Multi-Domain End-to-End (E2E) Routing with multiple QoS Parameters

Considering Real World User Requirements and Service Provider Constraints

Mark Yampolskiy[1,2,3], Wolfgang Hommel[2,3], Bernhard Lichtinger[2], Wolfgang Fritz[1,2], Matthias K. Hamm[3]

[1]*German Research Network (DFN), Alexanderplatz 1, 10170 Berlin, Germany*

[2]*Leibniz Supercomputing Centre (LRZ), Boltzmannstraße 1, 85748 Garching, Germany*

[3] *Munich Network Management (MNM) Team, Oettingenstraße 67, 80538 München, Germany*

myy@dfn.de, {hommel,lichtinger,fritz}@lrz.de, hamm@mnm-team.org

***Abstract*—The growing amount of international collaborations in research, education, and business fields has raised once again the demand for quality assurance of network connections, which the projects and applications are realized upon. A large spectrum of examples with diverse requirements is found in areas such as *GRID* and *cloud computing*, eLearning, video on demand, and video-conferencing. The consequences of these diverse project and application requirements culminate in the urgent necessity to provide an End-to-End (E2E) guarantee for any user-tailored combination of service-specific Quality of Service (QoS) parameters. The quality of the overall network connections provided to users directly depends on the quality of the involved connection parts. This means that already during the routing process the quality of available connection parts has to be considered. Especially for international connections it is common that multiple service providers (SPs) realize different connection segments. At the same time the inter-domain routing is driven mostly by the combination of business interests and restrictive information policies. This means that during the routing not only the optimality of the available connection parts has to be considered, but also the preferences and restrictions of the involved provider domains. In this paper, we present an inter-domain routing algorithm for distinguishing the E2E route for dedicated point-to-point connections. The proposed algorithm considers both the E2E user requirements for service quality and the service provider constraints. The proposed algorithm is not restricted to consider a sole quality parameter and can therefore be used for the establishment of connections with the user-tailored combination of connection properties, including service quality as well as connection management functionality.**

***Keywords*-routing, quality of service, multi-constrained path finding**

I. Introduction

As the answer to the growing demand for network connections with guaranteed End-to-End (E2E) Quality of Service (QoS), multiple national and international research projects have been established. Currently, the Dynamic Circuit Network (DCN) cooperation led by Internet2 can be named as the most advanced of all of these projects [1]. Among others, projects like OSCARS, DRAGON, Phosphorus and the Géant-developed AutoBAHN are involved in this cooperation [2][3][4][5]. All these projects are focused on two aspects: techniques of dynamic circuit switching within a single administrative domain and interoperability between developed management systems as well as between the network technologies used in these domains in order to automatically switch multi-domain connections.

Despite all achievements of these projects, their crucial drawback is the consideration of only a sole QoS parameter — the bandwidth of the E2E connection. Support of further QoS parameters like jitter is planned for the future but not implemented yet. This restriction allows the mentioned projects to reuse the well established techniques and algorithms like OSPF, which is based on Dijkstra's algorithm [6]. However, an international research cooperation or any other large scale project will require E2E guarantees for more than just a single QoS parameter. Moreover, the combination of the required QoS parameters can vary regarding the application area of the network connections. For instance, in order to distribute raw experimental data of the Large Hadron Collider (LHC) project, additionally to the bandwidth also the high availability of the connections has to be guaranteed [7]. In the GRID cooperation DEISA, more than a dozen European supercomputing centers must be interconnected with guaranteed bandwidth and low jitter [8]. Currently, no automated establishment of connections considering combinations of multiple QoS parameters is supported by the existing connection services and they are neither tackled by the currently ongoing research projects. Therefore, the establishment of connections with multiple user-tailored QoS parameters remains a subject of manual connection planning and setup. For the LHC and DEISA projects mentioned above, such connections and their respective required properties are realized based on manually planned *Géant E2E Links* (also referred to as *GÉANT Lambda*) [9][10]. Due to the massively increasing demand for such high-quality links, a higher degree of automation is critical to ensure sufficient scalability and efficiency.

The core of the successful and efficient provisioning of multi-domain network connections is the routing algorithm: Based on the information about connections available within different administrative domains, the routing algorithm determines the connection parts of the new E2E connection. The algorithm should be able to cope with the multi-domain specifics, especially with preferences of the service providers

on the available connection parts.
The remainder of this paper is structured as follows: In the next section we discuss related work. Section III outlines the requirements and SP constraints considered in our proposal. Sections IV and V outline the information and communication models on which our routing algorithm is based on. The details of this routing algorithm are presented in Sections VI and VII: Section VI contains a detailed explanation of our novel inter-domain routing algorithm. Section VII focuses on the ordering aspects of the chosen route segments. A discussion about real-world application areas is given in Section VIII. An evaluation of the proposed solution is given in Section IX. A short outline of our future research plans concludes this paper.

