

Internet Draft

Francois Le Faucheur
Michael Dibiasio
Bruce Davie
Cisco Systems, Inc.

Michael Davenport
Chris Christou
Booz Allen Hamilton

Jerry Ash
Bur Goode
AT&T

[draft-ietf-tsvwg-rsvp-dste-00.txt](#)

Expires: January 2006

July 2005

Aggregation of RSVP Reservations over MPLS TE/DS-TE Tunnels

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

Abstract

This document provides specification for aggregation of RSVP end-to-end reservations over MPLS Traffic Engineering (TE) tunnels or MPLS Diffserv-aware MPLS Traffic Engineering (DS-TE) Tunnels. This approach is based on [RFC 3175](#) and simply modifies the corresponding procedures for operations over MPLS TE tunnels instead of aggregated RSVP reservations. This approach can be used to achieve admission control of a very large number of flows in a scalable manner since the devices in the core of the network are unaware of the end-to-end RSVP reservations and are only aware of the MPLS TE tunnels.

Copyright Notice

Copyright (C) The Internet Society. (2005)

Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [\[RFC2119\]](#).

1. Introduction

The Integrated Services (Intserv) [\[INT-SERV\]](#) architecture provides a means for the delivery of end-to-end Quality of Service (QoS) to applications over heterogeneous networks.

[RSVP] defines the Resource reSerVation Protocol which can be used by applications to request resources from the network. The network responds by explicitly admitting or rejecting these RSVP requests. Certain applications that have quantifiable resource requirements express these requirements using Intserv parameters as defined in the appropriate Intserv service specifications ([\[GUARANTEED\]](#), [\[CONTROLLED\]](#)).

The Differentiated Services (DiffServ) architecture ([\[DIFFSERV\]](#)) was then developed to support differentiated treatment of packets in very large scale environments. In contrast to the per-flow orientation of Intserv and RSVP, Diffserv networks classify packets into one of a small number of aggregated flows or "classes", based on the Diffserv codepoint (DSCP) in the packet's IP header. At each Diffserv router, packets are subjected to a "per-hop behavior" (PHB), which is invoked by the DSCP. The primary benefit of Diffserv is its scalability. Diffserv eliminates the need for per-flow state and per-flow processing and therefore scales well to large networks.

However, DiffServ does not include any mechanism for communication between applications and the network. Thus, as detailed in [\[INT-DIFF\]](#), significant benefits can be achieved by using Intserv over Diffserv

including resource based admission control, policy based admission control, assistance in traffic identification /classification and traffic conditioning. As discussed in [\[INT-DIFF\]](#), Intserv can operate over Diffserv in multiple ways. For example, the Diffserv region may be statically provisioned or may be RSVP aware. When it is RSVP aware, several mechanisms may be used to support dynamic provisioning and topology aware admission control including aggregated RSVP reservations, per flow RSVP or a bandwidth broker. The advantage of using aggregated RSVP reservations is that it offers dynamic, topology-aware admission control over the Diffserv region without the scalability burden of per-flow reservations and the associated level of RSVP signaling in the Diffserv core. [\[RSVP-AGGR\]](#) describes in detail how to perform such aggregation of end to end RSVP reservations over aggregated RSVP reservations in a Diffserv cloud. It establishes an architecture where multiple end-to-end RSVP reservations sharing the same ingress router (Aggregator) and the same egress router (Deaggregator) at the edges of an "aggregation region", can be mapped onto a single aggregate reservation within the aggregation region. This considerably reduces the amount of reservation state that needs to be maintained by routers within the aggregation region. Furthermore, traffic belonging to aggregate reservations is classified in the data path purely using Diffserv marking.

[\[MPLS-TE\]](#) describes how MPLS TE Tunnels can be established via [\[RSVP-TE\]](#) and how these tunnels can be used to carry arbitrary aggregates of traffic. MPLS TE uses Constraint Based Routing to compute the path for a TE tunnel. Then, CAC (Call Admission Control) is performed during the establishment of TE Tunnels to ensure they are granted their requested resources.

[\[DSTE-REQ\]](#) presents the Service Providers requirements for support of Diff-Serv-aware MPLS Traffic Engineering (DS-TE). With DS-TE, separate DS-TE tunnels can be used to carry different Diffserv classes of traffic and different resource constraints can be enforced for these different classes. [\[DSTE-PROTO\]](#) specifies RSVP-TE signaling extensions as well as OSPF and ISIS extensions for support of DS-TE.

In the rest of this document we will refer to both TE tunnels and DS-TE tunnels simply as "TE tunnels".

TE tunnels have much in common with the aggregate RSVP reservations used in [\[RSVP-AGGR\]](#):

- a TE tunnel is subject to CAC and thus is effectively an aggregate bandwidth reservation
- In the data plane, packet scheduling relies exclusively on Diff-Serv classification and PHBs
- Both TE tunnels and Aggregate RSVP reservations are controlled

by "intelligent" devices on the edge of the "aggregation core"

- (Head-end and Tail-end in the case of TE tunnels, Aggregator and Deaggregator in the case of Aggregated RSVP reservations
- Both TE tunnels and Aggregate RSVP reservations are signaled using the RSVP protocol (with some extensions defined in [RSVP-TE] and [[DSTE-PROTO](#)] respectively for TE tunnels and DS-TE tunnels).

This document provides a detailed specification for performing aggregation of end-to-end RSVP reservations over MPLS TE tunnels (which act as aggregated reservations in the core). This document builds on the RSVP Aggregation procedures defined in [[RSVP-AGGR](#)], and only changes those where necessary to operate over TE tunnels. With [[RSVP-AGGR](#)], a lot of responsibilities (such as mapping end-to-end reservations to Aggregate reservations and resizing the Aggregate reservations) are assigned to the Deaggregator (which is the equivalent of the Tunnel Tail-end) while with TE, the tunnels are controlled by the Tunnel Head-end. Hence, the main change over the RSVP Aggregations procedures defined in [[RSVP-AGGR](#)] is to modify these procedures to reassign responsibilities from the Deaggregator to the Aggregator (i.e. the tunnel Head-end).

[LSP-HIER] defines how to aggregate MPLS TE Label Switched Paths (LSPs) by creating a hierarchy of such LSPs. This involves nesting of end-to-end LSPs into an aggregate LSP in the core (by using the label stack construct). Since end-to-end TE LSPs are themselves signaled with RSVP-TE and reserve resources at every hop, this can be looked at as a form of aggregation of RSVP(-TE) reservations over MPLS TE Tunnels. This document capitalizes on the similarities between nesting of TE LSPs over TE tunnels and RSVP aggregation over TE tunnels and reuses the procedures of [[LSP-HIER](#)] wherever possible.

