DSL Forum

Proposed Draft
PD-021
Revision 2

Layer 2 Control Mechanism

For
Architecture and Transport Working Group

August 2004

Abstract:

The document reflects the work done by this WG so far. It explains motivation as well as the concept of the layer 2 control plane. This is illustrated by several use cases and message flows, e.g. Topology Discovery, Layer 2 OAM and Multicasting. It also contains functional and design goals for a possible protocol mapping.

Revision History

Revision Number	Date of Last Modification	Edited by	Changes
Revision 1	11-May-04	Norbert Voigt, Jerome Moisand	Outline and scope for Toronto meeting based on input from DSL Forum contributions dslf2003.367, dslf2003.368, dslf2003.376, dslf2003.434, dslf2004.034 and dslf2004.087

Revision 2	06-Aug-04	Norbert Voigt, Jerome Moisand	Update following Toronto meeting plus interim group discussions:
			• Updated history section
			  o Section 1.8
			• Clarified scope of applicability
			  o New section 1.7
			• More details on DSL line resynchronization
			  o Section 3.1
			  o new Appendix A
			• More specifics on possible DSL parameters relevant for topology discovery and line configuration
			  o Section 3.1
			  o new Appendix B
			• More specifics on distributed architectures involving remote terminals
			  o Section 3.1
			  o new Appendix C
			• IGMP-Proxy scenario: added L2-control subscriber feedback
			  o Section 3.3.6 and 4.2.8
			• New text about protocol mapping comparison reflecting narrowed-down protocol choices
			  o Section 5.2
			• Miscellaneous editorial improvements
			Toronto contributions used as inputs: dslf2004.152, dslf2004.183, dslf2004.185.

Table 1: Revision History

5

Technical comments or questions about this document should be directed to:

Jerome Moisand
Juniper Networks, Inc.
10 Technology Park Drive
Westford, MA 01886
U.S.A.
jmoisand@juniper.net

Norbert Voigt
Siemens AG
Siemensallee 1
D-17489 Greifswald
Germany
norbert.voigt@siemens.com

Table of Content

List of Figures

List of Tables

1 Introduction

1.1 DSL architectures, value-added services

DSL is a widely deployed access technology for Broadband Access for Next Generation Networks. Several documents like DSL-Forum TR-058, DSL-Forum TR-059 and ITU H.610 describe possible architectures for these access networks.

In the scope of these specifications are the delivery of voice, video and data services.

Corresponding architectures require permanent virtual circuit(s) per subscriber. Such virtual circuit is configured on layer 2 and terminated at the first layer 3 device (e.g. BRAS). Beside the data plane, the models define the element management architecture. Each Network Node is usually controlled by one element manager. These element managers should be connected to one set of network and service management systems, but practically speaking this is not always the case because of organizational boundaries between departments operating the local loop, departments operating the ATM network, and departments operating the IP network.

In any case, the management networks are usually not designed to transmit management data between the different entities in real time, in an interoperable and standardized manner.

When deploying value-added services across DSL access networks, special attention regarding quality of service and service control is required, which implies a tighter coordination between Network Nodes (e.g. Access Nodes and BRAS), without burdening the OSS layer with unpractical expectations. The following sections will provide examples of such need.

1.2 Topology discovery, motivation and concept

TR-059 identified various queuing/scheduling mechanisms to avoid congestion in the access network while dealing with multiple flows with distinct QoS requirements. Such mechanisms require that the BRAS gains knowledge about the topology of the access network, the various links being used and their respective rates. Some of the information required is somewhat dynamic in nature (e.g. DSL sync rate), hence cannot come from a provisioning and/or inventory management OSS system. Some of the information vary less frequently (e.g. capacity of a DSLAM uplink), but nevertheless needs to be kept strictly in sync between the actual capacity of the uplink and the image the BRAS has of it. Practically speaking, OSS systems are rarely able to enforce in a reliable and scalable manner the consistency of such data, notably across organizational boundaries.

Additionally, in case of an Ethernet-based access network there might not longer be a "logical circuit" or "logical path" terminated on the layer 3 device (e.g. BRAS), as a representation of the subscriber local loop. That creates in turn some challenges to properly configure the BRAS and its hierarchical scheduler.

Dynamic and automated discovery of the access network topology would help to address these issues, notably when performed by an interoperable and standardized protocol. An inband control plane is proposed that allows the Access Node (e.g. DSLAM) to communicate to the BRAS access network topology information and any corresponding updates.

1.3 Line configuration, motivation and concept

Following dynamic line identification (subscriber local loop) as assisted by the mechanism described in the previous section (topology discovery), the BRAS could then query a subscriber management OSS system (e.g. RADIUS server) to retrieve subscriber authorization data (service profiles, aka user entitlement). Most of such service mechanisms are typically enforced by the BRAS itself, but there are a few cases where it might be useful to push such service parameter to the DSLAM for local enforcement of a mechanism (e.g. DSL-related) on the corresponding subscriber line. Using an inband control plane, the BRAS could achieve such goal by an interoperable and standardized protocol.

Using such approach would dramatically simplify the OSS infrastructure for service management, allowing to fully centralize subscriber-related service data (e.g. RADIUS server back-end) and avoiding complex cross-organization B2B interactions.

1.4 Optimized multicasting, motivation and concept

Multicasting is the suitable technology for simultaneous delivery of multimedia streams to multiple recipients. Particularly, in the case of deploying video services across DSL access networks special attention regarding quality of service and service control is required.

Today, multicast streams are typically replicated on a subscriber basis at the BRAS (layer 3 Network Node) and forwarded to the Access Node (layer 2 Network Node), e.g. DSLAM. Usually, the DSLAM is located either in a central office or even further out to the subscriber as remote DSLAM whereas the BRAS is the aggregating device more centralized in a POP. Since the information is replicated at the BRAS and is sent in parallel as often as requested by the subscribers high facility costs are the result.

In order to offer profitable video services to consumers and to reuse broadband access infrastructure already in place, multicast mechanisms may need to be optimized (notably in cases where Video-On-Demand services are not yet deployed).

The efficiency of multicast services also directly affects the quality of service provided.

The proposed control plane allows to control the delivery of information to multiple destinations with a single transmission for commonly used multicast streams. Thus, the actual replication functionality at layer 2 can be provided by the Access Node (e.g. DSLAM), in order to use the bandwidth available between the Access Node and the IP network in an economic manner. The replication is still controlled by the layer 3 Network Node (BRAS), based on centralized (e.g. RADIUS-based) subscriber authorization data. The enforcement of user entitlement rules (multicast channels that can be legitimately joined by this user) continues to be performed by the device in charge of multicast routing protocols (e.g. IGMP, PIM, etc), i.e. the BRAS.

The proposed approach allows to leverage on the strength of Access Nodes (e.g. ATM replication or Ethernet bridging replication) while not disrupting normal IP multicast processing (performed by the BRAS) and keeping a simple and centralized OSS infrastructure (e.g. RADIUS) for subscriber-related service data.

1.5 OAM in a mixed ATM/Ethernet environment

Traditionally, ATM circuits are point to point connections between BRAS and DSLAM/CPE. In order to test the connectivity on layer 2 appropriate OAM functionality is used for operation and troubleshooting. Ethernet is getting a more important role as layer 2 technology and getting more attention from operators. From the operator's perspective, it is required to keep ways to test and troubleshoot connectivity in the case of a mixed Ethernet and ATM access network (including the local loop). Corresponding control plane functions must be envisioned. Considering existing ATM architecture an end to end OAM loopback is performed between the edge devices (BRAS and DSL CPE) of the broadband access network. To reach consistency in operation of a broadband access network an appropriate functionality must be implemented in Ethernet.

Statistics retrieval is another important OAM aspect to be foreseen which would require additional interactions at the control plane.

1.6 Overall concept

All these applications can be viewed as examples where there is a requirement for a control plane between a service-oriented layer 3 edge device (the BRAS) and a layer 2 Access Node (e.g. DSLAM) in order to perform QoS-related, service-related and subscriber-related operations.

It has to be noted that alternate broadband access technologies (e.g. Metro-Ethernet, Passive Optical Networking) will have similar challenges to address, and could benefit

from the same approach of a control plane between a BRAS and an Access Node (e.g. OLT), providing a unified control and management architecture for multiple access technologies, hence facilitating migration from one to the other and/or parallel deployments.

This document provides more details on how to achieve such a control plane and address the various use cases.

1.7 Scope of applicability

Currently, there are three main parallel architecture tracks supported by the DSL Forum via a technical Report, or being defined in the context of a Working Text.

