Y.RACF-DistribMPLS "Distributed RACF Architecture for MPLS Networks"

ABSTRACT

This is the new baseline document for Draft Recommendation Y.RACF-DistribMPLS. The following revisions were approved during the January 2008 Q4/13 meeting:

- Clause 1 (Scope) revised to reflect distributed nature of TRC-FE
- Clause 5 (Conventions) added
- Clause 6 (Distributed RACF Architecture and Supporting Requirements) – New clause added with TRC-FE architecture description (sub-clause 6.1) and supporting requirements (sub-clause 6.2)
- Appendix I (MPLS CAC Capabilities) revised
- Appendix II (CAC Process with Distributed Architecture) – New appendix describing CAC process.

This is the output baseline document from the January 2008 Q4/13 meeting.

TSB Note: All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

1. Scope

ITU-T Recommendation Y.2111 [Y.2111] defines general requirements for resource and admission control functions for Next Generation Networks (NGN). These requirements are independent of the underlying transport technology.

Multi-Protocol Label Switching (MPLS) is considered to be a key transport technology in core networks. In particular, MPLS with Traffic Engineering (MPLS-TE) capabilities provide significant assurance in the delivery of desired Quality of Service (QoS) for a variety of services and applications.

RACF entities control the flow of traffic from access networks into the MPLS core [Y.2111]. The architectural structure of the Transport Resource Control Functional Entity (TRC-FE) is considered in this Recommendation from a distributed perspective whereby MPLS transport resource information is available in Label Edge Routers (LER) and Gateways connected to the LERs. This Recommendation defines the distributed architecture for the TRC-FE and specifies supporting requirements.

2. References

Editor's Note: All references and their formats/style have to be updated in the document.

1. ITU-T Recommendation Y.2111, Resource and Admission Control Functions
2. IETF RFC 3270 "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services"
3. IETF RFC 3814 "Multiprotocol Label Switching (MPLS) Forwarding Equivalence Class To Next Hop Label Forwarding Entry (FEC-To-NHLFE) Management Information Base (MIB)"
4. ITU-T Recommendation Y.1711
5. ITU-T Recommendation Y.2171, "Admission Control Priority Levels in Next Generation Networks"
6. RFC 3031 – MPLS Architecture
7. RFC 2702 – RSVP Traffic Engineering
8. RFC 3564 – DiffServ Aware MPLS Traffic Engineering
9. RFC 4127 - Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering
10. RFC 4125 - Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering
11. RFC 4126 – Max Allocation with Reservation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering & Performance Comparisons
12. RFC4377 - OAM Requirements for MPLS Networks
13. RFC 4216- MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements
14. [RFC3175] "Aggregation of RSVP for IPv4 and IPv6 Reservations", IETF RFC 3175, September 2001.
15. [RFC4804] "Aggregation of Resource Reservation Protocol (RSVP) Reservations over MPLS-TE/DS-TE Tunnels", IETF RFC 4804, February 2007.
16. [RFC4090] "Fast Re-route Extensions for RSVP-TE for LSP Tunnels", IETF RFC 4090, May 2005.

3. Definitions

For the consistency between RACF, this document uses the same definitions as defined in Y.2111.

4. Abbreviations

Editor's Note: All abbreviations/acronyms have to be checked and updated in this document.

MPLS: Multiprotocol Label Switching
LSP: Label Switching Path
LER: Label Edge Router
LSR: Label Switching Router
QoS: Quality of Service
RACF: Resource and Admission Control Functions
TRC-FE: Transport Resource Control Functional Entity
TRE-FE: Transport Resource Enforcement Functional Entity
E-LSP: EXP inferred LSP
L-LSP: Label inferred LSP
CAC: Connection Admission Control
FTN: FEC-To-NHLFE
FEC: Forward Equivalence Class
NHLFE: Next Hop Label Forwarding Entry
TE: Traffic Engineering
OAM: Operation, Administration and Maintenance
TEF: Traffic Engineering Function.
OSS: Operations and Support System
NMS: Network Management System