This document also builds on the "RSVP over Tunnels" concepts of [RFC 2746](#) [[RSVP-TUN](#)]. It differs from that specification in the following ways

- Whereas [RFC 2746](#) describes operation with IP tunnels, this draft describes operation over MPLS tunnels. One consequence of this difference is the need to deal with penultimate hop popping (PHP).
- MPLS-TE tunnels inherently reserve resources, whereas the tunnels in [RFC 2746](#) do not have resource reservations by default. This leads to some simplifications in the current draft.
- There is exactly one reservation per MPLS-TE tunnel, whereas [RFC 2746](#) permits many reservations per tunnel.
- We have assumed in the current draft that a given MPLS-TE tunnel will carry reserved traffic and nothing but reserved traffic, which negates the requirement of [RFC 2746](#) to distinguish reserved and non-reserved traffic traversing the

same tunnel by using distinct encapsulations.

- There may be several MPLS-TE tunnels that share common head and tail end routers, with head-end policy determining which tunnel is appropriate for a particular flow. This scenario does not appear to be addressed in [RFC 2746](#).

At the same time, this draft does have many similarities with [RFC 2746](#). MPLS-TE tunnels are "type 2 tunnels" in the nomenclature of [RFC 2746](#):

"

The (logical) link may be able to promise that some overall level of resources is available to carry traffic, but not to allocate resources specifically to individual data flows.

"

Aggregation of end-to-end RSVP reservations over TE tunnels combines the benefits of [[RSVP-AGGR](#)] with the benefits of MPLS including the following:

- dynamic, topology-aware resource-based admission control can be provided to applications over any segment of the end to end path including the core
- as per regular RSVP behavior, RSVP does not impose any burden on routers where such admission control is not needed (for example if the links upstream and downstream of the MPLS TE core are vastly over-engineered compared to the core capacity, admission control is not required on these links and RSVP need not be processed on the corresponding router hops)
- the core scalability is not affected (relative to the standard MPLS TE deployment model) since the core remains unaware of end-to-end RSVP reservations and only has to maintain aggregate TE tunnels and since the datapath classification and scheduling in the core relies purely on Diffserv mechanism (or more precisely MPLS Diffserv mechanisms as specified in [[DIFF-MPLS](#)])
- the aggregate reservation (and thus the traffic from the corresponding end to end reservations) can be network engineered via the use of Constraint based routing (e.g. affinity, optimization on different metrics) and when needed can take advantage of resources on other paths than the shortest path
- the aggregate reservations (and thus the traffic from the corresponding end to end reservations) can be protected against failure through the use of MPLS Fast Reroute

This document, like [[RSVP-AGGR](#)], covers aggregation of unicast sessions. Aggregation of multicast sessions is for further study.

[1.1.1](#). Changes from previous versions

The significant changes from version [draft-lefaucheur-rsvp-dste-02](#)

to version [draft-ietf-tsvwg-rsvp-dste-00.txt](#) of this draft are:

Le Faucheur, et al.

[Page 5]

- added a SHOULD for use of Make-Before-Break when resizing TE tunnel
- added clarification text about E2E Resv hiding from Transit LSRs
- added reference to [[RSVP-AGG-IPSEC](#)] in [section 5](#).
- added definition of E2E reservation in [section 2](#).
- removed the case where E2E reservation is a TE tunnel (already covered in [[LSP-HIER](#)])

The significant changes from version [draft-lefaucheur-rsvp-dste-01](#) to version [draft-lefaucheur-rsvp-dste-02](#) of this draft are:

- Alignment with RSVP operations of [draft-ietf-mpls-lsp-hierarchy](#)
- Addition of an appendix providing an example usage scenario for information purposes

The significant changes from version [draft-lefaucheur-rsvp-dste-00](#) to version [draft-lefaucheur-rsvp-dste-01](#) of this draft were:

- added discussion of the relationship to [RFC 2746](#) [[RSVP-TUN](#)]
- added discussion of mapping policy at aggregator
- added discussion of "RSVP proxy" behavior in conjunction with the aggregation scheme described here
- added discussion on TTL processing on Deaggregator

2. Definitions

For readability, a number of definitions from [[RSVP-AGGR](#)] as well as definitions for commonly used MPLS TE terms are provided here:

Aggregator This is the router at the ingress edge of the aggregation region (with respect to the end to end RSVP reservation) and behaving in accordance with [[RSVP-AGG](#)]. In this document, it is also the TE Tunnel Head-end.

Deaggregator This is the router at the egress edge of the aggregation region (with respect to the end to end RSVP reservation) and behaving in accordance with [[RSVP-AGG](#)]. In this document, it is also the TE Tunnel Tail-end

E2E End to end

E2E reservation This is an RSVP reservation such that:
(i) corresponding Path messages are initiated upstream of the Aggregator and terminated downstream of the Deaggregator, and

- (ii) corresponding Resv messages are initiated downstream of the Deaggregator and terminated upstream of the Aggregator, and
- (iii) this RSVP reservation is to be aggregated over an MPLS TE tunnel between the Aggregator and Deaggregator.

An E2E RSVP reservation may be a per-flow reservation. Alternatively, the E2E reservation may itself be an aggregate reservation of various types (e.g. Aggregate IP reservation, Aggregate IPsec reservation). See [section 4](#) and 5 for more details on the types of E2E RSVP reservations. As per regular RSVP operations, E2E RSVP reservations are unidirectional.

Head-end

This is the Label Switch Router responsible for establishing, maintaining and tearing-off a given TE tunnel.

Tail-end

This is the Label Switch Router responsible for terminating a given TE tunnel

Transit LSR This is a Label Switch router that is on the path of a given TE tunnel and is neither the Head-end nor the Tail-end

3. Operations of RSVP Aggregation over TE with pre-established Tunnels

[RSVP-AGG] supports operations both in the case where aggregate RSVP reservations are pre-established and in the case where Aggregating and De-aggregating routers have to dynamically discover each other and dynamically establish the necessary Aggregated RSVP reservations.

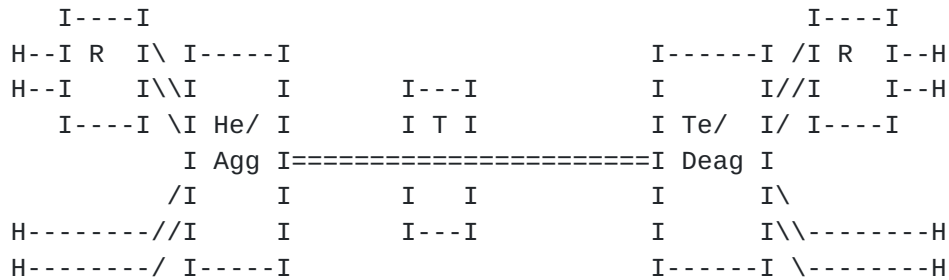
Similarly, RSVP Aggregation over TE tunnels could operate both in the case where the TE tunnels are pre-established and in the case where the tunnels need to be dynamically established.