Both TR-059 and WT-099 (based on ITU H.610) are ATM-based architectures, describing Broadband Access for Next Generation Networks delivering voice, video and data services in a reliable way. WT-101 expands the TR-59 logic towards Ethernet-based aggregation.

By means of the proposed layer 2 control mechanism, several improvements are possible to these three architectures, without changing their fundamental characteristics.

Layer 2 control allows for the dynamic nature of the access network's topology in order to adjust the various queuing and scheduling mechanisms inside the BRAS accordingly. This applies mostly to TR-59 and WT-101.

Multicast mechanisms need to be optimized since its efficiency directly affects the quality of the service and business models. Also WT-099 specifies the multicast replication point as close as possible to the subscriber. Even when IGMP termination and multicast replication is done in the DSLAM there is a need for subscriber multicast entitlement information at subscriber session establishment. A channel change based on the proposed layer control scheme would remove the burden of subscriber authentication, authorization and dynamic determination of multicast subscription rights from the DSLAM. All three architectures (TR-59, WT-099, WT-101) present challenges with multicast scenarios which would benefit from improvements brought by L2-control.

In case of an Ethernet-based aggregation network that is the scope of WT-101 there might not longer be a "logical circuit" or "logical path" terminated on the layer 3 device (e.g. BRAS), as a representation of the subscriber local loop. That creates in turn some challenges to properly configure the BRAS and its hierarchical scheduler. Also in this case the proposed layer 2 control mechanism could help to address these issues by a dynamic and automated discovery of the access network topology. Additionally, it could help to maintain similar layer 2 end-to-end OAM capabilities that were available with an ATM-based aggregation network.

1.8 History, Acknowledgements

At the Q4 meeting 2003 in Paris contributions by several members stating the need for a control mechanism between Access Node and BRAS were discussed:

- Layer 2 Control Mechanism (multicast, topology discovery, line configuration) – Juniper, Siemens, Ericsson (dslf2003.367 [4], dslf2003.368 [5])
- Aspects of Multicast Services – Deutsche Telekom (dslf2003.376 [6])
- DSLAM as Service Proxy For Video Distribution – ECI (dslf2003.434 [7])
- A Subscriber Update Protocol – Redback (dslf2003.441 [8]).

Architecture and Transport Working Group has decided to continue the analysis by focusing on use cases, and elaborate on multicast models (data plane and control plane). All parties having submitted contributions decided to collaborate further.

At the Q1 meeting in Brussels 2004, two contributions about several use cases as well as initial message flows were discussed:

- Layer 2 Control Mechanism – Use Cases
  by Juniper, Siemens, Deutsche Telekom, ECI, Redback (dslf2004.034) [9]
- Layer 2 Control Mechanism – Message Flows
  by Juniper, Siemens, Deutsche Telekom, ECI, Redback (dslf2004.087) [11].

At the Q2 meeting in Toronto, the first version of PD-021 was issued, plus several contributions discussed various aspects of this work:

- Comparison between possible protocol mappings for Layer2 control messages
  by Juniper, Siemens, DTAG, ECI, Redback (dsl2004.152)
- Layer2 control, additional considerations
  by Alcatel (dslf2004.183)
- Multicast in access architecture
  by Alcatel (dslf2004.185)
- Multicast scenarios, views from the customer premises
  by Juniper (dslf2004.200)

For full acknowledgment see the detailed list of references in chapter 7.

The document summarizes several use cases to give examples for applications. Different scenarios like topology discovery, line configuration, layer 2 OAM and multicast are in the scope. Furthermore, the required message types and message flows are defined.

2 Concept of Layer 2 Control Plane

2.1 Architecture

TR-058 and TR-059 propose a renovated multi-service architecture where the Access Node is the first aggregation point and specific support for IP applications is intended in the Access and Regional Network, e.g. IP-QoS and multicasting.

TR-059 defines a model where the Routing Gateway and the BRAS are the main Policy Enforcement Points receiving NSP/ASP profiles from a policy server (e.g. RADIUS and/or Configuration Server) and enforcing corresponding IP forwarding behavior on a per service/subscriber basis. Picture 20 of TR-059 shows the network topology, illustrated below.

In such architecture, the access network is not aware about service requirements which may impact its behavior, and strong assumptions are made on the OSS layer to sync up information between Network Nodes.

The approach described in this document does not propose a change to the TR-058/059 architecture but functional additions. An embedded communication (Layer 2 control) between BRAS and Access Node is intended to support control mechanisms facilitating economical service delivery, on a per subscriber session basis. The Access Node and the BRAS would then act as layer 2 control points, with a tight coordination via an in-band control protocol. To a certain extent, this could be viewed as a distributed switching fabric logic between the BRAS and the Access Node.

New service introduction and service delivery (unicast and multicast) is then enabled while minimizing the number of points to be modified and/or integrated in an extended management paradigm, yet allowing dynamic coordination between network and access nodes when necessary.

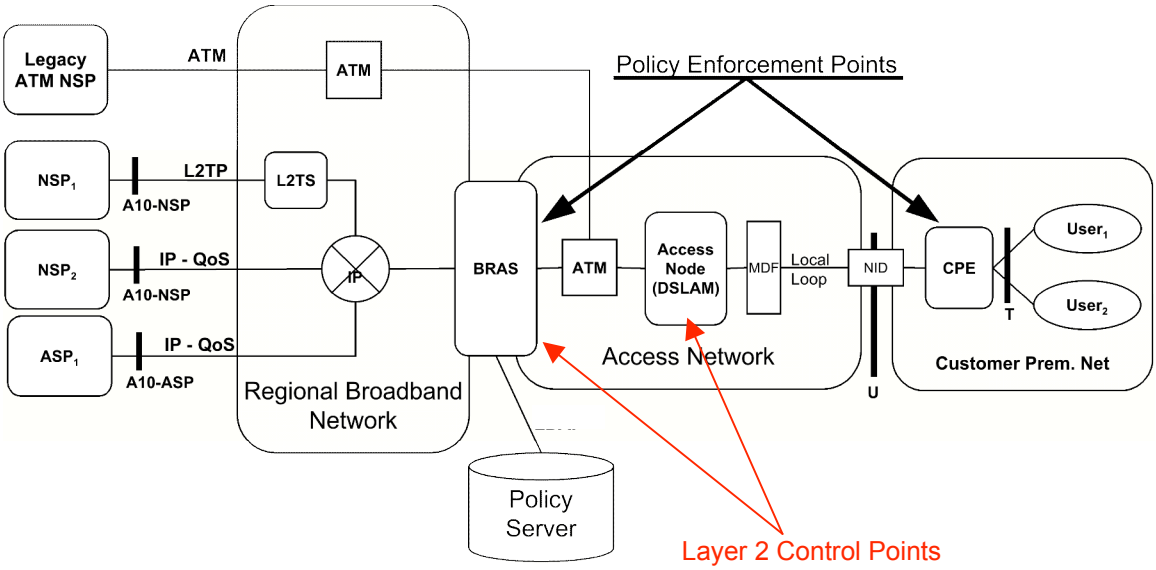


Figure 1 Network Topology [2]

As described in the introduction, this will allow multiple types of bi-directional control interactions:

- Dynamically discovering the access topology (actually the subset of it relevant for the BRAS hierarchical scheduling and policing)
- Dynamically communicating service parameters of relevance for the Access Node (in relation with a given subscriber line)
- Dynamically managing layer 2 multicast entries in real time when processing IGMP requests
- Statistics and end-to-end OAM exchanges

2.2 General Guidelines

The general guidelines that led to this approach of a an in-band control plane are the following:

- Stay as close as possible to the model defined by TR-58 and 59 where service & subscriber management is distributed between the routing gateway (CPE) and the BRAS, and decoupled from the access network settings.
- Keep IP and AAA processing in the BRAS
  - This includes companion protocols like PPP, DHCP, IGMP, PIM, RADIUS, COPS, etc
- Keep the Access Nodes focused on layer 2 processing (e.g. ATM and Ethernet), try to avoid mechanisms implying hardware upgrades
- The approach must allow to significantly optimize the overall OPEX and CAPEX of the access network
- The approach must accommodate various cases of administrative scope of responsibility, notably situations where the Access Node is **not** operated by the same department (or even company) as the BRAS

- The approach must be general-purpose and accommodate multiple types of access technology (e.g. ATM, Ethernet, VLANs, etc) and also foresee other types of local loops and Access Nodes (e.g. PON and Optical Line Terminations).
- The approach must foresee a non-disruptive step by step transition between the existing paradigm (e.g. 1 VC per subscriber, ATM-based access network) and a more general paradigm.
- The approach must be flexible enough to address multiple distinct use cases (e.g. multicast, topology discovery, line configuration, OAM, etc) and to open the door to future extensions.