5. Conventions

None

6. Distributed RACF Architecture and Supporting Requirements

A high level view of the role of RACF in MPLS core networks is depicted in Figure 1. Note that this figure is logical. It does not preclude the case where the service stratum spans across multiple service provider domains. Real-time traffic is aggregated by access networks into the MPLS core; there may be multiple core networks – also referred to as autonomous administrative domains. Inside the MPLS core, Label Edge Routers are assumed to be connected via pre-provisioned Label Switched Paths (LSP) or Traffic Engineering (TE) tunnels. The terms LSP and TE tunnels are synonymous in this document. A brief overview of MPLS-TE capabilities to support admission control is described in Appendix 1.

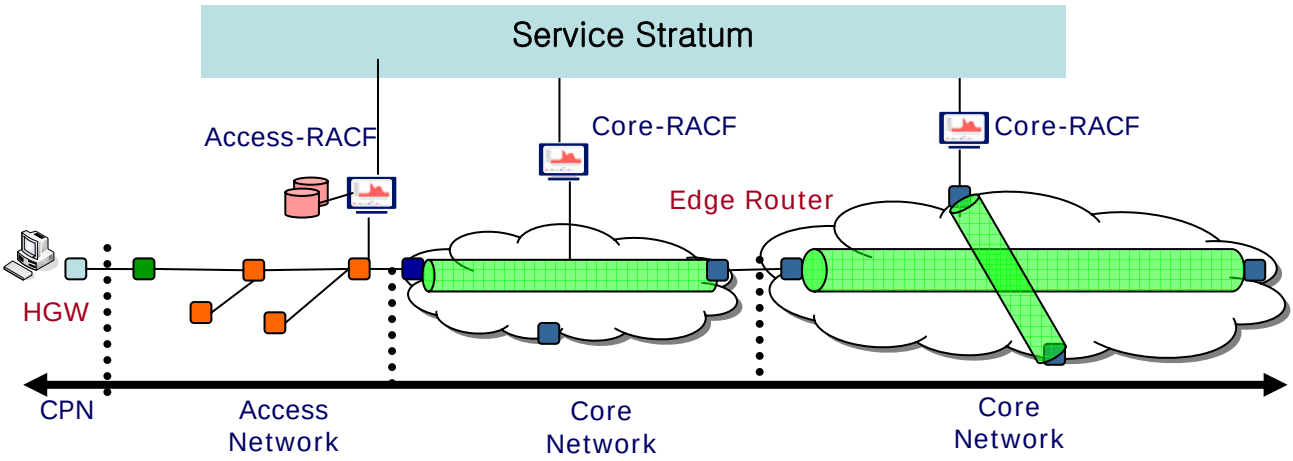

Figure 1 – Network Architecture Overview

Admission of traffic flows into these tunnels requires careful attention. The admission control process needs to determine the type of incoming service/application flows (real-time calls, interactive sessions, or data flows) as well as their QoS and bandwidth requirements. The initial point of contact for incoming flows is typically a Gateway which is connected to the LER. The Session Border Controller (SBC) serves as the Gateway for real-time applications [Y.2012-Sup1]. The processing of incoming flows commences at the Gateways (or SBC) prior to admission control processing at the Policy Decision Functional Entity (PD-FE). Once the QoS and priority requirements have been established by the PD-FE, the flows are directed to the LER where they are mapped into the proper tunnels. Information on tunnel status and bandwidth utilization conditions is assumed to be available at the LERs and the connected Gateways – tunnel selection for incoming flows is based on this information.

The architectural structure of the Transport Resource Control Functional Entity (TRC-FE) is considered in this Recommendation from a distributed perspective whereby MPLS transport resource information is available in Label Edge Routers (LER) and Gateways connected to the LERs. This clause defines the distributed architecture for the TRC-FE and specifies supporting requirements.