In this document we provide a detailed description of the procedures in the case where TE tunnels are already established. These procedures are based on those defined in [[LSP-HIER](#)].

Pre-establishment of the TE tunnels may be triggered by any mechanisms including for example manual configuration or automatic establishment of a TE tunnel mesh through dynamic discovery of TE Mesh membership as allowed in [[AUTOMESH](#)].

Procedures in the case of dynamically established TE tunnels are for further studies.

3.1. Reference Model



H = Host requesting end-to-end RSVP reservations

R = RSVP router

He/Agg = TE tunnel Head-end/Aggregator

Te/Deag = TE tunnel Tail-end/Deaggregator

T = Transit LSR

-- = E2E RSVP reservation

== = TE Tunnel

3.2. Receipt of E2E Path message By the Aggregator

The first event is the arrival of the E2E Path message at the Aggregator. Standard RSVP procedures are followed for this path message (including update of the PHOP field to a local Aggregator address) augmented with the extensions documented in this section.

The Aggregator first attempts to map the E2E reservation onto a TE tunnel. This decision is made in accordance with routing information as well as any local policy information that may be available at the Aggregator. Examples of such policies appear in the following paragraphs. Just for illustration purposes, among many other criteria, such mapping policies might take into account the Intserv service type, the Application Identity [[RSVP-APPID](#)] and/or the signaled preemption [[RSVP-PREEMP](#)] of the E2E reservation (for example, the aggregator may take into account the E2E reservations RSVP preemption priority and the MPLS TE Tunnel set-up and/or hold priorities when mapping the E2E reservation onto an MPLS TE tunnel).

There are situations where the Aggregator is able to make a final mapping decision. That would be the case, for example, if there is a single TE tunnel towards the destination and if the policy is to map any E2E RSVP reservation onto TE Tunnels.

There are situations where the Aggregator is not able to make a final determination. That would be the case, for example, if routing

identifies two DS-TE tunnels towards the destination, one belonging to DS-TE Class-Type 1 and one to Class-Type 0, if the policy is to map Intserv Guaranteed Services reservations to a Class-Type 1 tunnel and Intserv Controlled Load reservations to a Class-Type 0 tunnel, and if the E2E RSVP Path message advertises both Guaranteed Service and Controlled Load.

Whether final or tentative, the Aggregator makes a mapping decision and selects a TE tunnel. Before forwarding the E2E Path message towards the receiver, the Aggregator should update the ADSPEC inside the E2E Path message to reflect the impact of the MPLS TE cloud onto the QoS achievable by the E2E flow. This update is a local matter and may be based on configured information, on information available in the MPLS TE topology database, on the current TE tunnel path, on information collected via RSVP-TE signaling, or combinations of those.

The Aggregator then forwards the E2E Path message. In accordance with [\[LSP-HIER\]](#), the E2E Path message is:

- sent with an IF_ID RSVP_HOP object instead of an RSVP_HOP object. The data interface identification identifies the TE Tunnel
- addressed directly to the Deaggregator. The destination address of the E2E Path message is set to the Deaggregator address and the Router Alert is not set. Thus, the E2E Path message will not be visible to Transit routers along the path of the TE tunnel. Thus, in contrast to the procedures of [\[RSVP-AGGR\]](#), the IP Protocol number need not be modified to "RSVP-E2E-IGNORE"; it is left as is (indicating "RSVP").

[3.3.](#) Handling of E2E Path message By Transit LSRs

Since the E2E Path message is addressed directly to the Deaggregator and does not have Router Alert set, it is hidden from all transit LSRs.

[3.4.](#) Receipt of E2E Path Message by Deaggregator

On receipt of the E2E Path message addressed to it, the Deaggregator will notice that the IP Protocol number is set to "RSVP" and will thus perform RSVP processing of the E2E Path message.

As with [\[LSP-HIER\]](#), the IP TTL vs. RSVP TTL check must not be made. The Deaggregator is informed that this check must not be made because of the presence of the IF_ID RSVP HOP object. As with [\[LSP-HIER\]](#), the following checks should be made by the receiver Y of the IF_ID RSVP_HOP object:

1. Make sure that the data interface identified in the IF_ID RSVP_HOP object actually terminates on Y.
2. Find the "other end" of the above data interface, say X. Make sure that the PHOP in the IF_ID RSVP_HOP object is a control channel address that belongs to the same node as X.

3.5. Handling of E2E Resv Message by Deaggregator

The Deaggregator follows standard RSVP procedures on receipt of the E2E Resv message. This includes performing admission control for the segment downstream of the Deaggregator and forwarding the E2E Resv message to the PHOP signaled earlier in the E2E Path message and which identifies the Aggregator. Since the E2E Resv message is directly addressed to the Aggregator and does not carry the Router Alert option (as per regular RSVP Resv procedures), the E2E Resv message is hidden from the transit LSRs which handle the E2E Resv message as a regular IP packet.

3.6. Handling of E2E Resv Message by the Aggregator

The Aggregator is responsible for ensuring that there is sufficient bandwidth available and reserved over the appropriate TE tunnel to the Deaggregator for the E2E reservation.

On receipt of the E2E Resv message, the Aggregator first performs the final mapping onto the final TE tunnels (if it was only a tentative mapping). If needed the Aggregator updates the ADSPEC and immediately generates an E2E Path refresh in order to provide the accurate ADSPEC information to the receiver as soon as possible.

The aggregator then calculates the size of the resource request using standard RSVP procedures. That is, it follows the procedures in [[RFC2205](#)] to determine the resource requirements from the Sender Tspec and the Flowspec contained in the Resv. It then compares the resource requests with the available resources of the selected TE tunnel.

If sufficient bandwidth is available on the final TE tunnel, the Aggregator updates its internal understanding of how much of the TE Tunnel is in use and forwards the E2E Resv messages to the corresponding PHOP.

As noted in [[RSVP-AGGR](#)], a range of policies may be applied to the re-sizing of the aggregate reservation (in this case, the TE tunnel.) For example, the policy may be that the reserved bandwidth of the tunnel can only be changed by configuration. More dynamic policies are also possible, whereby the aggregator may attempt to increase the reserved bandwidth of the tunnel in response to the amount of

allocated bandwidth that has been used by E2E reservations. Furthermore, to avoid the delay associated with the increase of the Tunnel size, the Aggregator may attempt to anticipate the increases in demand and adjust the TE tunnel size ahead of actual needs by E2E reservations. In order to reduce disruptions, the aggregator SHOULD use "make-before-break" procedures as described in [[RSVP-TE](#)] to alter the TE tunnel bandwidth".