For multicasting some additional guidelines have to be considered:

- Leverage on the best performing multicast technique natively supported by the Access Node
- Multicast "zapping" must be fast enough to not impact typical behaviors of TV viewers (e.g. response time of a couple of hundreds of milliseconds)
- Multicast traffic for premium channels or commonly used channels (which are typically not very numerous) may require traffic optimization (e.g. replication at the DSLAM level) while multicast traffic for less commonly used (but possibly very numerous) channels might not.

2.3 Communication Framework

The aim is to define a general-purpose solution for multiple network scenarios with an extensible communication scheme, addressing the applications described in the introduction as well as future applications.

These control plane interactions are transactional in nature, and imply a reliable communication channel to share states. Bidirectional operations are needed, as well as dynamic negotiation of capabilities to address transition issues.

A high-level communication framework for such control plane can therefore be defined, as described by the following figure:

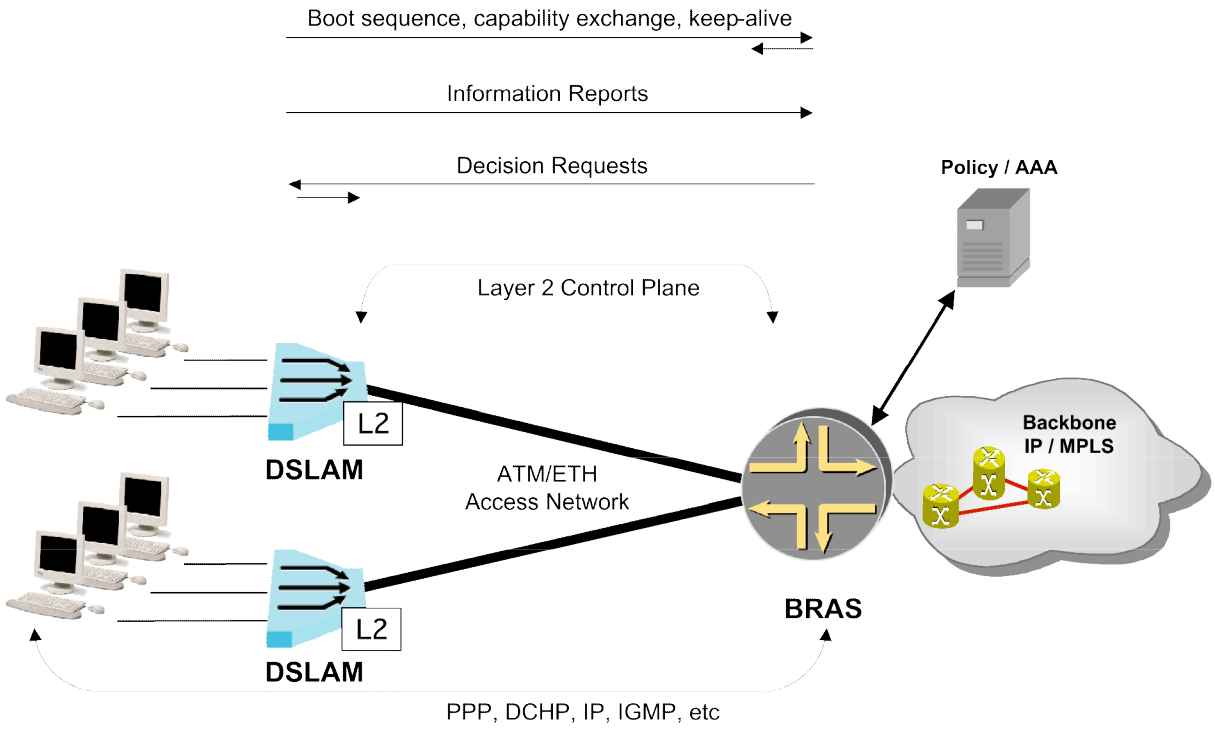


Figure 1: Layer 2 Control Plane

Since the connectivity between DSLAM and BRAS differs, identification of unicast & multicast flows/channels will also differ. So the communication will have a common framework and may have technology-specific variations.

Finally, such a protocol should be defined in a way to simplify specification, implementation, debugging & troubleshooting, but also has to be easily extensible in order to support other types of in-band remote control and update operations between Access Nodes and BRAS

The messages in this document are described in an abstract way, independent from any actual protocol mapping which is discussed in a separate contribution dslf.2004.152

3 Use Cases for Layer 2 Control Mechanism

3.1 Topology Discovery and Line Configuration

3.1.1 Overview and Motivation

The following figure comes from TR-059 and describes a typical access network topology and the information that is necessary to be discovered. The topology and the rates of the various links need to be accurately known by the BRAS to enable its hierarchical scheduling and policing mechanisms.

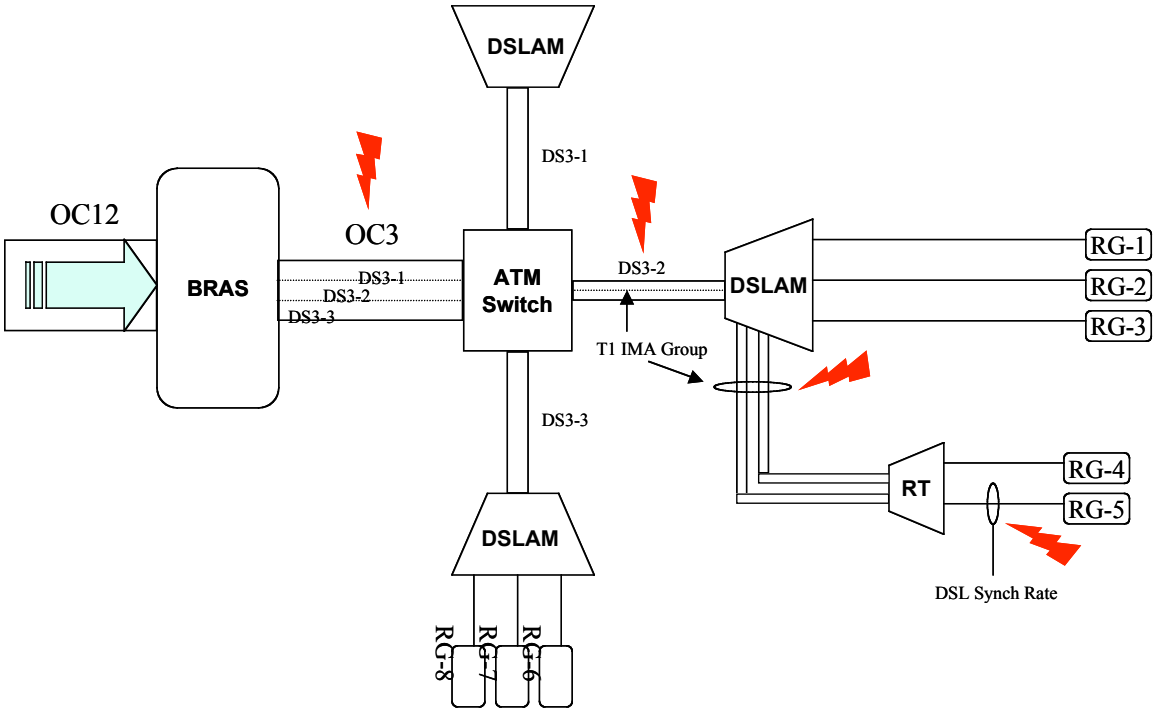


Figure 2: Access Network Topology

Key data to be discovered are the following:

- the identification and speed (rate) of the DSL local loop (sync rate)
- the identification and speed (rate) of the RT/DSLAM link (when relevant)
- the identification and speed (rate) of the DSLAM/ATM-switch link

Note that the ATM-switch/BRAS link is obviously known by the BRAS since this physical link is terminated on the BRAS. Any links between a hierarchy of ATM switches would typically not be oversubscribed, hence are not relevant for the BRAS. A similar picture could be drawn for an Ethernet access topology, replacing the ATM switch by an Ethernet switch. A key difference though is that the traffic issued or

received by the BRAS may not be encapsulated in a two-level hierarchy of layer2 “logical” paths & circuits (assuming stacked VLANs are not used for such purpose).