II. Related Work

The actual inter-domain routing protocol in the internet is the Border Gateway Protocol (BGP), which does not take any QoS constraints into account [11]. Instead, BGP relies on provider-defined policies to influence the routing decisions. There are proposals to extend BGP to support QoS routing, like [12], which adds support for the exchange of information about supported classes of services, but this is only a coarse approach to add QoS, as no individual QoS parameters can be specified.
The DRAGON project, a typical example of DCN projects, uses a modified version of OSPF-TE to share a link state database between domains, which uses abstracted views of each domain to guarantee confidentiality. The path computation follows the same principle as the IETF path computation element (PCE) [13]; DRAGON calls the equivalent component *resource computation element* (RCE) [14]. The RCEs calculate the strict paths for their domain and communicate then with the RCE in the next domain. DRAGON adds time schedule and AAA policy constraints to the traffic engineering. But as the traffic engineering constraints are limited to reserve bandwidth, there are no guarantees regarding other QoS parameters like delay or jitter.
A variety of path computation problems in graphs with multiple weights has been investigated in [15]. In addition, [16] gives a broad overview of constraint-based path selection algorithms for QoS routing regarding one domain, for which a global knowledge about the QoS capabilities of the network is given. The *multi-constrained path* (MCP) problem is well understood and there are algorithms to solve the problem. E.g., SAMCRA [17] gives an exact solution for the MCP problem, whereas H_MCOP [18] provides a heuristic algorithm to find a multi-constrained path while simultaneously minimizing a path length function. All these algorithms have in common that they solve the MCP problem only within one domain and only for additive and multiplicative constraints like delay and availability.
[19] defines an inter-domain QoS routing algorithm, which is based on pre-computed transit paths through each domain and the possibility to calculate on-demand paths, if a demand cannot be fulfilled by pre-computed paths. It operates on a given domain chain for a path request, building a tree of path segments from the destination domain to the source domain of the request. This tree contains only abstract information about the path (entry and exit node with corresponding QoS parameters in each domain), to maintain the confidentiality of each domain. If the tree reaches the source it can make a selection among the found paths, which fulfill the constraints. The calculation of transit paths inside a domain is based on algorithms as mentioned in [16], so only the same classes of constraints are considered. The main drawback of the existing algorithms in graph theory considering multiple QoS parameters is the requirement of complete network topology information. This is, however, barely applicable to inter-domain routing. Furthermore, non-additive/non-multiplicative constraints like time slots for maintenance are not considered at all.

III. Focus, motivation and considered aspects

A pure packet switching approach, as it is used in the internet, has proven to be a very good concept for the optimization of resource utilization in practice. At the same time, it has been shown that the guarantee of QoS parameters for connections in packet switched networks is a very difficult task. The reason lies in the concurrence between communication flows for the available resources, which consequently lead to interferences between flows. On the contrary, a line switching in combination with resource reservation has proven to be a viable solution for QoS guarantees in PSTN networks.
As line switching can be emulated in a packet switched network, e.g., with combination of MPLS and RSVP techniques, we focus in our work on the routing for line-switched connections. Consequently, we adopt the definition of routing as a designation of all connection segments along the E2E path. We consider only routing on the same network layer. For the sake of simplicity, we assume that routing is performed at the network layer of services provided to the customer.
Furthermore, our work focuses on routing for Concatenated Services (CS), which are — regarding their planning and operation — probably the most challenging type of point-to-point connections. The following properties are characteristic for CS [20]:

- **User perspective:** E2E guarantees for the combination of relevant QoS parameters and/or management functionality are required for each connection.
- **Service composition:** The E2E service is composed of multiple horizontally (i.e., at the same network layer) concatenated connection segments, which are realized by different SPs.
- **Organizational relationships:** All SPs involved in this service's provisioning are independent organizations

and act as equal partners.

A detailed requirements analysis for CS can be found in [20]. In this section we only present a short outline of the most important requirements, challenges, and design criteria, which have played a decisive role during the development of the presented routing algorithm.

The necessity for a novel inter-domain routing approach is caused primarily by **user demand**. Instances of CS can be used in various types of user specific applications, which will require varying connection characteristics. This means that already during the routing process the user-tailored combinations of QoS parameters have to be considered. However, established routing algorithms like Dijkstra-based ones cannot be used for path finding with multiple criteria, because they are based on Bellman's optimality principle [21]. The fulfillment of this criterion is however not guaranteed if more than one parameter has to be considered [22]. Therefore, an inter-domain routing algorithm is needed, which can cope with multiple connection properties at the same time.

The **service composition** influences how the overall quality of an E2E connection is derived from the quality of each of the involved service parts. In this regard, especially the often-neglected fact has to be recalled that the overall E2E quality (QoS$_{E2E}$) is not solely influenced by the quality of connection parts provided within a single administrative SP domain, but also by the quality of — typically very short, but highly relevant — connections between those domains (see Figure 1). For the routing algorithm this means in the first place that not only the connectivity between two endpoints but also all relevant quality parameters of all involved parts have to be considered.