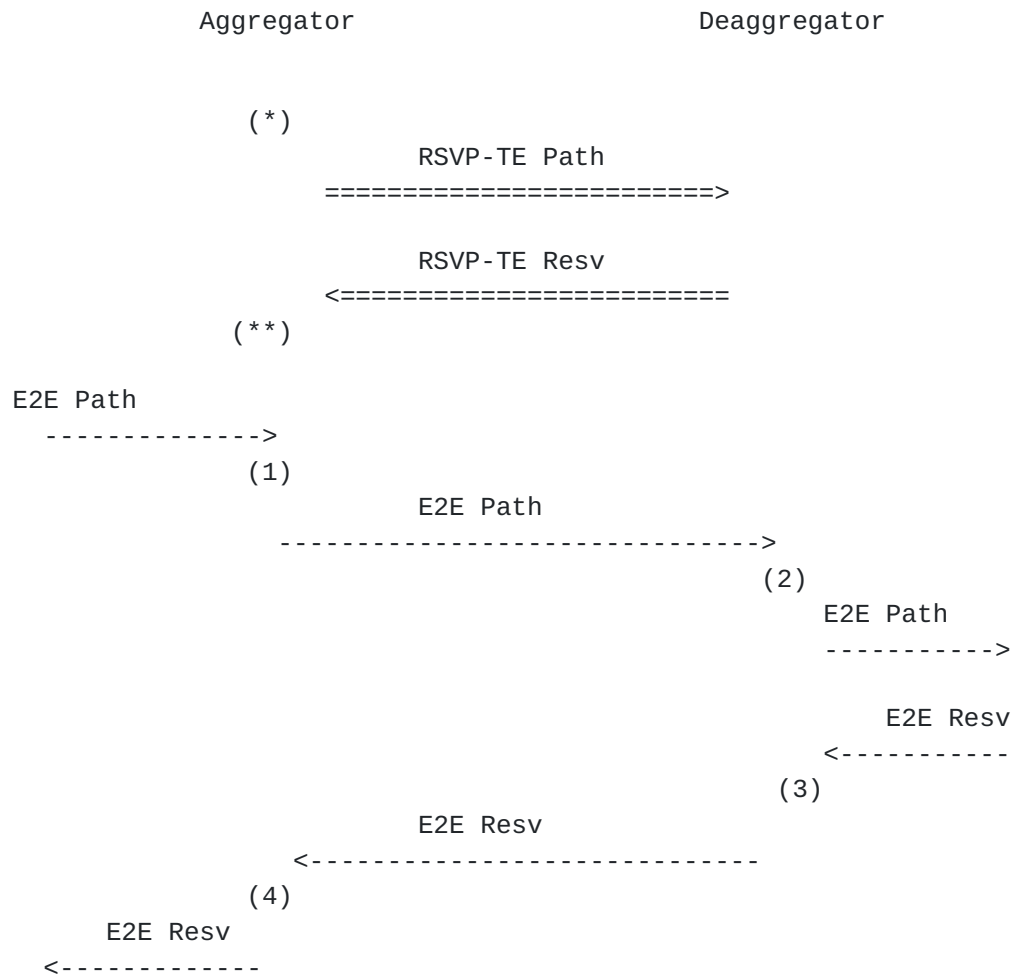
If sufficient bandwidth is not available on the final TE Tunnel, the Aggregator must follow the normal RSVP procedure for a reservation being placed with insufficient bandwidth to support this reservation. That is, the reservation is not installed and a ResvError is sent back towards the receiver.

3.7. Removal of E2E reservations

E2E reservations are removed in the usual way via PathTear, ResvTear, timeout, or as the result of an error condition. When a reservation is removed, the Aggregator updates its local view of the resources available on the corresponding TE tunnel accordingly.

3.8. Removal of TE Tunnel

Should a TE Tunnel go away (presumably due to a configuration change, route change, or policy event), the aggregator behaves much like a conventional RSVP router in the face of a link failure. That is, it may try to forward the Path messages onto another tunnel, if routing and policy permit, or it may send Path_Error messages to the sender if no suitable tunnel exists. In case the Path messages are forwarded onto another tunnel which terminates on a different Deaggregator, or the reservation is torn-down via Path Error messages, the reservation state established on the router acting as the Deaggregator before the TE tunnel went away, will time out since it will no longer be refreshed.

3.9. Example Signaling Flow

(*) Aggregator is triggered to pre-establish the TE Tunnel(s)

(**) TE Tunnel(s) are pre-established

(1) Aggregator tentatively selects the TE tunnel and forwards E2E path to Deaggregator

(2) Deaggregator forwards the E2E Path towards receiver

(3) Deaggregator forwards the E2E Resv to the Aggregator

(4) Aggregator selects final TE tunnel, check there is sufficient bandwidth on TE tunnel and forwards E2E Resv to PHOP

4. IPv4 and IPv6 Applicability

The procedures defined in this document are applicable to all the following cases:

- (1) Aggregation of E2E IPv4 RSVP reservations over IPv4 TE Tunnels
- (2) Aggregation of E2E IPv6 RSVP reservations over IPv6 TE Tunnels
- (3) Aggregation of E2E IPv6 RSVP reservations over IPv4 TE tunnels, provided a mechanism such as [\[6PE\]](#) is used by the Aggregator and Deaggregator for routing of IPv6 traffic over an IPv4 MPLS core,
- (4) Aggregation of E2E IPv4 RSVP reservations over IPv6 TE tunnels, provided a mechanism is used by the Aggregator and Deaggregator for routing IPv4 traffic over IPv6 MPLS.