The following figure comes from TR-059 and describes the BRAS 5-level queuing scheduler:

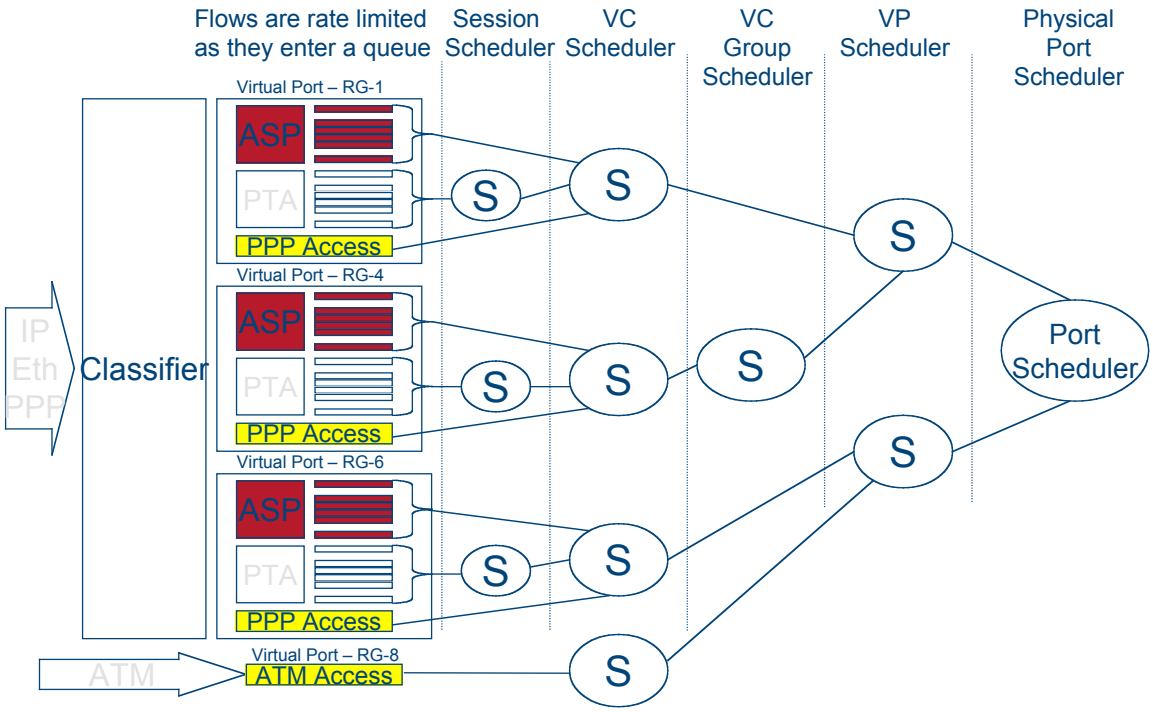


Figure 3: Hierarchical Scheduler

Topology discovery function will allow the BRAS to perform advanced functions like hierarchical scheduling and policing without having to depend on an error-prone & possibly complex integration with an OSS system. Such a function will also allow the BRAS to communicate back to the DSLAM service & subscriber-related line configuration parameters, alleviating the need for OSS integration with the DSLAM for subscriber-related data.

In case the net data rate on the DSL line or any other link in the access network is modified, it is necessary to control and check all elements along the data path (BRAS and DSLAM), so that the desired data rate is in line with the available data rate. The latter is usually limited by noise conditions on the DSL line.

Topology discovery is specifically important in case the net data rate the DSL line changes overtime. More background information about the circumstances of such rate change can be found in Appendix A.

Topology discovery may actually include more information than link identification and corresponding data rates, notably for the local loop. A more complete list of such DSL parameters can be found in Appendix B.

3.1.2 Control Interactions

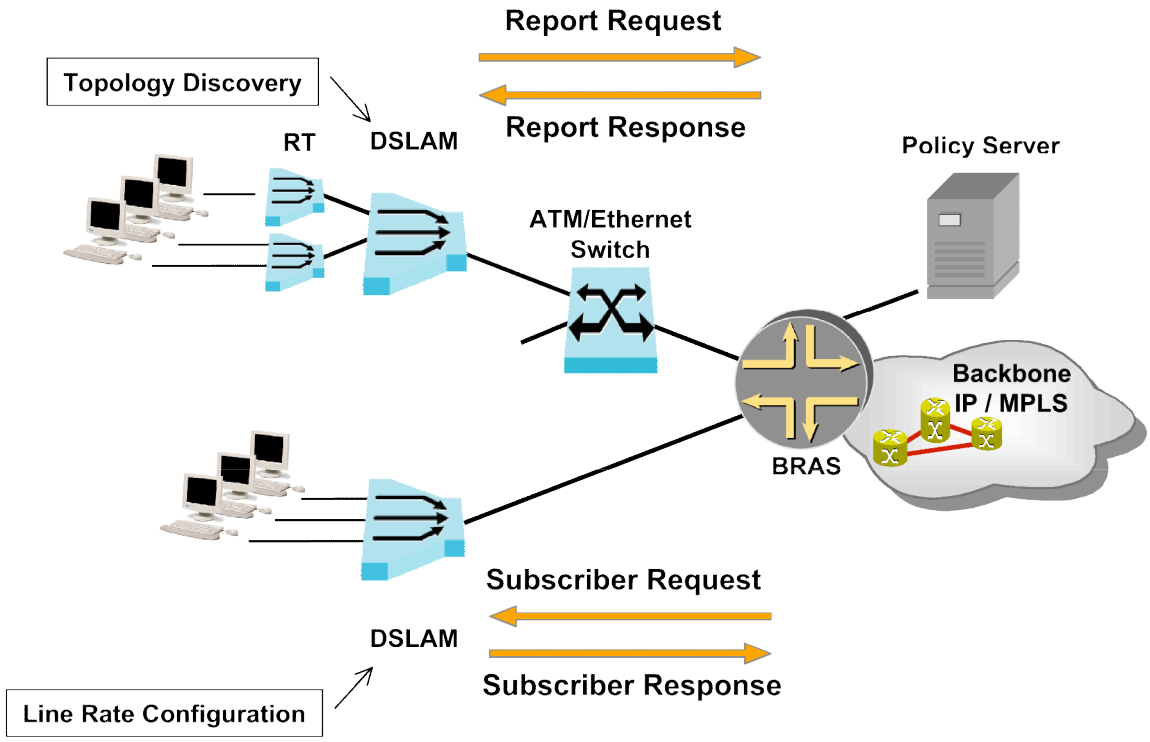


Figure 4: Topology Discovery and Line Configuration

UC-1: Topology Discovery

In networks that use a BRAS performing advanced functions like hierarchical queuing (as required by TR-059), the BRAS needs to have an accurate view of the access network topology. To allow the BRAS to dynamically adapt his schedulers, the DSLAM informs the BRAS with a Report Message when a new link is discovered or when the data rate on one of its links changes (e.g. the data rate on the DSL line, i.e. sync rate). This allows the BRAS to rebuild the hierarchy of links in the access network and automatically (re-)configure its hierarchical QoS packet scheduler. A policy server can also use such information for admission control purposes.

- independent on layer 2 at U-interface
- independent on layer 2 at V-interface
- DSLAM informs BRAS about line rate for specific ports of

    - the DSL local loop (sync rate)
    - the RT/DSLAM link (when relevant)
    - the DSLAM/ATM-switch uplink
- BRAS can adjust downstream shaping to current DSL line rate, and more generally (re-configure) all nodes of its hierarchical scheduler (support of advanced capabilities according. to TR-059)

It is expected that Remote Terminals may not implement such a layer 2 control mechanism, but that “master” DSLAMs may report RT DSL local loops and RT/DSLAM links on their behalf. See appendix C for more details on distributed architectures involving Remote Terminals.

UC-2: Line Configuration

DSL line rates are typically configured in a static way. If a subscriber wants to change its line rate, this requires an OPEX intensive reconfiguration of the line configuration via the network operator, possibly implying a business-to-business transaction between an ISP and an Access Provider. The proposed layer 2 control mechanism would support a more flexible approach for supporting “bandwidth on demand”.

More generally, several service/subscriber DSL parameters (e.g. rate, inter-leaving delay) could benefit from such flexible approach to enable a “service on demand” model. Multicast-related parameters (e.g. user entitlement) may also be relevant, notably in the context of use case UC-11.

The best way to change line parameters would be by using profiles. These profiles (DSL profiles for different services) are pre-configured by the EMS managing the DSLAMs. The L2-control interaction then only needs to transmit a reference to the right DSL profile. In the future, we may also consider extending this scheme to convey discrete DSL parameters in the L2-control interaction.

Triggered by topology information reporting a new DSL line, the BRAS may send line configuration information (e.g. reference to a DSL profile) to the DSLAM using Subscriber Request Messages. The BRAS may get such line configuration data from a policy server (e.g. RADIUS). The BRAS may update the line configuration due to a subscriber service change (e.g. triggered by the policy server).