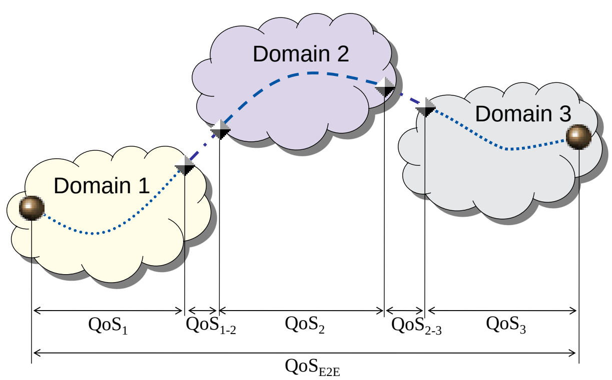

Figure 1. Composition of E2E connection quality

Finally, the **administrative boundaries** of SP domains cause very restrictive information and management policies. By the restrictive information policies especially the definition of information model is influenced. Indirectly, it also has an impact on the routing algorithm operating on the basis of information model. Significantly bigger impact is caused by the autonomy and independence of the involved SPs. Autonomy means in the first place the freedom of decision making. Furthermore, SPs often have preferences for or even constraints about acceptable routes. The reasons are manifold and vary from pure technical load balancing inside the own network to the contractual agreements with other SPs. Characteristic for the only two actually established inter-domain routing systems — BGP and SS7 — is the consideration of the individual SP interests. Multiple alternative proposals, like [23], neglected this aspect and have thus not evolved beyond the state of a research project. Consequently, in order to gain an SP's acceptance, the routing algorithm has to consider its preferences for realizable service parts.

Furthermore, a multi-domain environment raises several challenges related to inter-domain relationships, e.g., regarding trust and contractual relationships between SPs. Between all neighboring SPs and within a small tight provider cooperation the knowledge of each other as well as good trust relationships can be presumed. However, especially in a cooperation that is either big open and/or characterized by the high dynamic of SPs' participation, the trust relationships might become an issue. A good example for a service instance in a big cooperation is a telephone connection between two villages located near Sevilla and Beijing. In this case, multiple national and international scale PSTN providers can be involved in the provisioning of all needed connection parts; especially the national scale providers in Spain and China do not necessarily have any knowledge of each other and/or established trust relationship. Therefore, these SPs might be unwilling to participate in direct negotiations. Consequently, the routing algorithm should be able to cope with such situations.

IV. Key aspects of the used information model

In this section, we outline the information model defined in [24], which our routing algorithm is based on. This model allows the derivation of a multi-domain view with realizable connections from the local single-domain views of involved SPs. The speciality of our model is the possibility to associate multiple connection properties with every realizable connection part, i.e., QoS parameters and supported management functionality.

In order to illustrate the aspects of the information model that are especially relevant for the explanation of the proposed routing algorithm, we use an example with five SP domains and the connection parts they can realize (see Figure 2). The picture represents the knowledge of service provider SP$_1$ before any information exchange has been performed. For the routing algorithm, the following aspects are important:

- We distinguish between already known and not yet known but realizable connection parts. In fig. 2 they are depicted correspondingly as thick dash-pointed and thin dashed lines. Also, known realizable connections

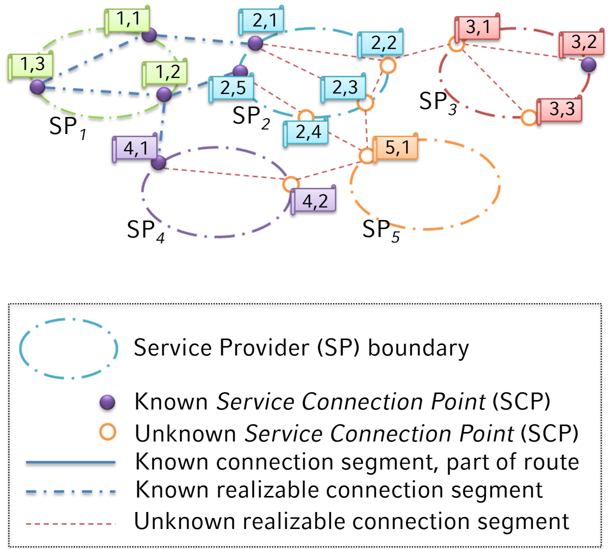

Figure 2. Representation of an example with five SPs

can be considered as parts of the possible route. In this case, such parts are depicted as bold continuous lines.

- From an SP-perspective, every connection part is a service, which can be realized upon the used network infrastructure. Therefore, we will use the term *service part* as a synonym for a connection part realized within a single SP or interconnecting two neighboring SPs. In order to reflect the fact whether the connection is realized within a single SP or interconnecting two SPs, correspondingly the terms *Domain Link* and *InterDomain Link* are used. Due to the fact that all connection parts are services from an SP's perspective, their endpoints are referred to as *Service Connection Points* (SCPs). Each SCP has a globally unique ID assigned to it. At the abstraction level of the network infrastructure, each SCP corresponds to one or more UNI/NNI[1] interfaces. Similar to the connection parts, we distinguish between already known and not yet known but existing SCPs.
- An SP's organizational boundary is depicted as an ellipse in Figure 2. In our work, we consider administrative SP boundaries because they define the scope of available information as well as responsibility areas, in which SPs can guarantee service quality. We assume that each SP has exact and up-to-date knowledge about the connection parts within its own domain as well as the connection parts interconnecting it with neighboring SP domains. Every SP has exclusive responsibility for the connection parts provided within its boundaries. The responsibility of the interconnecting parts is normally shared between the neighboring SPs. In order to simplify and standardize management processes, we propose that every SP acts as a proxy by offering the quality assurance for the whole InterDomain Link. The contractual details as well as the aspects of partial responsibility should (and can only) be negotiated by each pair of neighboring SPs.