5. E2E Reservations Applicability

The procedures defined in this document are applicable to many types of E2E RSVP reservations including the following cases:

- (1) the E2E RSVP reservation is a per-flow reservation where the flow is characterized by the usual 5-tuple
- (2) the E2E reservation is an aggregate reservation for multiple flows as described in [\[RSVP-AGG\]](#) where the set of flows is characterized by the <source address, destination address, DSCP>
- (3) the E2E reservation is a reservation for an IPsec protected flow. For example, where the flow is characterized by the <source address, destination address, SPI> as described in [\[RSVP-IPSEC\]](#)
- (4) the E2E reservation is an aggregate reservation for multiple flows and where the set of flows are protected by IPsec [\[RSVP-AGG-IPSEC\]](#)

6. Example Deployment Scenarios

6.1. Voice and Video Reservations Scenario

An example application of the procedures specified in this document is admission control of voice and video in environments with very high numbers of hosts. In the example illustrated below, hosts generate end-to-end per-flow reservations for each of their video streams associated with a video-conference, each of their audio streams associated with a video-conference and each of their voice

calls. These reservations are aggregated over MPLS DS-TE tunnels over the packet core. The mapping policy defined by the user maybe that all the reservations for audio and voice streams are mapped onto DS-TE tunnels of Class-Type 1 while reservations for video streams are mapped onto DS-TE tunnels of Class-Type 0.

```

-----
I H  I# -----
I    I\#I      I      -----
-----I \I Agg I      I T I      I Deag I/ -----
      I      I=====I      I
----- /I      I:::I      I:::I      I\ -----
I H  I/#I      I      -----
I    I# -----
-----

```

H = Host

Agg = Aggregator (TE Tunnel Head-end)

Deagg = Deaggregator (TE Tunnel Tail-end)

T = Transit LSR

/ = E2E RSVP reservation for a Voice flow

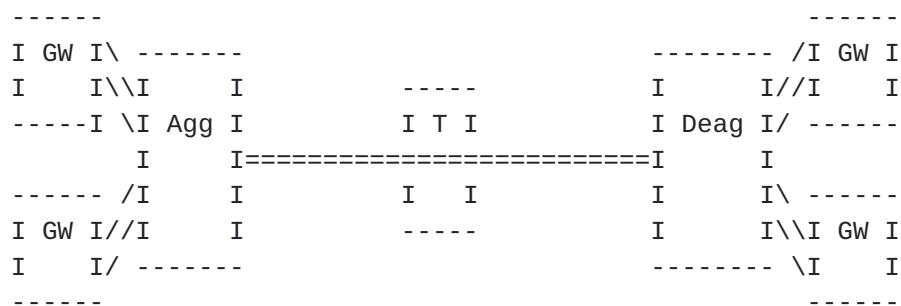
= E2E RSVP reservation for a Video flow

== = DS-TE Tunnel from Class-Type 1

:: = DS-TE Tunnel from Class-Type 0

6.2. PSTN/3G Voice Trunking Scenario

An example application of the procedures specified in this document is voice call admission control in large scale telephony trunking environments. A Trunk VoIP Gateway may generate one aggregate RSVP reservation for all the calls in place towards another given remote Trunk VoIP Gateway (with resizing of this aggregate reservation in a step function depending on current number of calls). In turn, these reservations may be aggregated over MPLS TE tunnels over the packet core so that tunnel Head-ends act as Aggregators and perform admission control of Trunk Gateway reservations into MPLS TE Tunnels. The MPLS TE tunnels may be protected by MPLS Fast Reroute. This scenario is illustrated below:



GW = VoIP Gateway

Agg = Aggregator (TE Tunnel Head-end)

Deagg = Deaggregator (TE Tunnel Tail-end)

T = Transit LSR

/ = Aggregate Gateway to Gateway E2E RSVP reservation

== = TE Tunnel

7. Optional Use of RSVP Proxy on RSVP Aggregator

A number of approaches ([[RSVP-PROXY](#)], [[L-RSVP](#)]) have been, or are being, discussed in the IETF in order to allow a network node to behave as an RSVP proxy which:

- originates the Resv Message (in response to the Path message) on behalf of the destination node
- originates the Path message (in response to some trigger) on behalf of the source node.

We observe that such approaches may optionally be used in conjunction with the aggregation of RSVP reservations over MPLS TE tunnels as specified in this document. In particular, we consider the case where the RSVP Aggregator/Deaggregator also behaves as the RSVP proxy.

As discussed in [[RSVP-PROXY](#)]:

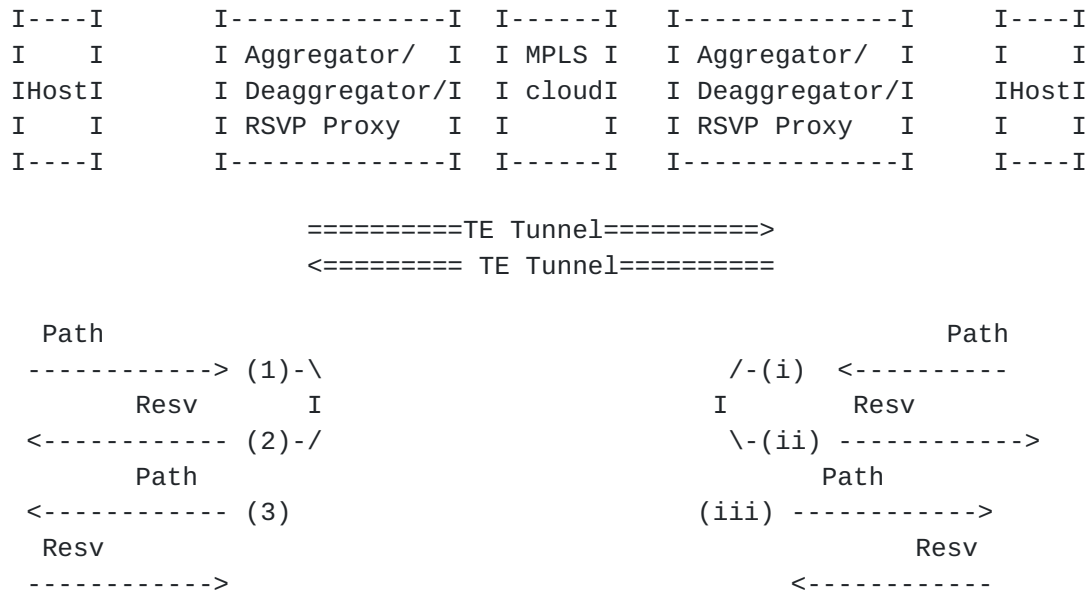
"The proxy functionality does not imply merely generating a single Resv message. Proxying the Resv involves installing state in the node doing the proxy i.e. the proxying node should act as if it had received a Resv from the true endpoint. This involves reserving resources (if required), sending periodic refreshes of the Resv message and tearing down the reservation if the Path is torn down."

Hence, when behaving as the RSVP Proxy, the RSVP Aggregator may effectively perform resource reservation over the MPLS TE Tunnel (and hence over the whole segment between the RSVP Aggregator and the RSVP

Deaggregator) even if the RSVP signaling only takes place upstream of the MPLS TE Tunnel (i.e. between the host and the RSVP aggregator).

Also, the RSVP Proxy can generate the Path message on behalf of the remote source host in order to achieve reservation in the return direction (i.e. from RSVP aggregator/Deaggregator to host).

The resulting Signaling Flow is illustrated below, covering reservations for both directions:



(1)(i) : Aggregator/Deaggregator/Proxy receives Path message, selects the TE tunnel, performs admission control over the TE Tunnel. (1) and (i) happens independently of each other.