6.1 – Distributed TRC-FE Architecture in MPLS-TE Core Networks

A distributed implementation for the TRC-FE is shown in Figure 2. It is based on the capabilities of the MPLS Label Edge Router (LER) and the Session Border Controller (SBC) which serves as the Gateway connected to the LER.

IETF RFC 3175 [RFC3175] and RFC 4804 [RFC4804] define a process by which Gateways connected to LERs in an MPLS/DS-TE network, are involved in the admission control of real-time applications into pre-established DiffServ-Aware MPLS Traffic Engineering (DS-TE) tunnels between the LER pairs. In an NGN environment, a Session Border Controller provides key functionality as a “Gateway” for real-time applications [Y.2012-Sup1] and these SBC’s are connected to the LERs. SBC functions are defined as Signaling Path functions (e.g., Signaling Protocol Translation and Inter-working, Authentication, Authorization, and Accounting, Session-based Routing) and Media Path Functions (e.g., Policing and Marking, Resource and Admission Control, Opening and Closing of Pinhole/Firewall). Media Path functions such as Policing and Marking and Opening and Closing of Pinhole/Firewall provide policy enforcement functionality for incoming flows. The SBC therefore serves the role of PE-FE in this architecture.

The Media Path Resource and Admission Control or CAC function also provides an initial level of TRC-FE functionality. It is required to be RSVP-enabled. It is also required to keep track of

available reserved bandwidth with other SBCs in the MPLS-TE core. Thus decision making for additional bandwidth reservations is carried by this CAC function based on information regarding current bandwidth states between pairs of SBCs. The CAC function and associated requirements defined below is described here as part of the SBC Media Path. However, it should be noted that such a CAC function can be implemented in any generic Gateway or Border Element, rather than being limited as an SBC Media Path function.

The bandwidth status of DS-TE tunnels is not required in the SBC CAC function – tunnel states are maintained by the LERs in the core. The LER is part of the Autonomous System domain and it receives the entire route and link status information via the IP and MPLS routing protocols. LERs are required to be RSVP enabled in order to establish DS-TE tunnels and reserve bandwidth in these tunnels [RFC3175]. Thus, they maintain the state of all originating and terminating DS-TE tunnels. State information includes tunnel attributes as well as tunnel bandwidth characteristics (available bandwidth, bandwidth upper bound, etc). In this role, the LER provides the next layer of TRC-FE functionality.