[1] User-Network Interface (UNI) and Network-Network Interface (NNI) are used to refer interfaces interconnecting SP with users or with neighboring SPs correspondingly.

V. Inter-Domain communication, outline

The purpose of the routing algorithm described in our work is the selection of the path at the abstraction level of SP services. Consequently, also the communication between SPs has to be defined at this abstraction level. In contrast to network management, which sees the UNI/NNI interfaces as the combination of connection points and inter-domain communication interfaces, we treat these concepts separately.

In IT Service Management, the well established MNM Service Model [25] defines two communication interfaces: *Service Access Points* (SAP) for service delivery to the user and *Customer Service Management* (CSM) for delivering management functionality to the customer. We propose to extend this model with the *Domain Service Management* (DSM) interface, which should be used for communication between collaborating SPs.

Furthermore, a communication protocol is needed in order to request functionality through the DSM interface. We argue that an exact technical specification of the DSM interface as well as of the communication protocol cannot be given at this point, as it might vary with respect to multiple aspects like the decision for *in-band* or *out-band* signaling or the management functionality available through this interface. Therefore, we only discuss the following aspects needed during the routing process:

As the routing algorithm can operate only on the data about available service parts, the protocol should allow for requesting such information from different SPs. For the multi-domain ordering process, additional requests for service parts ordering and — in order to minimize interferences between simultaneous ordering processes — reservation are needed.

VI. Routing

Our proposal for inter-domain routing consists of two parts. In the regular case, *source routing with semi-global knowledge* should be performed. This part of the algorithm is described in Section VI-A. In the case of missing or insufficient direct trust relationships between communicating SPs, an *on-demand delegation* has to be performed; this part of the algorithm is described in Section VI-B. In order to clarify the proposal, we will illustrate the algorithm on the basis of an example.

A. Source-Routing with semi-global knowledge

From the SP-perspective, the ordering process begins when the customer requests a new service instance through the CSM interface (see Figure 3). The customer has to specify two endpoints of the requested connection. We will refer to such endpoints as SCPs, because from an SP's perspective, they reside on the border of its own administrative domain. The specification of endpoints can be performed either explicitly or implicitly, like it is done for telephone connections. In order to enable user-tailored services, the customer requirements to the E2E service instance properties also have to be specified. In the depicted example, a customer requests a permanent 1Gbps connection between SCPs with IDs "1,3" and "3,2" (subsequently we will refer to these SCPs as $SC_{1,3}$ and $SCP_{3,2}$). The delay of the whole connection shall be below 40ms. Maintenance of the connection, which might lead to service interruption, is acceptable to the customer. However, maintenance can only take place between 6am and 10am GMT and should take no longer than 1 hour.

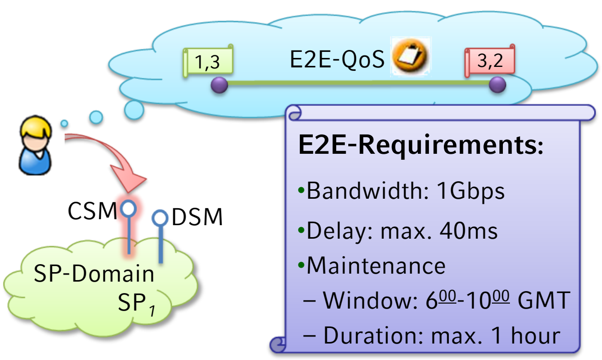

Figure 3. Customer requests new service instance

From the user and customer perspective, the connection SP (in the example SP_1) should take the role of the service provider for the whole E2E service instance. In this case, the connection SP hides the internal service composition and details about further participating SPs or about interactions between SPs from the customer.

We derive the strategy of the routing algorithm directly from depth first search. The **order** in which different alternative service parts have to be considered for the route corresponds to the preferences of SPs providing these service parts. The main reason for this decision is to meet the SP demand for being able to influence the route finding process (see also corresponding discussion in Section III).

For the sake of simplicity, we assume that each SP can perform the inter-domain routing by itself. Therefore, the connection SP — based on the information about its own network, its own preferences, and customer requirements — can decide, which available connection part has to be tried first as part of the route finding. In the example, SP_1 selects the Domain Link between $SCP_{1,3}$ and $SCP_{1,1}$ and the InterDomain Link between $SCP_{1,1}$ and $SCP_{2,1}$ (see Figure 4).

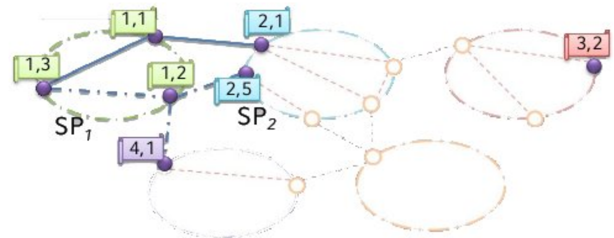

Figure 4. Considered route (1)

As the knowledge of SP_1 is restricted to the own Domain Links and to InterDomain Links in whose provisioning it is involved, it has to request information about further available service parts from the next SP en route. As the last SCP en route ($SCP_{2,1}$) belongs to SP_2, the missing information has to be requested from this provider (see Figure 5).