- independent on layer 2 at U-interface
- independent on layer 2 at V-interface
- BRAS is informed by a policy server, e.g. via COPS about requested bandwidth
- BRAS instructs DSLAM to configure line rate for a specific port
- DSLAM configures line rate and informs BRAS about new line rate

3.2 Layer 2 OAM Use Cases

3.2.1 Overview and Motivation

As long as there is ATM end-to-end between BRAS and CPE, OAM loopback cells can be used for on-demand connectivity monitoring, fault localization and pre-service connectivity verification.
The subscriber is identified by the PVC/PVP assignment. Therefore the loopback endpoint can be addressed by one default loop back location ID which is equal to all endpoints. That simplifies the operation because no administration of different subscriber specific Loop back ID is needed.
Once Ethernet technology is used between the BRAS and the DSLAM this end-to-end ATM OAM test can't be used, and Ethernet OAM (as being specified by ITU and IEEE) will be limited in scope to the Ethernet segments.
The motivation is to close this gap by:

- restoring a fault localization mechanism between the BRAS and the DSLAM
- avoiding any change to the ATM-based U-interface (local loop)

One possible solution is to have the desired test capability between BRAS and CPE provided by a new type of operation performed via the layer 2 control plane.[1]

Another aspect of subscriber management is to gather proper statistics about the services being delivered and made them available to proper OSS systems (e.g. performance management tools or accounting tools).

Moving most of the actual multicast replication creates in turn a challenge for the BRAS to issue proper statistics.

Consequently, it might be desirable to have a "reporting" feedback mechanism available by which the DSLAM indicates traffic statistics on a per subscriber line basis. Beside multicast, unicast traffic statistics may also be required in the future. Then the BRAS can aggregate such statistics with other subscriber-related statistics and make them available to proper OSS systems (e.g. via Radius accounting, SNMP, bulk statistics, etc).

[1] It is expected that WT-101 will address OAM topics at large when Ethernet is used as an access network. The specific topic of fault localization will be studied in this context. It is possible that other solutions than using L2 control operations for this purpose will emerge.

3.2.2 Control Interactions

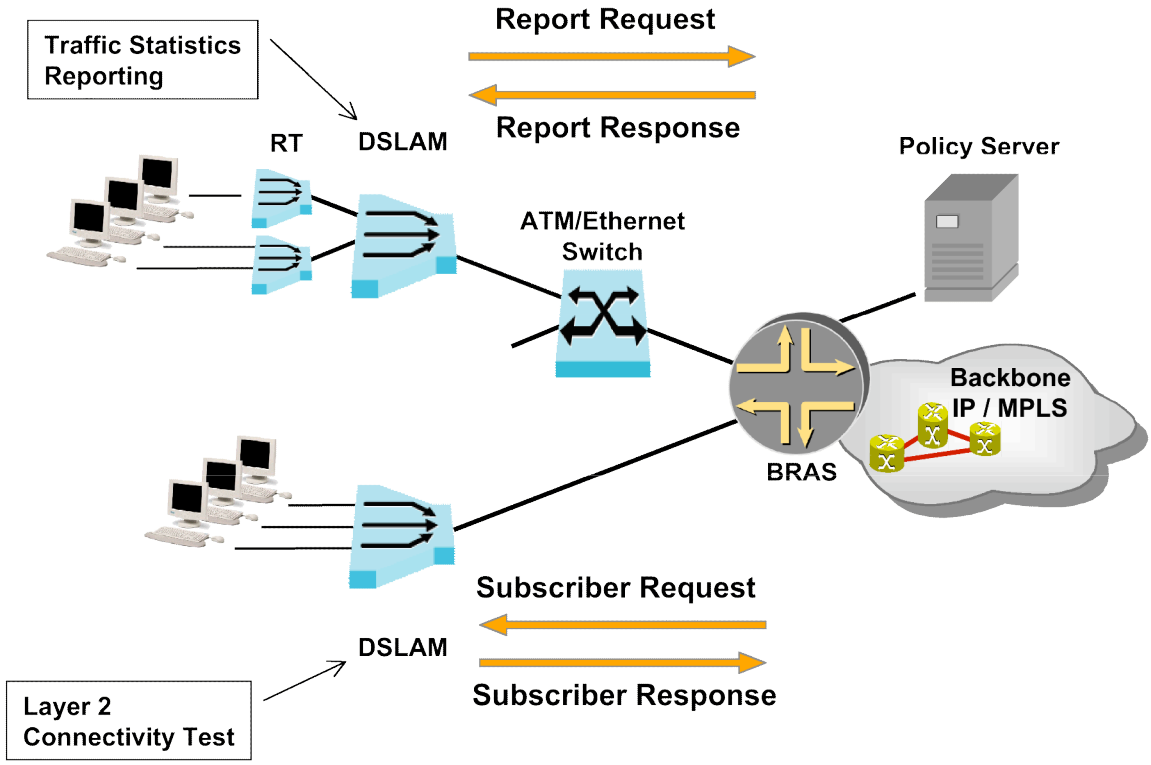


Figure 5: Layer 2 OAM

UC-3: Layer 2 Connectivity Test

Triggered by a local management interface the BRAS can initiate a loop test between DSLAM and DSL modem via a layer 2 control operation. A Subscriber Message is sent to the DSLAM, which triggers an ATM F4/F5 loopback test (or an Ethernet loopback test according to EFM) on the DSL line (local loop). A Subscriber Response is sent by the DSLAM to the BRAS, which may report results via a local management interface.
Thus, the connectivity between BRAS and DSL modem (CPE) can be monitored by a single trigger event.

- Independent on Layer 2 at U-Interface
- Ethernet on Layer 2 at V-Interface
- BRAS can initiate ATM (or Ethernet) local loop test
- DSLAM sends ATM F4/F5 OAM loopback cell in case of ATM over DSL or DSLAM sends EFM loopback frame in case of Ethernet over DSL
- DSLAM informs BRAS about test result

UC-4: Traffic Statistics Reporting

The DSLAM can use Report Messages to send traffic statistics, which may be on per subscriber basis, to the BRAS. The BRAS can use local packet statistics to generate aggregated per-subscriber statistics or accounting data may then aggregate such statistics.

3.3 Multicasting Use Cases

3.3.1 Overview and Motivation

The intention is to avoid replicating multicast traffic going from BRAS to the same Access Node, at least for commonly used channels.
A first approach is for the DSLAM to not acquire any BRAS functionality, i.e. IGMP authorization/filtering and routing decisions remain completely with the BRAS. All subscriber management mechanisms (Authentication, Authorization, Accounting and Address management) also stay with the BRAS, including multicast policies such as per-channel authorization. The complex multicast routing protocols (e.g. PIM, DVMRP) stay with the BRAS, too. There are also no changes to regular unicast traffic (based on PPP or bridged-2684) as well as to upstream multicast traffic, including IGMP itself.

A concept for non-premium and premium channels is proposed where the latter one is distinguished by a guaranteed quality of service. In case of availability of the proposed extension the BRAS would replicate multicast streams for premium channels on a per-DSLAM basis, i.e. actually, the DSLAM performs the multicast replication by layer 2 means (e.g. ATM point-to-multipoint cell replication, or Ethernet data-link bridging). All non-premium channel packet replication remains with the BRAS. That way, pay-per-view and premium channels models are facilitated.

In contrast to this first approach, ITU-T recommendation H.610 recommends a second approach, which requires multicast traffic replication and IGMP processing implemented in the DSLAM and may load the DSLAM with the burden of subscriber authentication, authorization and dynamic determination of multicast subscription rights. In addition, the CPE/modem may have to filter the user's upstream traffic for IGMP messages and to forward them into a special "zapping" PVC. For this second approach, a layer 2 control plane can be added to offload the DSLAM from subscriber authentication, authorization and dynamic determination of multicast subscription rights and improve the ITU H.610 scheme.
When the layer 2 control model is used in combination with DSLAMs supporting the H.610 model, inband control messages will be used, not on a per IGMP message basis, but more as a line configuration mechanism, allowing the BRAS to push subscriber-related authorization data to the DSLAM. This approach would protect the investment already made by some operators towards the H.610 architecture.

In order to give an overview the following table shows relevant use cases depending on the switching technology used at the DSLAM and the technology the multicast replication takes place.