(2)(ii) : Aggregator/Deaggregator/Proxy generates the Resv message towards Host. (2) is triggered by (1) and (ii) is triggered by (i). Before generating this Resv message, the Aggregator/Proxy performs admission control of the corresponding reservation over the TE tunnel that will eventually carry the corresponding traffic.

(3)(iii) : Aggregator/Deaggregator/Proxy generates the Path message towards Host for reservation in return direction. The actual trigger for this depends on the actual RSVP proxy solution. As an example, (3) and (iii) may simply be triggered respectively by (1) and (i).

Note that the details of the signaling flow may vary slightly depending on the actual approach used for RSVP Proxy. For example, if the [[L-RSVP](#)] approach was used instead of [[RSVP-PROXY](#)], an additional PathRequest message would be needed from host to

Aggregator/Deaggregator/Proxy in order to trigger the generation of the Path message for return direction.

But regardless of the details of the call flow and of the actual RSVP Proxy approach, RSVP proxy may optionally be deployed in combination with RSVP Aggregation over MPLS TE Tunnels, in such a way which ensures (when used on both the Host-Aggregator and Deaggregator-Host sides, and when both end systems support RSVP) that:

- (i) admission control and resource reservation is performed on every segment of the end-to-end path (i.e. between source host and Aggregator, over the TE Tunnel between the Aggregator and Deaggregator which itself has been subject to admission control by MPLS TE, between Deaggregator and destination host)
- (ii) this is achieved in both direction
- (iii) RSVP signaling is localized between hosts and Aggregator/Deaggregator, which may result in significant reduction in reservation establishment delays (and in turn in post dial delay in the case where these reservations are pre-conditions for voice call establishment), particularly in the case where the MPLS TE tunnels span long distances with high propagation delays.

8. Security Considerations

The security issues inherent to the use of RSVP, RSVP Aggregation and MPLS TE apply. Those can be addressed as discussed in [[RSVP](#)], [[RSVP-AGG](#)] and [[RSVP-TE](#)].

In addition, in the case where the Aggregators dynamically resize the TE tunnels based on the current level of reservation, there are risks that the TE tunnels used for RSVP aggregation hog resources in the core which could prevent other TE Tunnels from being established. There are also potential risks that such resizing results in significant computation and signaling as well as churn on tunnel paths. Such risks can be mitigated by configuration options allowing control of TE tunnel dynamic resizing (maximum Te tunnel size, maximum resizing frequency,...) and/or possibly by the use of TE preemption.

9. IANA Considerations

This document has no actions for IANA.

10. Acknowledgments

This document builds on the [[RSVP-AGGR](#)], [[RSVP-TUN](#)] and [[LSP-HIER](#)] specifications. Also, we would like to thank Tom Phelan, John Drake and Arthi Ayyangar for their input into this document.

11. Normative References

- [RFC2119] S. Bradner, Key words for use in RFCs to Indicate Requirement Levels, [RFC2119](#), March 1997.
- [RFC3668] S. Bradner, Intellectual Property Rights in IETF Technology, [RFC 3668](#), February 2004.
- [BCP-78], S. Bradner, IETF Rights in Contributions, [RFC3978](#), March 2005.
- [INT-SERV] Braden, R., Clark, D. and S. Shenker, Integrated Services in the Internet Architecture: an Overview, [RFC 1633](#), June 1994.
- [GUARANTEED] Shenker et al., Specification of Guaranteed Quality of Service, [RFC2212](#)
- [CONTROLLED] Wroclawski, Specification of the Controlled-Load Network Element Service, [RFC2211](#)
- [DIFFSERV] Blake et al., An Architecture for Differentiated Services, [RFC 2475](#)
- [INT-DIFF] A Framework for Integrated Services Operation over Diffserv Networks, [RFC 2998](#), November 2000.
- [RSVP-AGGR] Baker et al, Aggregation of RSVP for IPv4 and IPv6 Reservations, [RFC 3175](#), September 2001.
- [RSVP-TE] Awduche et al, RSVP-TE: Extensions to RSVP for LSP Tunnels, [RFC 3209](#), December 2001.
- [DSTE-PROTO] Le Faucheur et al, Protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering, [RFC 4124](#), June 2005.
- [LSP-HIER] Kompella et al, LSP Hierarchy with Generalized MPLS TE, work in progress

12. Informative References

[DIFF-MPLS] Le Faucheur et al, MPLS Support of Diff-Serv, [RFC3270](#), May 2002.

[DSTE-REQ] Le Faucheur et al, Requirements for support of Diff-Serv-aware MPLS Traffic Engineering, [RFC3564](#), July 2003.

[6PE] De Clercq et al, Connecting IPv6 Islands over IPv4 MPLS using IPv6 Provider Edge Routers (6PE), work in progress

[RSVP-IPSEC] Berger et al, RSVP Extensions for IPSEC Data Flows, [RFC 2207](#)

[RSVP-AGG-IPSEC] Le Faucheur et al, Generic Aggregate RSVP Reservations, [draft-lefaucheur-rsvp-ipsec-01.txt](#), work in progress

[RSVP-TUN] Terzis et al., RSVP Operation Over IP Tunnels, [RFC 2746](#), January 2000

[RSVP-PREEMP] Herzog, Signaled Preemption Priority Policy Element, [RFC 3181](#)

[L-RSVP] Manner et al., Localized RSVP, [draft-manner-lrsvp-04.txt](#), work in progress.

[RSVP-PROXY] Gai et al., RSVP Proxy, [draft-ietf-rsvp-proxy-03.txt](#) (expired), work in progress.

[RSVP-APPID] Bernet et al., Identity Representation for RSVP, [RFC 3182](#).

[AUTOMESH] Vasseur and Leroux, Routing extensions for discovery of Multiprotocol (MPLS) Label Switch Router (LSR) Traffic Engineering (TE) mesh membership, [draft-vasseur-ccamp-automesh-00.txt](#), work in progress.

[SIP-RSVP] Camarillo, Integration of Resource Management and Session Initiation Protocol (SIP), [RFC 3312](#)

13. Authors Address:

Francois Le Faucheur
Cisco Systems, Inc.
Village d'Entreprise Green Side - Batiment T3
400, Avenue de Roumanille
06410 Biot Sophia-Antipolis
France
Email: flefauch@cisco.com

Michael DiBiasio
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719
USA
Email: dibiasio@cisco.com

Bruce Davie
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719
USA
Email: bdavie@cisco.com

Christou Christou
Booz Allen Hamilton
8283 Greensboro Drive
McLean, VA 22102
USA
Email: christou_chris@bah.com

Michael Davenport
Booz Allen Hamilton
8283 Greensboro Drive
McLean, VA 22102
USA
Email: davenport_michael@bah.com

Jerry Ash
AT&T
200 Laurel Avenue
Middletown, NJ 07748, USA
USA
Email: gash@att.com

Bur Goode
AT&T
32 Old Orchard Drive
Weston, CT 06883
USA
Email: bgoode@att.com

14. IPR Statements

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard.