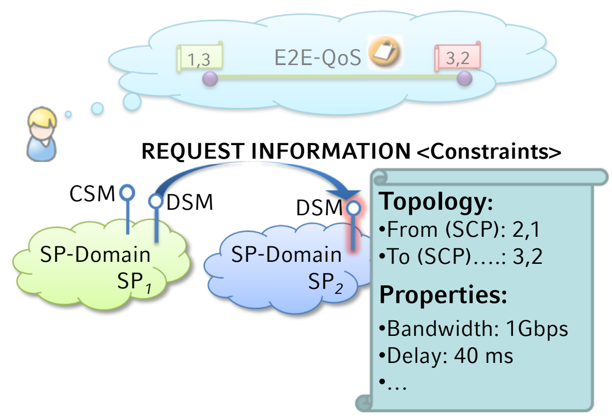

Figure 5. Requesting missing information

As discussed in Section III, one of the most critical aspects in the multi-domain environment corresponds to the very restrictive SP information policies. In order to mitigate the SP concerns, we propose to **restrict the requested information** based on two aspects:

- The topology parameters, which the reported service parts have to meet. We propose to always use the SCP on the end of the considered route and the remote SCP specified by the customer as endpoint. All service parts reported by SP should be adjacent to the SCP at the end of the considered route.
- The relevant connection properties (QoS parameters and management functionality), which the reported service parts have to match. We propose to reuse the E2E requirements specified by the customer. The reason

is that the information about service parts, even if they might be inacceptable for the particular route, might be reused by considering an alternative route.

In addition to the improvement of SP acceptance, such reduction of information has several additional advantages. First of all, background information of particular SPs about which service parts within this SP lead into the direction of the desired endpoint can be reused. Consequently, it should reduce the amount of considered alternatives, which in turn reduces the runtime complexity of the routing algorithm. Also, there is no need for the global knowledge of all service parts realizable by all SPs. This in turn reduces the amount of communication between SPs needed to collect all necessary information.

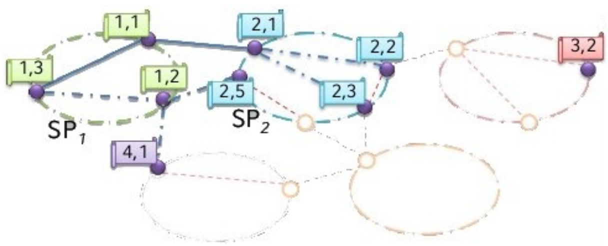

Figure 6. Considered route (2)

In the example, SP_2 can suggest two alternative service parts leading to $SCP_{3,2}$: one going via $SCP_{2,2}$ and another one via $SCP_{2,3}$ (see Figure 6). From these two alternatives, the one going via $SCP_{2,3}$ is preferable for SP_2. Therefore, this service part should be considered first as the next part of the possible route. In order to give a deeper insight into the proposed algorithm, Table I contains several imaginary values for the properties of the two new service parts (last two columns) as well as of the considered path (second column). The values are prepared the way that all parameters of each of the service parts comply with E2E requirements. However, the value aggregation of intermediate sum (2nd column) and the most preferable service part (3rd column) will break the E2E requirements for the delay QoS parameter. Therefore, the second best alternative (4th column) has to be considered as the next one.

In the depicted example, the second alternative can be used as a part of the route. In this case SP_1 has to contact SP_2 again with information request. This time the topology constraint has to be changed to "from $SCP_{2,2}$ to $SCP_{3,2}$". In the best case scenario, the re-iteration of the proposed algorithm should lead to a path between two endpoints complying with all E2E requirements.

B. *On-demand routing by delegation*

As discussed above, sufficient trust relationships between the communicating SPs cannot always be preconditioned.

	E2E-Requirements	Intermediate Σ	Service-Part 1	Service-Part 2
From (SCP)	$SCP_{1,3}$	$SCP_{1,3}$	$SCP_{2,1}$	$SCP_{2,1}$
To (SCP)	$SCP_{3,2}$	$SCP_{2,1}$	$SCP_{2,3}$	$SCP_{2,2}$
Bandwidth	1 Gbps	1 Gbps	1 Gbps	1 Gbps
Delay	40 ms	23 ms	**20 ms**	10 ms
Maintenance Window	6^{00}-10^{00} GMT	6^{00}-10^{00} GMT	6^{00}-8^{00} GMT	7^{00}-9^{00} GMT
Maintenance Duration	1 hr	1 hr	40 min	1 hr

Table I
INTERMEDIATE VALUES FOR CONSIDERED ROUTE

Especially in a large open and/or highly dynamic cooperation it should not necessarily be the case. Therefore, an emergency solution is needed in case that trust relationships are insufficient.

In order to illustrate this situation, we assume that a route between $SCP_{1,3}$ and $SCP_{3,1}$ could be found with the help of the algorithm described in Section VI-A (see Figure 7). According to the algorithm defined in Section VI-A, SP_1 has to contact SP_3 with an information request about realizable service parts leading to the endpoint $SCP_{3,2}$. If SP_3 has very restricted information policies and accepts such request, e.g., allowing them only from neighboring domains, the request from SP_1 will be rejected (see Figure 8).