DSLAM		Technology at V-interface	Use Case #
Switching/ Bridging	**MC replication**		
ATM	ATM cell replication based on VPI/VCI information	ATM dedicated PVCs per subscriber	UC-5
		Ethernet dedicated VLANs per subscriber	UC-6
		Ethernet no dedicated VLAN per subscriber	UC-7
Ethernet	Ethernet frame replication; MAC addresses used as MC stream identifier	Pure Ethernet (no VLAN at all)	UC-8
		VLAN-tagged Ethernet no dedicated VLANs per subscriber	UC-8
		Ethernet over ATM (RFC 2684) dedicated PVC per subscriber	UC-9
		Ethernet over ATM (RFC 2684) no dedicated PVC per subscriber	UC-10

Table 2: Layer 2 Multicast Use Cases (first approach)

In the case of the H.610 approach, the replication is based on ATM technology and dedicated ATM PVCs (hence similar to use case UC-5 from a data plane standpoint), but the dynamics of the layer 2 control exchanges are quite different, and will be explained in use case UC-11.

3.3.2 End-To-End Data Plane, CPE Considerations

In many deployments, PPP is used from an IP host (e.g. PC, Set-Top Box) connected to the home network or from a CPE/RG (Routing Gateway) at the residence. These PPP sessions will run between Host or RG and network BRAS, with functions such as authentication, authorization, accounting and address assignment performed by PPP processing in the BRAS, usually combined with RADIUS back-end server(s). In the case of multicast, IGMP would run between PPP endpoints (Host/RG and BRAS) and would control group joining and leaving.

When PPP sessions run between Host and BRAS, traffic between those endpoints is carried in the PPP session. When the PPPoE session runs between the RG and BRAS, connectivity from multiple host devices may be sustained by the PPP session between RG and BRAS, with the CPE/RG performing a bridging or routing function

(often combined with network address translation) for traffic received on the PPP session.

Multicast solutions such as H.610 that have IGMP terminating and replication occurring in a DSLAM assume that IGMP is not PPP-encapsulated, and cannot provide a DLSAM based multicast replication solution for deployments that run PPP between Host/RG and BRAS.

A solution using layer2 control for distributed multicast will support replication occurring in the DLSAM but with IGMP termination and multicast control being performed in the BRAS while PPP is being used. More details about such approach will be provided in a separate contribution.

3.3.3 Data Plane: Multicast with ATM-Based DSLAMs

Use cases UC-5, UC-6, and UC-7 have in common, that the internal forwarding, switching and multicast decisions are based on ATM VPI/VCI information.

Between the customer premises equipment (CPE) and the DSLAM there is one ATM PVC per customer for normal unicast traffic (upstream/downstream), plus upstream multicast traffic including IGMP. For downstream multicast traffic one ATM PVC per active multicast channel is used.

The number of circuits depends on the actual DSL capacity, and also on the subscriber service contract.

- Multiple PVCs at U-Interface
- ATM Switching in DSLAM
- ATM Multicast Cell-based Replication in DSLAM
- Real-time remote zapping

Based on IGMP joins and leaves, the BRAS communicates to the DSLAM what channels are to be replicated for a particular subscriber. Thus, multicast replication takes place closer to the subscriber. In case of rejecting the request, the BRAS could decide to fall-back to a normal user channel for downstream multicast.

Following use cases are sub-classified depending on their connectivity between DSLAM and BRAS (V-interface).

UC-5: ATM at V-interface with dedicated PVCs per subscriber

Between the DSLAM and the BRAS there is one PVC per connection containing one pre-configured multicast channel per premium multicast group. Such a premium multicast channel is dynamically cross-connected with user-facing multicast PVCs.

At the V-interface there are:

- 1 ATM PVC per subscriber for Internet traffic
- 1 ATM PVC per multicast stream
- 1 ATM PVC for layer 2 control channel per DSLAM

UC-6: Ethernet at V-interface with dedicated VLANs per subscriber

Between the DSLAM and the BRAS there is one VLAN per connection containing one pre-configured multicast channel per premium multicast group. Such a premium multicast channel is dynamically cross-connected with user-facing multicast PVCs.

At the V-interface there are:
- 1 VLAN per subscriber for Internet traffic
- 1 VLAN per multicast stream
- 1 VLAN for layer 2 control channel per DSLAM

Beside the ATM multicast replication the DSLAM has to support a VLAN/PVC mapping function.

UC-7: Ethernet at V-interface without dedicated VLANs per subscriber

At the V-interface there are:
- 1 or more VLANs per service or DSLAM for all subscribers' Internet traffic (Unicasts)
- 1 or more VLANs for multicast streams
- 1 VLAN for layer 2 control channel per DSLAM
- Subscriber identification mechanism such as proposed in contribution dslf2004.071.

Beside the ATM multicast replication the DSLAM has to support a VLAN/PVC mapping function.

3.3.4 Data Plane: Multicast with Ethernet-Based DSLAMs

Use cases UC-8, UC-9, UC-10 have in common, that the internal forwarding, switching and multicast decision are based on Ethernet MAC address information (the DSLAM acting as a data-link bridge). This does not exclude configuration with Ethernet encapsulated in ATM on the V-interface.

Between the customer premises equipment (CPE) and the DSLAM there is one ATM PVC per customer for all upstream and downstream traffic including IGMP. Instead of ATM there can also be Ethernet over DSL at the U-interface.
Premium multicast channel are dynamically cross-connected towards the user-facing PVC. Based on IGMP joins and leaves the BRAS communicates to the DSLAM what channels are to be replicated for a particular subscriber on base of their multicast MAC addresses.

- 0 or 1 PVC at U-Interface (0 in case of Ethernet over DSL)
- Ethernet Switching in DSLAM (incl. MAC Address - Filtering)
- Ethernet Multicast in DSLAM
- MAC addresses used as multicast stream identifier
- Subscriber identification mechanism such as proposed in contribution dslf2004.071.

Note: If VLAN-tagging is used at the V-interface with dedicated VLANs per subscriber then the VLAN ID can be used as multicast stream identifier. The DSLAM performs VLAN switching instead of data-link bridging.

Following use cases are sub-classified depending on their connectivity between DSLAM and BRAS (V-interface).

UC-8: Ethernet at V-interface

For the Ethernet frame replication at the DSLAM it is not significant whether the Ethernet at the V-interface is VLAN tagged or pure Ethernet is used. The different traffic is always identified based on MAC addresses.
If VLANs are used then they are not dedicated per subscriber but used for traffic segregation.

- 1 VLAN for all subscriber Internet traffic
- 1 (or N) VLANs for multicast traffic
- 1 VLAN for layer 2 control channel per DSLAM
- or maybe 1 single VLAN per DSLAM, or any combination

UC-9: Ethernet over ATM (RFC 2684) at V-interface, dedicated PVC per subscriber

Ethernet encapsulation in the next two sub cases allows for enabling multicast services by upgrading to Ethernet switching without changing the existing ATM access network. This requires the use of Ethernet data-link bridging within the DSLAM.

At the V-interface there are:
- 1 ATM PVC for each subscriber Internet traffic
- 1 (or N) ATM PVCs for multicast streams
- 1 ATM PVC for layer 2 control channel per DSLAM
- optional ATM switching for specific ports

UC-10: Ethernet over ATM (RFC 2684) at V-interface, no dedicated PVC per subscriber

At the V-interface there are:

- 1 ATM PVC for all subscribers Internet traffic
- 1 (or N) ATM PVCs for multicast streams
- 1 ATM PVC for layer 2 control channel per DSLAM

3.3.5 Data Plane: Multicast with IGMP Proxy running in DSLAM

In this case the DSLAM will process IGMP joins and leaves, build multicast replication tables and perform the packet or cell replication function. Subscriber multicast entitlement information will be sent from BRAS to DSLAM at subscriber session establishment or when a change is made to subscriber entitlement.