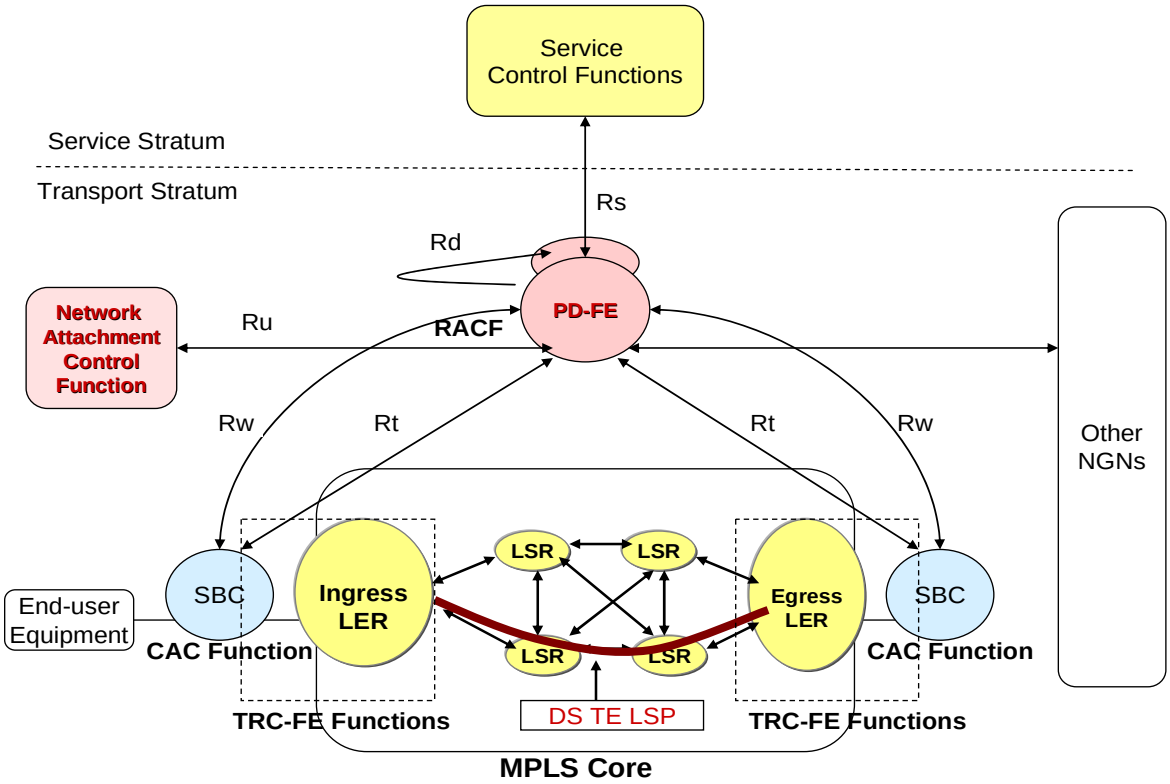

Figure 2 - Distributed TRC-FE Architecture

The Rc interface functionality is already provided at the LER by virtue of supporting IP and MPLS protocols as well as DS-TE tunnel state information. It is also supported by requiring the SBC CAC functions to keep track of bandwidth status between all SBC pairs. Further, the LERs also play the role of Transport Resource Enforcement Functional Entity (TRE-FE) by virtue of their capability to establish TE tunnels and reserve bandwidth. Hence, the Rc and Rn Reference Points are not required in this implementation. However, the Rt interface is required between the PD-FE and the SBC CAC function for uni-directional and symmetric bi-directional flows.

Since this MPLS-specific transport knowledge is distributed across all LER-SBC combinations in the core, this architecture is considered to be distributed. The TRC-FE function is jointly carried out by the SBC CAC function (determination of bandwidth state with other SBCs) and the LERs

(maintenance of DS-TE tunnels between LER pairs). Once available bandwidth has been established, the SBC CAC function admits the media flow request.

Bandwidth and QoS requirements for incoming flows are submitted to the initiating SBC CAC function by the PD-FE. The PD-FE also determines the destination SBC for the call/session flow. The detailed mechanisms for determining the destination SBC are for further study. The SBC CAC function determines if sufficient bandwidth is available between itself and the terminating SBC. If so, the SBC CAC function admits the media flow. If not, then the initiating and terminating SBC CAC functions submit RSVP reservation requests for additional bandwidth prior to admitting the flow request. Bandwidth reservation can be on a per-call or aggregated basis. The respective LERs identify the specific bi-directional tunnel – based on inputs received from the PD-FE - and then attempt to reserve additional bandwidth. If the reservation is successful, the SBC admits the media flow. A detailed description of the admission control process is provided in Appendix II.

6.2 – Requirements to Support Distributed Architecture

The SBC CAC function provides an initial level of TRC-FE functionality (tracking reserved bandwidth availability between pairs of SBCs). The following requirements are necessary to support the CAC function:

1. The CAC function is required to track bandwidth reservation status between the SBC and all other SBCs in the core network. For each SBC pair, the CAC function needs to keep track of the total amount of bandwidth reserved and the current value of reserved bandwidth utilization as described in Appendix 2.
2. The CAC function is required to interact with the PD-FE via the Rt Reference Point in order to receive necessary call/session flow input such as Source, Destination, and DSCP. Protocol requirements to support this Reference Point are for further study.
3. The CAC function is required to be RSVP-enabled per IETF RFC 3209 [RFC3209]. This allows the CAC function to initiate RSVP reservations [RFC3175], [RFC4804].