Please address the information to the IETF at ietf-ipr@ietf.org.

15. Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

16. Copyright Notice

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This Appendix presents an example scenario where the mechanisms described in this document are used, in combination with other mechanisms specified by the IETF, to achieve Call Admission Control of Voice over IP (VoIP) traffic over the packet core.

The information in this Appendix is purely informational and illustrative.

Consider the scenario depicted in Figure A1. VoIP Gateways GW1 and GW2 are both signaling and media gateways. They are connected to an MPLS network via edge routers PE1 and PE2, respectively. In each direction, a DSTE tunnel passes from the head-end edge router, through core network P routers, to the tail-end edge router. GW1 and GW2 are RSVP-enabled. The RSVP reservations established by GW1 and GW2 are aggregated by PE1 and PE2 over the DS-TE tunnels. For reservations going from GW1 to GW2, PE1 serves as the aggregator/head-end and PE2 serves as the de-aggregator/tail-end. For reservations going from GW2 to GW1, PE2 serves as the aggregator/head-end and PE1 serves as the de-aggregator/tail-end.

To determine whether there is sufficient bandwidth in the MPLS core to complete a connection, the originating and destination GWs each send for each connection a Resource Reservation Protocol (RSVP) bandwidth request to the network PE router to which it is connected. The bandwidth request takes into account VoIP header compression, where applicable. As part of its Aggregator role, the PE router effectively performs admission control of the bandwidth request generated by the GW onto the resources of the corresponding DS-TE tunnel.

In this example, in addition to behaving as Aggregator/Deaggregator, PE1 and PE2 behave as RSVP proxy. So when a PE receives a Path message from a GW, it does not propagate the Path message any further. Rather, the PE performs admission control of the bandwidth signaled in the Path message over the DSTE tunnel towards the destination. Assuming there is enough bandwidth available on that tunnel, the PE adjusts its book-keeping of remaining available bandwidth on the tunnel and generates a Resv message back towards the GW to confirm resources have been reserved over the DSTE tunnel.

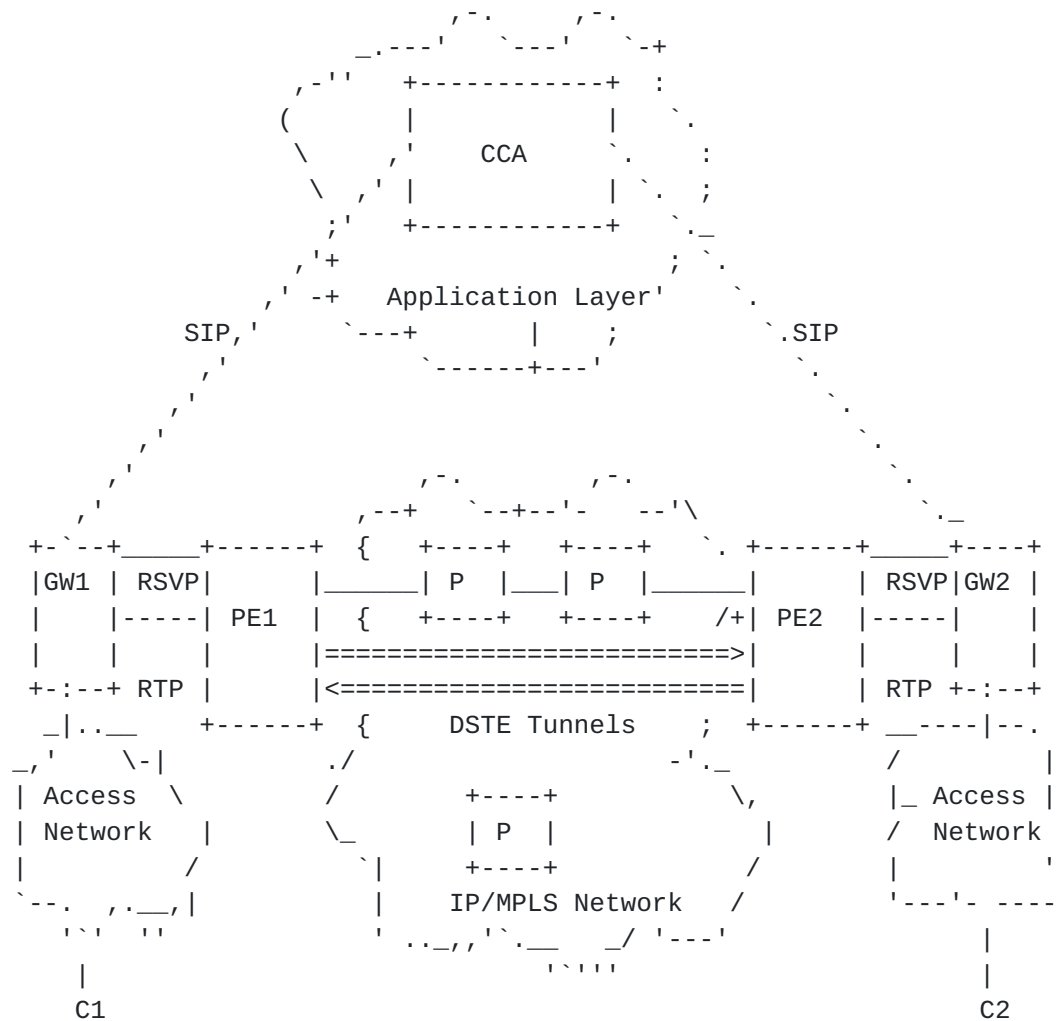


Figure A1. Integration of SIP Resource Management, DSTE and RSVP Aggregation

[SIP-RSVP] discusses how network quality of service can be made a precondition for establishment of sessions initiated by the Session Initiation Protocol (SIP). These preconditions require that the participant reserve network resources before continuing with the session. The reservation of network resources are performed through a signaling protocol such as RSVP.

Our example environment relies of [SIP-RSVP] to synchronize RSVP bandwidth reservations with SIP. For example, the RSVP bandwidth requests may be integrated into the call setup flow as follows (See call setup flow diagram in Figure A2):

- Caller C1 initiates a call by sending a SIP INVITE to VoIP

gateway GW1, which passes the INVITE along to the call control

agent (CCA). The INVITE message may contain a list of codecs that the calling phone can support.