3.3.6 Control Plane Interactions

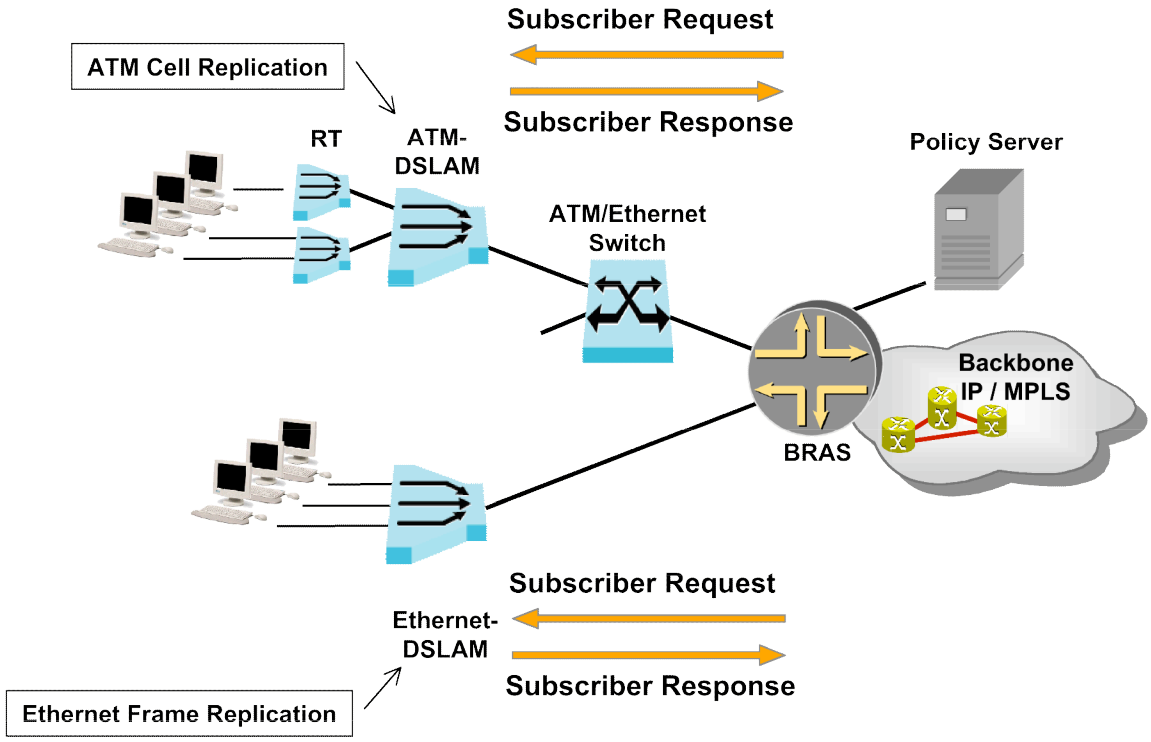


Figure 6: Multicast with ATM/Ethernet-Based DSLAMs

UC-5, UC-6, UC-7: Multicast with ATM-Based DSLAMs

On reception of the IGMP request at the BRAS, the BRAS may verify subscriber multicast and/or group entitlement. Successful group membership requests may result in a Subscriber Response being sent on a control channel to the DSLAM, with Subscriber VC, and Multicast channel information. An ATM cell replication will occur on the DLSAM between appropriate multicast channel and subscriber line.

UC-8, UC-9, UC-10: Multicast with Ethernet-Based DSLAMs

On reception of the IGMP request at the BRAS, the BRAS may verify subscriber multicast and/or group entitlement. The BRAS may use some of the mechanisms outlined in dslf2004.071to identify the DSLAM and DSL line which the Subscriber is associated with. Successful group membership requests may result in a Subscriber Response being sent to the DSLAM that identify the subscriber and multicast group. Replication will occur at the DSLAM for appropriate multicast channel and subscriber line.

UC-11: Multicast with IGMP-Proxy in the DSLAM

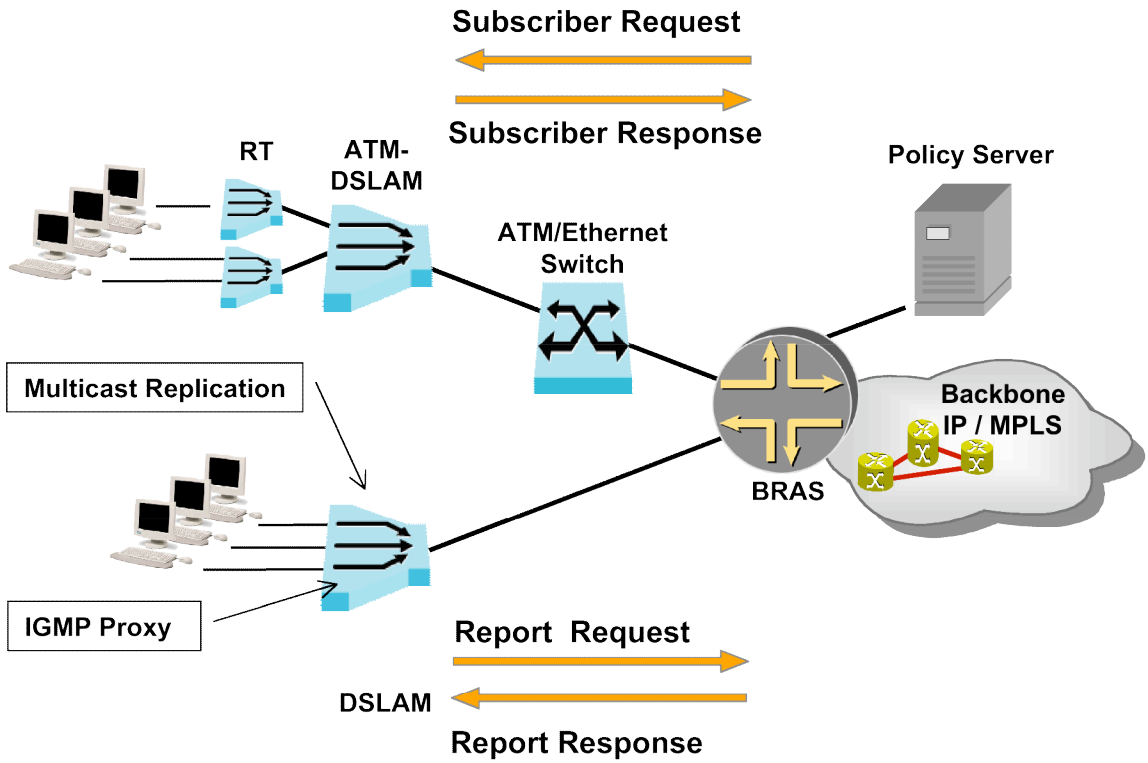


Figure 7: Multicast with IGMP-Proxy in the DSLAM

In order to offload the DSLAM from subscriber authentication, authorization and dynamic determination of multicast subscription rights, the layer 2 control mechanism is used similar to the Line Configuration use case. This allows the BRAS to push subscriber-related authorization data to the DSLAM using Subscriber Messages, e.g. a list of multicast channels allowed for this subscriber. Dynamic updates of such data may also be pushed when needed.

When replication memberships change at the DSLAM for a given subscriber, the available per subscriber bandwidth for unicast traffic changes. The DSLAM will notify the BRAS of this change using a per subscriber Report message. The BRAS can use such information to adjust TR-59 compliant hierarchical scheduler nodes.

Note: this use case is illustrated with an IGMP-Proxy function in the DSLAM, but would be equally applicable if IGMP is terminated by the DSLAM (then acting as an IGMP router), or if transparent IGMP snooping is implemented in the DSLAM.

4 Message Flows for Layer 2 Control Mechanism

4.1 Message Descriptions

The following message types may exist to facilitate the exchange of this control information

1) A Boot Request message sent by BRAS or Access Node
2) A Boot Response message
3) A Subscriber Request message sent by the BRAS to Access Node
4) A Subscriber Response message that is sent in response by the Access Node
5) A Report message to support an asynchronous exchange of control data
6) A Report response message
7) A Heartbeat message to be exchanged between Access Node and BRAS

4.1.1 Boot Request

The Boot Request is sent in a directed or broadcast manner by BRAS or Access Node, typically at start-up time. It is intended to solicit capability information from an Access Node for a BRAS or from a BRAS for an Access Node. The Boot Request message may contain the following parameters

- <Sequence Number>
- <Holding Time>
- <Device Type>
- <Topology Capable >
- <Statistics Capable>
- <Multicast Capable>
- <“Other capability” Capable>
- <Message Authentication>

The receiver, to foresee future extensions, must ignore unknown capabilities.

4.1.2 Boot Response

The Boot Response message may contain the following parameters

- <Sequence Number>
- <Holding Time>
- <Device Type>
- <Topology Capable (Y/N) >
- <Statistics Capable (Y/N)>
- <Multicast Capable (Y/N)>
- <“Other Capability” Capable (Y/N>
- <Message Authentication>

4.1.3 Subscriber Request Message

The Subscriber Request messages allows the BRAS to configure subscriber level information in the DSLAM, query data from or initiate an action in the DLSAM.
Examples of when a Subscriber Request message may be sent by the BRAS include

- When a user is first identified and authenticated
- In the case of multicast when an IGMP join or leave is received.
- When a subscriber service level changes
- When OAM tests need to be performed

The operation is intended to be transactional in nature (performed in an atomic way; in case of failure, no change occurred on the DSLAM side).