The LER provides an additional level of TRC-FE function in this architecture by establishing DS-TE tunnels and maintaining tunnel status information which is necessary to determine further RSVP reservations. The following requirements are necessary to support the LER:

1. The LER is required to be MPLS-TE/DS-TE enabled per IETF RFC 3564 [RFC3564]. All LERS in the core should be enabled with an appropriate Bandwidth Constraint algorithm. Three choices are available: Maximum Allocation Model (MAM) [RFC4125], Maximum Allocation with Reservation Model (MAR) [RFC4126], and Russian Dolls Model (RDM) [RFC4127]. Note that all LERs in the administrative domain must have the same Bandwidth Constraint algorithm.
2. The LER is required to map the flow into the proper tunnel according to the process described in IETF RFC 3814 [RFC3814].

The LER is required to be RSVP-enabled such that bandwidth reservations in the TE/DS-TE tunnels are supported as described in IETF RFC 3175 [RFC3175] and RFC 4804 [RFC4804]. Note that bandwidth reservations can be on a per-call/session basis or on an aggregated basis – the choice is up to the network operator.

7. Security Considerations

Editor’s Note: Text for this section needs to be provided.

Appendix I - MPLS Traffic Engineering Capabilities to Support CAC

QoS in the core network is enhanced by the traffic engineering capabilities of MPLS-TE. MPLS has been designed to overcome IP's least cost routing limitation and provides ATM-like QoS capabilities via Traffic Engineering methods. It relies on the IP routing protocols to build the network map and assigns labels to links that meet certain quality criteria via the Label Distribution Protocol (LDP) or signaling protocols such as RSVP-TE [RFC2702]. Packet forwarding is based on swapping labels. Differentiated Services (DiffServ) provides a QoS treatment to traffic aggregates. It is scalable and does not require per-flow signaling. However, it cannot guarantee QoS and it does not influence the path of a packet. MPLS can force packets into specific paths and, in combinations with constraint-based routing, can guarantee bandwidth for Forwarding Equivalency Classes. But MPLS by itself cannot specify class-based differentiated treatment of packet flows.

Combining DiffServ and MPLS-based TE can lead to true QoS in IP packet backbones. To achieve this functionality, networks have to be carefully engineered with traffic engineering applied on a per-class basis – DiffServ-Aware MPLS Traffic Engineering (DS-TE) [RFC3564]. The goal of DS-TE is to guarantee bandwidth separately for critical traffic types (e.g., emergency telecommunications) such that the QoS requirements compliance for that traffic type is improved and optimized. It is assumed that the majority of traffic is of the "Best Effort" traffic class. Under congestion or failure conditions, this traffic class will have reduced bandwidth available in order to ensure that more critical traffic classes have the required bandwidth to meet their priority and QoS requirements. An aggregated grouping of Traffic Trunks based on the class of service requirements such that they share the same bandwidth reservation is called Class Type (CT). Up to eight Class Types are allowed. Each CT has two attributes:

- Bandwidth Constraint (BC) – A limit on the percentage of a link's bandwidth that a particular CT may request. Three Bandwidth Constraint models have been developed: Maximum Allocation Model (MAM) [RFC4125], Maximum Allocation with Reservation Model (MAR) [RFC4126], and Russian Dolls Model (RDM) [RFC4127]. Note that each MPLS network domain can be provisioned to support only one of these BC models. All tunnels in a domain have to be governed by the rules of one model only.
- Preemption Priority (p) – The relative importance of a given CT compared to others. This priority enables the DS-TE bandwidth constraint models to release shared bandwidth occupied by a lower priority CT when higher priority CT traffic arrives during conditions of congestion or failure.