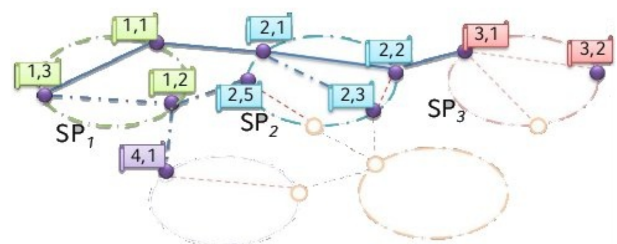

Figure 7. Considered route (3)

We see two possible alternative procedures to cope with this issue. The first one is to try to find an alternative route around the SP refusing to provide the requested information. We argue *against* using this approach, as it has at least the following drawbacks: The "way around" can only be possible by transit SPs. However, if the SP refuses to respond with the required information, this method will not work. Second, without partial-services available in the SP with restrictive policies, the alternative route might not be optimal (from the SP perspective) or even not possible; and third, re-routing will require additional communication with other SPs.

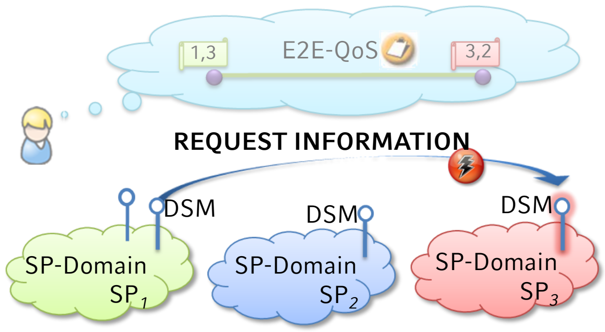

Figure 8. Requesting information rejected by SP_3

Instead we propose to use *Routing-by-Delegation* as an *on-demand* workaround in such situations. In this case, the routing task has to be delegated to the last SP in the found route, which is directly connected to the SP rejecting information requests. In the example this is SP_2 (see Figure 9). The SP to which the routing task has been delegated is responsible for the selection of the remaining route. The following information has to be passed to this SP:

- SCPs between which the routing has to be performed. These are the last SCP in the already chosen part of the route and the distant endpoint.
- The E2E requirements specified by the customer. These are properties and values, which have to be fulfilled by the complete route between two endpoints.
- The intermediate sum of relevant values along the already found path.

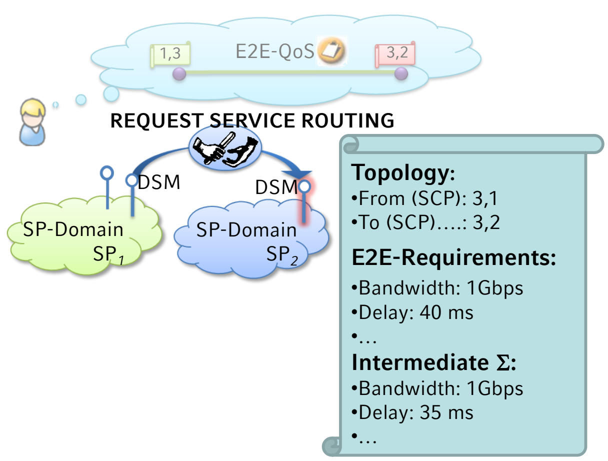

Figure 9. Delegation of routing task to SP_2

We propose that the SP to which the routing task has been delegated performs this as specified in Section VI-A. Further delegation of routing subtasks remains possible, but should be used as an on-demand exceptional solution.
Once the SP that is responsible for choosing the remaining route (in the example SP_2) has finished its task, the result has to be reported to the SP that has requested routing for the remaining part. If no acceptable route has been found via delegation of the routing task to the next SP in the route, an alternative branch should be examined as described above.
In case the routing was successful, only the properties of the found route are relevant. They should be passed back, as they can be used in the customer-provider negotiation process. The exact details about service parts involved in the remaining route as well as about the SPs providing these parts can be omitted. At the same time, the SP which has performed the routing should now take a proxy role for all requests related to the remaining part, map them to the chosen service parts and pass these requests to the corresponding SPs (see also more elaborate discussion in Section VII). The necessity for this procedure lies in the trust relationships: If information requests have been rejected, it cannot be assumed that reservation and/or ordering requests will be accepted.

C. Drawbacks of on-demand routing

The main drawback of the proxy role is the necessity to manage mapping information as well as the necessity to recognize when mapping information is not needed anymore and therefore the resources allocated for information management can be freed. Different approaches are conceivable as a solution for this situation, e.g., not to save the mapping information in the proxy SP at all and just provide the result of the route choice together with the mapping to the requesting SP. Another possibility is to save the mapping information at the proxy SP and to notify the requesting SP about the lifetime of this information. Further technical solutions are also imaginable. As all these solutions have different advantages and disadvantages, no general recommendation can be given at this point. The final decision has to be made based on the specific requirements of the user-faced service as well as of aspects of provider collaboration.
Another major drawback of *routing-by-delegation* compared to *source-routing* is the comparatively bigger difficulty of loop prevention. For *source-routing* it is enough to mark SCPs as "used in path" as it is usually done in graph theory. In order to prevent loops by *routing-by-delegation*, we propose to associate a globally unique ID with every instance ordered by a customer. For information requests, this ID should always be passed together with the constraints on service parts. If an SP receives a second information request for the same service instance and between the same SCPs, it should reject the request in case all constraint parameters are as previously given. Please note that multiple requests for the same service instance should be accepted, if different SCPs have been specified. In order to recognize loops, a corresponding information management is needed. Similar to the discussion above, different solutions can be elaborated. Alternatives are time life management implemented