- VoIP gateway GW2, chooses a compatible codec from the list and responds with a SIP message 183 Session Progress.
- When GW1 receives the SIP response message and learns the codec to be used, it knows how much bandwidth is required for the call.
- GW1 sends an RSVP Path message to PE1, requesting bandwidth for the call.
- GW2 also sends an RSVP Path message to PE2.
- Assuming that the tunnel (from left to right) has sufficient bandwidth, PE1 responds to GW1 with a Resv message
- Again assuming the tunnel (from right to left) has sufficient bandwidth, PE2 responds to GW2 with a Resv message
- GW2 sends a SIP 200 OK message to GW1.
- GW1 sends a SIP UPDATE message to GW2.
- Upon receiving the UPDATE, GW2 sends the INVITE to the destination phone, which responds with SIP message 180 RINGING.
- When (and if) the called party answers, the destination phone responds with another SIP 200 OK which completes the connection and tells the calling party that there is now reserved bandwidth in both directions so that conversation can begin.
- RTP media streams in both directions pass through the DSTE tunnels as they traverse the MPLS network.

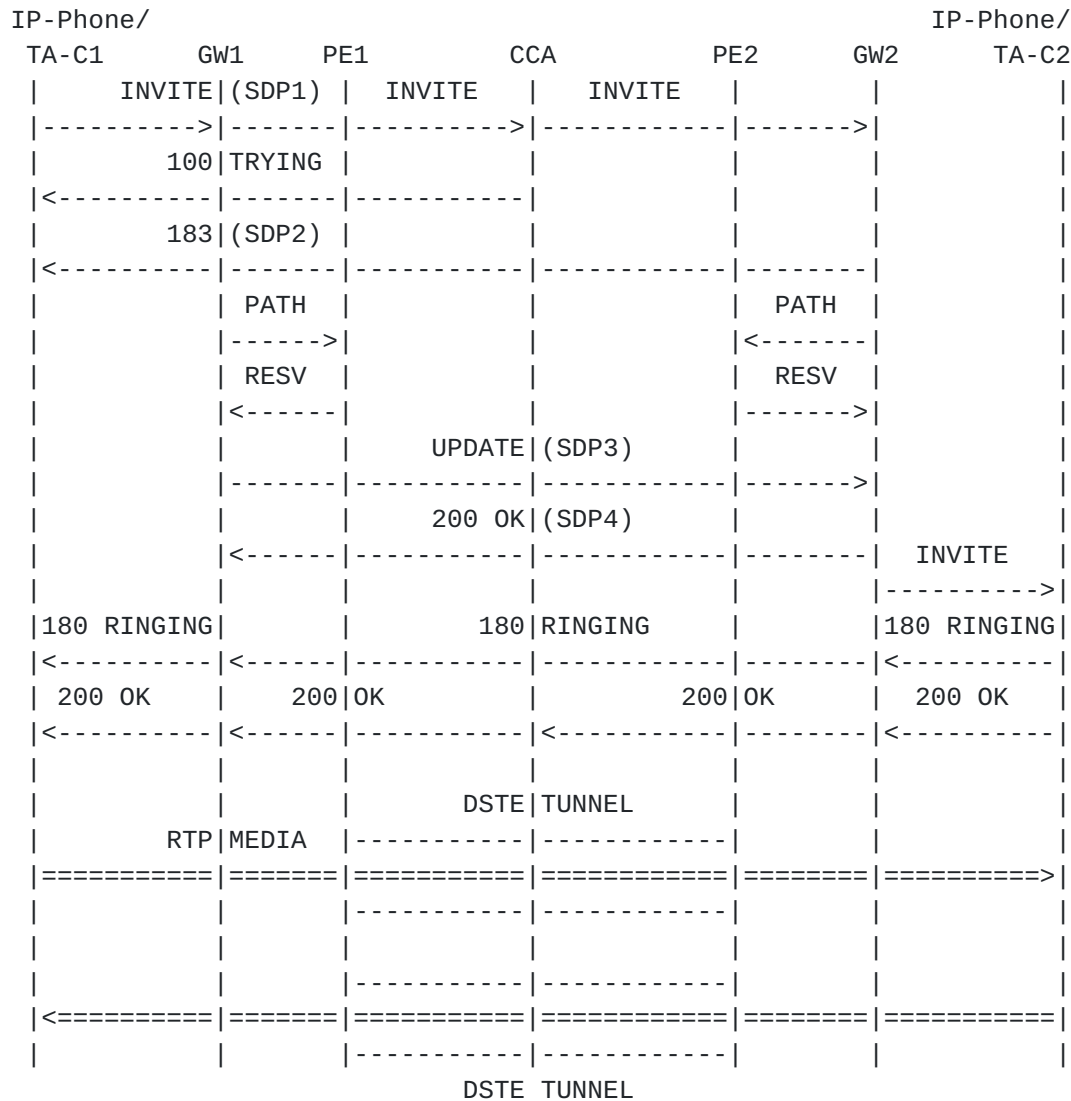


Figure A2. VoIP QoS CAC using SIP with Preconditions

Through the collaboration between SIP resource management, RSVP signaling, RSVP Aggregation and DS-TE as described above, we see that:

- the PE and GW collaborate to determine whether there is enough bandwidth on the tunnel between the calling and called GWs to accommodate the connection,
- the corresponding accept/reject decision is communicated to the GWs on a connection-by-connection basis, and
- the PE can optimize network resources by dynamically adjusting the bandwidth of each tunnel according to the load over that tunnel. For example, if a tunnel is operating near capacity, the network may dynamically adjust the tunnel size within a set of parameters.

We note that admission Control of voice calls over the core network capacity is achieved in a hierarchical manner whereby:

- DSTE tunnels are subject to CAC over the resources of the MPLS TE core
- Voice calls are subject to CAC over the DSTE tunnel bandwidth

This hierarchy is a key element in the scalability of this CAC solution for voice calls over an MPLS Core.

It is also possible for the GWs to use aggregate RSVP reservations themselves instead of per-call RSVP reservations. For example, instead of setting one reservation for each call GW1 has in place towards GW2, GW1 may establish one (or a small number of) aggregate reservations as defined in [[RSVP-AGGR](#)] which is used for all (or a subset of all) the calls towards GW2. This effectively provides an additional level of hierarchy whereby:

- - DSTE tunnels are subject to CAC over the resources of the MPLS TE core
 - Aggregate RSVP reservations from GW to GW are subject to CAC over the DSTE tunnels (as per the "RSVP Aggregation over TE Tunnels" procedures defined in this document)
 - Voice calls are subject to CAC by the GW over the aggregate reservation towards the appropriate destination GW.

This pushes even further the scalability limits of this voice CAC architecture.