Additional attributes associated with incoming flows are as follows:

- DiffServ Code Point (DSCP): The desired QoS of an incoming flow can be characterized by delay, packet loss, and jitter requirements. These requirements can be summarized by assigning appropriate DSCP values to the media stream. For example, stringent delay, loss, and jitter requirements are characteristics of Voice over IP (VoIP) services. These services can then be characterized by the Expedited Forwarding (EF) DSCP as this code point refers to stringent QoS requirements.
- Restoration/Re-route Priority: This is the priority with which a tunnel can be restored [Y.2172] or mapped into an MPLS Fast Re-route priority [RFC4090].

From the perspective of admission control, an incoming service or application seeking entry into the MPLS network needs to be "mapped" into the TE tunnel of the appropriate Class Type. Such a mapping can be done by linking the service CAC priority as defined in ITU-T Recommendation Y.2171 [Y.2171] with the priority attribute of the appropriate Class Type that supports the desired QoS requirements of the incoming service.

Admission control policy enforcement for unicast real-time applications (e.g., Voice over IP calls, video services) may be accomplished for MPLS-TE networks based on functionality described in IETF RFC 3175 [RFC3175] and RFC 4804 [RFC4804]. Policies for multicast applications are for further study. The underlying premise is the setup of TE Tunnels between pairs of LERs with sufficient bandwidth such that a significant number of calls/sessions for various applications can traverse these tunnels without additional processing of signaling messages on the intermediate routers along the path of the tunnel.

A network that is DiffServ-enabled and is RSVP-aware offers several mechanisms to support topologically aware admission control that can be per-flow control or on aggregated flows. The advantage of aggregated RSVP reservations is that it offers dynamic admission control without the per-flow reservations and associated RSVP signaling in the DiffServ core. With DS-TE tunnels, different types of services/applications can be directed into designated tunnels identified by a unique Class Type. Further, appropriate bandwidth allocation/constraints can be enforced for these different Class Type tunnels such that service/application QoS requirements are satisfactorily met.

The benefits of aggregating RSVP reservations over TE Tunnels are as follows:

- RSVP signaling messages for bandwidth reservations are ignored by intermediate routers along the tunnel path per RFC 3175 [RFC3175]. RSVP message processing is further minimized by initiating them for aggregated reservations as opposed to per-call reservations.
- Core scalability is unaffected because the core has to simply maintain aggregated TE tunnels.
- Aggregate reservations can be network engineered with Constraint Based Routing taking advantage of alternate paths when needed.
- Aggregated reservations over TE Tunnels can be protected against failure via MPLS Fast Reroute [RFC4090].

Appendix II - Admission Control Process with Distributed Architecture

It is assumed that DS-TE tunnels are pre-established between pairs of LERs in the MPLS-TE core network and that some pre-determined amount of bandwidth is reserved in these tunnels – this is appropriate in the case where aggregated RSVP reservations is the practice. Note that additional bandwidth reservation in these tunnels may be permitted up to a specified limit.

Incoming calls/sessions are signaled to the PD-FE via the SBC. The following information and requirements about these calls/sessions are determined via appropriate interactions between the PD-FE and the Service/Application Layer:

1. Terminating SBC
2. CAC Priority
3. Bandwidth requirement (based on appropriate codec negotiation between originating and terminating SBC as well as call/session QoS requirements if any)
4. DiffServ Code Point (DSCP) based on the type of call/session and it's priority requirements.

The next step is the determination of available bandwidth between the originating and terminating SBCs to support the flow. This determination can lead to per-call bandwidth reservation or aggregated bandwidth reservation [RFC3175].

Per-call reservation provides optimal use of bandwidth. However, if the volume of busy-hour calls/sessions is significant, then the resulting RSVP processing may create an unacceptable burden on the LERs. Hence, per-call reservations should be considered only in cases where busy-hour call/session volume is not excessive.