by all SPs or the passing of detailed information about previous route to the proxy SP. Time management keeps the network communication very low, requires, however, true management on the SP side. Information passing has complementary properties. As discussed above, no general recommendation for one solution can be given without considering concrete requirements of the offered user-faced service and/or SP cooperation. The objective of the presented discussion is to stress the necessity to solve the outlined aspects.

VII. Reservation and Ordering of service parts

It is obvious that routing considering the E2E requirements is a prerequisite for the provisioning of user-tailored connection services. However, in order to guarantee the properties of a chosen route, all service parts along the found route have to be ordered for the particular service instance. This means in turn that the providers of these service parts have to reserve and to assign the resources necessary to hold parameters specified during the information request to the service instance.

As multiple simultaneous instances can be ordered in a multi-domain environment, the reservation of a service (or service parts) before ordering it has proven to be a very good solution to reduce negative influence between concurrent ordering processes. In addition, several patterns have been elaborated for the negotiation of service quality. We propose to reuse the ideas of *bilateral* and *trilateral peer-to-peer negotiation patterns* as described in [26]. Corresponding to these patterns, if the reservation of a service is requested, the SP providing this service may reduce the requested quality. If the already reserved service is ordered, the SP providing this service should reserve resources to guarantee at least the requested quality. If this is not possible, the ordering request should fail. This behavior also means that information about actually reserved and/or ordered quality has to be passed back to the requester.

The described behavior should be reused during the reservation and ordering of all service parts. Furthermore, we define a communication during the reservation and ordering of all service parts along the found route as follows:

- Reservation of service parts should be performed sequentially along the found route. Reservation should begin at the service part connected to the starting endpoint. All further service parts should be adjacent to the previously reserved one. This reservation of a service part should be only requested if the previous one was successful. The main reason for such a procedure is that it allows an adaptation to the following two problem cases: (a) necessary resources are not available anymore and the SP has to confirm reservation with only reduced service quality, and (b) the SP is unable to realize the requested service part at all and notifies the requester about this failure. In the first case it might be possible — e.g., if the delay for the reserved service part is bigger than expected — to adjust requirements to the remaining service parts along the found route. Furthermore, in both cases a re-routing might be used in order to compensate for the reduced quality of already ordered route parts. In the best case the re-routing of the remaining route might lead to the desired route quality. In the worst case the reservation of the service parts shall be cancelled in reverse order, each time followed by a re-routing attempt. If re-routing has succeeded, the reservation procedure should be performed as described above.
- We assume that the probability that the reserved service parts cannot be ordered with the required quality is very low. Therefore, we propose to send the ordering requests simultaneously to all involved SPs.

Consequently, to order the service instance fulfilling E2E user requirements the following requests should be supported by the communication protocol:

- information request: it is needed for route finding
- reservation of service part, and
- cancellation of reservation: they are needed during E2E reservation
- orderings of reserved service part: it is needed during ordering of all reserved service parts
- finally the decommissioning of service parts: it is needed if one or more ordering requests have failed.

Decommissioning of service parts is also needed at the end of the service instance life cycle, when the whole E2E service instance has to be decommissioned. During the requests for reservation, cancellation of reservation, ordering and decommissioning, the relevant partial-service has to be identified unambiguously. As one concatenated service can contain only a single service part between the same SCPs, we propose to identify the service parts via the combination of the service instance ID and the IDs of two involved SCPs. Alternative identification possibilities could be applied, like the association of IDs to the particular service parts.

VIII. Application Areas and special cases

Compared to the established connectivity-oriented routing approaches, the proposed algorithm requires significantly more input, computations, and — what is most time-consuming — 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. Instead we see the application area of the developed algorithm 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. The proposed algorithm can provide near-real-time finding and ordering of a route with customer-tailored E2E connection properties. This algorithm is applicable to scenarios like video-on-demand, videoconferencing, on demand connectivity for GRID or Cloud collaborations and so on, where customers are willing to pay not only for the connectivity but also for the connection quality.

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* (see Section VI-B) 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 case of 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 instance requests, which neglect 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* (see Section VI-A) in which the routing task is 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.

Concluding, the proposed algorithm is applicable in the most challenging case of an open SP collaboration and with a variety of customer-specific QoS parameters.

