-Managementprozesse Verketteter Dienste* 2009;

[21] Khurram, K. (Editor) *Optical Networking Stan-dards: A Comprehensive Guide for Professionals* (1st edn). 2006;

[22] Brockners, F. Ethernet OAM Overview: Making Ethernet Manageable. In *Proceedings 1. DFN Forum Kommunikationstechnologien,* Müller P, Neumair B, Dreo Rodesek G (eds). Kaiserslautern, 2008; 35-44.

[23] Ziegelmann, M. *Constrained shortest paths and related problems* 2004;

[24] Hamm, M. K., Yampolskiy, M. IT Service Management verketteter Dienste in Multi-Domain Umgebungen. Modellierung und Teilaspekte. *PIK Praxis der Informationsverarbeitung und Kommunikation* 2008; **31**(2):82-89.

[25] Bellman, R. The theory of dynamic programming. In *Proceedings of the National Academy of Sciences of the United States of America,* 1952; 716-719.

[26] Jaffe, J.M. Algorithms for finding paths with multiple constraints. *Networks* 1984; **14**(1):95-116.

[27] Kuipers, F.A. *Quality of service routing in the internet: Theory, complexity and algorithms* 2004;

[28] Feng, G. A multi-constrained multicast QoS routing algorithm. *Computer Communications* 2006; **29**(10):1811-1822.

[29] Castineyra, I. and Chiappa, N. and Steenstrup, M. The Nimrod routing architecture. RFC-1992 [1996].

[30] ITU-T Generic functional architecture of transport networks. ITU-T: G.805 [2000].

[31] TeleManagement Forum (TMF) Common Information Model (CIM). [2009].

[32] Hamm, M. K., Yampolskiy, M., Hanemann, A. I-SHARe. *DFN Mitteilungen* 2008; **74**:39-41.

[33] De Marinis, E. and Hamm, M. and Hanemann, A. and Vuagnin, G. and Yampolskiy, M. and Cesaroni, G. and Thomas, S.-M. *Deliverable DS3.16.1: Use Cases and Requirements Analysis for I-SHARe* Géant2 technical paper 2008.

[34] Steinmetz, R. *Multimedia-Technologie. Grundlagen, Komponenten und Systeme* 2000;

[35] M. Yampolskiy. *Maßnahmen zur Sicherung von E2E–QoS bei Verketteten Diensten*. Ph.D. Thesis, Ludwig–Maximilians–Universität München, December 2009.

[36] Géant SA2-T1 working group *Static Dedicated Wavelength Service: Joint Infrastructure and Operational Level Agreement* Géant2 technical paper, Draft 2010.