Assuming that no bandwidth is reserved in the TE tunnels during pre-establishment, the per-call/session bandwidth reservation is as follows [RFC3175]:

1. PD-FE submits call/session input to the CAC functions in the originating and terminating SBCs. This input includes (Source, Destination, DSCP). Head-end and tail-end LERs identify appropriate DS-TE tunnel based on this input.
2. Originating SBC CAC function submits an RSVP PATH message with appropriate information to head-end LER. Similar RSVP PATH message is submitted by the terminating SBC CAC function to the tail-end LER to support the call/session in the opposite direction.
3. Head-end LER processes RSVP PATH message and determines the amount of requested bandwidth in the identified DS-TE tunnel. Similar action is carried out by the tail-end LER.
4. RSVP PATH messages are forwarded in both directions by the head-end and tail-end LERs to the terminating SBC and originating SBC. There are two cases here:
    a. If sufficient bandwidth is available in the DS-TE tunnel (both directions) to support the call/session, then the bandwidth is reserved in both directions and RSVP RESV messages are sent back to the originating and terminating SBCs which then admit the media flow in both directions.
    b. If sufficient bandwidth is not available, then the DS-TE Bandwidth Constraint algorithm (MAM, MAR, or RDM) attempts to increase the bandwidth upper limit in the tunnel by adjusting bandwidth limits from DS-TE tunnels of lower pre-emption priority:

i. If the DS-TE tunnel bandwidth limit is successfully increased, then the desired bandwidth to support the incoming call/session is reserved in both directions and the RSVP RESV message is sent back to the originating and terminating SBCs which then admit the media flow.

ii. Otherwise, an RSVP Error Code message is sent back to the initiating SBC and the call/session is denied admission.

5. Once the call/session is completed, the reserved bandwidth is released in both directions.

When busy-hour call/session volume is expected to be high, aggregated RSVP reservations are recommended. By reserving pre-determined chunks of bandwidth that can support a large number of simultaneous flows, RSVP processing on the LERs is significantly minimized; however bandwidth utilization may not be optimal.

To support aggregated reservations, the SBC CAC function is assumed to have the necessary knowledge to make the reservation decision. The SBC CAC function is required to keep track of the total amount of reserved bandwidth ***R*** and the amount of utilized bandwidth by calls/sessions in progress ***U*** between itself and all other SBCs in the network.

The detailed process for aggregated bandwidth reservations for an incoming call/session is as follows:

1. PD-FE submits call/session input to the CAC functions in the originating and terminating SBCs. This input includes (Source, Destination, DSCP).
2. Originating SBC CAC function compares the available reserved bandwidth ***R – U*** with the bandwidth requirements of the incoming call/session. If the available reserved bandwidth is sufficient, then the originating and terminating SBCs admit the media flow in both directions. The LERs map the flow into the appropriate DS-TE tunnels based on the signaled information (Source, Destination, DSCP). The originating and terminating SBCs update the respective ***U*** values for current bandwidth utilization between them.
3. If the available reserved bandwidth is not sufficient to support the bandwidth requirements of the incoming call/session, then the originating and terminating SBC CAC functions request an aggregated chunk of bandwidth reservation in the DS-TE tunnel via RSVP PATH messages as described in IETF RFC 4804.
    a. If bandwidth reservation is successful, then the media flow is admitted in both directions and the two SBCs update their ***R*** and ***U*** values.
    b. If the DS-TE tunnel bandwidth limit is reached, then the upper limit can be temporarily increased by adjusting the limits of DS-TE tunnels of lower pre-emption priorities as described above.
        i. If the adjustment is successful, the RSVP aggregated reservation process commences and the media flows can be admitted. The ***R*** and ***U*** values are then adjusted accordingly between the two SBCs.
        ii. If the adjustment is not successful, an RSVP Error Code message is sent back to the two SBCs and the call/session is denied admission.
4. Upon completion of the call/session, the respective SBC CAC functions need to adjust their reserved bandwidth utilization values (***U***) accordingly.

Periodically, per pre-determined rules, chunks of bandwidth can be released during low utilization periods. The respective SBCs need to adjust their ***R*** values accordingly.