IX. Evaluation of SP-Acceptance

As discussed in Section III, most critical 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 experience within the context of the ***I**nformation **S**haring across **H**eterogeneous **A**dministrative **Re**gions* (I-SHARe) activity. This activity has been established in Géant in order to foster information exchange during manual planning and operation of E2E Links. During the first phase of this activity, the established manual processes have been captured and described in [27]. The handling of the domain-internal information as well as the influence of SPs on the route planning show several important similarities with the algorithm described in Section VI-A. In both cases the information request about possible service parts is service instance related and also provides information about the connection point with the previous SP. The response about available service parts is also defined at abstraction level of SP-domains. In contrast to the proposal described in this paper, only one possible SP service part is considered at the same time in Géant processes. This is especially needed to simplify work during the manual planning. In the case that the service part cannot be used, a new information request is foreseen in the Géant procedure. 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 operations involved into planning of Géant E2E Links have shown the acceptance of SPs to provide information about multiple service parts during the routing process. The only true concern that has been repeatedly mentioned is the compliance of the service part choice to the SP preferences. The enforcement of this aspect has been reflected in both algorithm parts described in Section VI-A and VI-B.

X. Conclusion

As already mentioned in Section II, existing routing procedures are restricted to QoS parameters with additive and sometimes also multiplicative aggregation functions. In order to support a variety of customer-specific QoS parameters, a generalized solution for the QoS function is urgently needed.

In the presented work, we have focused on the E2E routing during the ordering of a new instance of a CS. The definition of management processes, e.g., E2E monitoring, needed during the operation of a CS-instance is the next logical step in our research. Furthermore, we plan to investigate the applicability of self-adaptation techniques as a means for multi-domain compensation of single-domain quality reduction.

Acknowledgment

The authors wish to thank the members of the Munich Network Management Team (MNM Team) for fruitful 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] "Dynamic Circuit Network (DCN) Homepage," [Online: http://www.internet2.edu/network/dc/].

[2] "On-Demand Secure Circuits and Advance Reservation System (OSCARS) Homepage," [Online: https://oscars.es.net/OSCARS/docs].

[3] "Dynamic Resource Allocation via GMPLS optical Networks (DRAGON) Homepage," [Online: http://dragon.maxgigapop.net/].

[4] "Phosphorus: Homepage," [Online: http://www.ist-phosphorus.eu/].

[5] "Automated Bandwidth Allocation across Heterogeneous Networks (AutoBAHN)," [Online: http://www.geant2.net/server/show/nav.756].

[6] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[7] "LHC – The Large Hadron Collider, Homepage," [Online: http://lhc.web.cern.ch/lhc/].

[8] "DEISA – Distributed European Infrastructure for Supercomputing Applications, Homepage," [Online: http://www.deisa.eu/].

[9] E. Apted, J. Chevers, M. Garcia Vidondo, and S. Tyley, "Deliverable DS2.0.3,3: Report on GÉANT2 Advanced Services – Lambdas and Switched Optical," Géant2, Tech. Rep., 2009.

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

[11] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," http://tools.ietf.org/html/rfc4271, Jan 2006.

[12] T. Knoll, "BGP Extended Community Attribute for QoS Marking, draft-knoll-idr-qos-attribute-05, work in progress," IETF, 2010.

[13] A. Farrel, J. Vasseur, and J. Ash, "RFC 4655: A Path Computation Element (PCE)-Based Architecture," *IETF, August*, 2006.

[14] "Network Aware Resource Broker (NARB) and Resource Computation Element (RCE) Architecture, Version 2.1b," http://dragon.east.isi.edu, April 2008.

[15] M. Ziegelmann, *Constrained Shortest Paths and Related Problems*. VDM, 2007.

[16] F. Kuipers, *Quality of service routing in the internet: Theory, complexity and algorithms*. Delft University Press, 2004.

[17] P. Van Mieghem, H. De Neve, and F. Kuipers, "Hop-by-hop quality of service routing," *Computer Networks*, vol. 37, no. 3-4, pp. 407–423, 2001.

[18] T. Korkmaz and M. Krunz, "Multi-constrained optimal path selection," in *IEEE INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 2, 2001.

[19] A. Frikha and S. Lahoud, "Hybrid Inter-Domain QoS Routing based on Look-Ahead Information," 2010, Publications Internes de I'IRISA.

[20] M. Hamm and M. Yampolskiy, "IT Service Management verketteter Dienste in Multi-Domain Umgebungen. Modellierung und Teilaspekte," *PIK-Praxis der Informationsverarbeitung und Kommunikation*, vol. 31, no. 2, pp. 82–89, 2008.

[21] R. Bellman, "The theory of dynamic programming," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 8, pp. 716–719, 1952.

[22] J. Jaffe, "Algorithms for finding paths with multiple constraints," *Networks*, vol. 14, no. 1, pp. 95–116, 1984.

[23] I. Castineyra, N. Chiappa, and M. Steenstrup, "The Nimrod routing architecture," 1996.

[24] M. Yampolskiy, W. Hommel, P. Marcu, and M. Hamm, "An information model for the provisioning of network connections enabling customer-specific End-to-End QoS guarantees," *Submitted to The 7th International Conference on Services Computing, SCC 2010*, 2010.

[25] M. Garschhammer, R. Hauck, H. Hegering, B. Kempter, I. Radisic, H. Rölle, H. Schmidt, M. Langer, and M. Nerb, "Towards generic service management concepts a service model based approach," in *2001 IEEE/IFIP International Symposium on Integrated Network Management Proceedings*, 2001, pp. 719–732.

[26] R. Steinmetz, *Multimedia-Technologie: Grundlagen, Komponenten und Systeme*. Springer, 1998.

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