The following parameters may be contained in this message type:

- <The Subscriber ID with which DSLAM can identify the subscriber (e.g.: MAC Address, ATM VC associated with a subscriber, DSL Line ID)>
- <Line Configuration Parameters | Optional>
- <OAM Test Parameters | Optional>
- <Multicast group address/Channel identifier | Optional>
- <Add/Delete Multicast Entry | Optional>
- <Message Authentication | Optional>

4.1.4 Subscriber Response Message

An access node sends the Subscriber Response message to a BRAS in response to a Subscriber Request message destined to it.
The parameters of this message may include:

- <Subscriber ID>
- <Sequence No>
- <Multicast group address/Channel identifier (S/F/Error Code)| Optional>
- <Add/Delete Multicast Entry (S/F/Error Code)| Optional>
- <DSL Line ID | Optional>
- <OAM Test Results | Optional>
- <Message Authentication | Optional>

While awaiting a response of Subscriber Request message, the BRAS may retransmit the SR messages every SR_Retrans_Timer seconds. If no Subscriber Response is received after MAX_PSR_COUNT messages, the BRAS should consider that the operation has failed.

4.1.5 Report Message

The Report Message will typically be sent from Access Node to BRAS used to transmit Data such as Topology information or Statistics to the BRAS and may contain some of the following parameters:

- <Sequence No>
- <Subscriber ID | Optional>
- <Statistics | Optional>
- <Topology Data | Optional>
  - <DSL Line ID & Bandwidth | Optional>
  - <RT/DSLAM Link ID & Bandwidth | Optional>
  - <DSLAM Uplink ID & Bandwidth | Optional>
- <Message Authentication | Optional>

The bandwidth description of a given link provides both the upstream and downstream capacity of such link. In the case of the DSL Line, this is actually the sync-up rate of the DSL line. Any change to the sync rate (or any change to a link capacity) should trigger a new report message.

4.1.6 Report Message Response

The Report message Acknowledgement may use the following parameters

- <Subscriber ID| Optional>
- <Sequence No>
- <Message Authentication | Optional>

While awaiting acknowledgement of Report message, the Access node may retransmit the Report messages every Report_Retrans_Timer seconds. If no Report Response is received after MAX_R_COUNT messages, the Access Node should consider that the operation has failed.

4.1.7 Heartbeat

Heartbeat messages can optionally be sent between Access Node and BRAS and can contain the following parameters.

- <Sequence Number>
- <Holding Time>
- <Authentication | Optional>

4.2 Message Flows

4.2.1 Capabilities Exchange during Boot Sequence

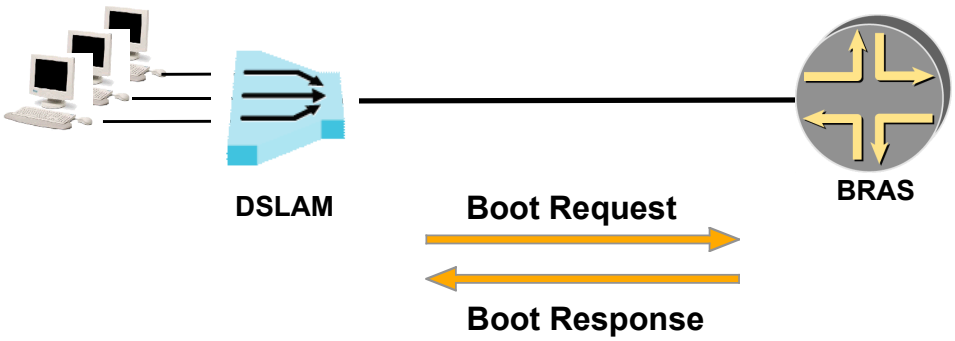


Figure 8: Boot Exchange Messages

BRAS and DSLAM exchange capabilities information via boot exchange messages for such functions as multicast capability, ability to accept topology information, use Statistics, etc. The capability exchange is expected to occur at startup and could also serve as a device discovery mechanism for both the Access Node and BRAS.

The boot data is valid for the duration of the Holding Time. Sending a null Holding Time allows performing an orderly shutdown.

→ BRAS or DSLAM Boot Request (Sequence Number, Holding Time, Device Type, Topology Capable, Statistics Capable, Multicast Capable, “Other capability” Capable, Message Authentication)
← BRAS or DSLAM sends Boot Response ()

4.2.2 Topology Discovery

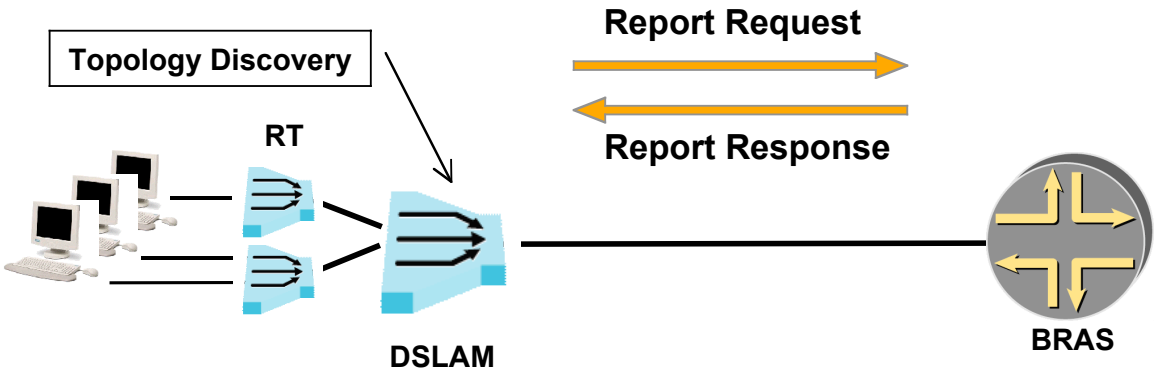


Figure 9: Topology Discovery

The DLSAM will send topology information (Link, Device, Bandwidth information) to the BRAS using Report messages. The Report message can also be used by the DLSAM to notify the BRAS of changes in the DSL line Sync Rate information.

The BRAS can use this data to parameterize its TR-59 Hierarchical Scheduler, or general access configuration elements (e.g. stack of interfaces). A policy server can also use such data for admission control purposes.

→ DSLAM sends Report Request (Topology Data, DSL Line ID & Bandwidth, RT/DSLAM Link ID & Bandwidth, DSLAM Uplink ID & Bandwidth, Message Authentication
← BRAS sends Report Response

4.2.3 Line Configuration

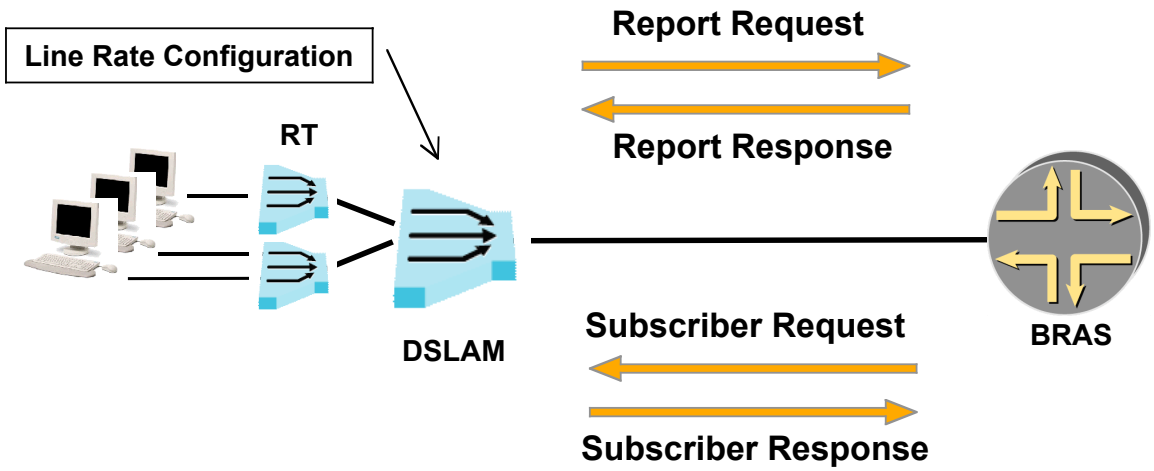


Figure 10: Line Configuration

The following message flows shows an upload of topology information from BRAS to DSLAM and the subsequent update of line configuration in the DLSAM by the BRAS.

→ DSLAM sends Report Request (DSL Line ID & Bandwidth, Message Authentication)
← BRAS sends Report Response
← BRAS sends Subscriber Request (Subscriber-ID, DSL Line ID & Bandwidth, Message Authentication)
→ DSLAM sends Subscriber Response()

4.2.4 Traffic Statistics Reporting

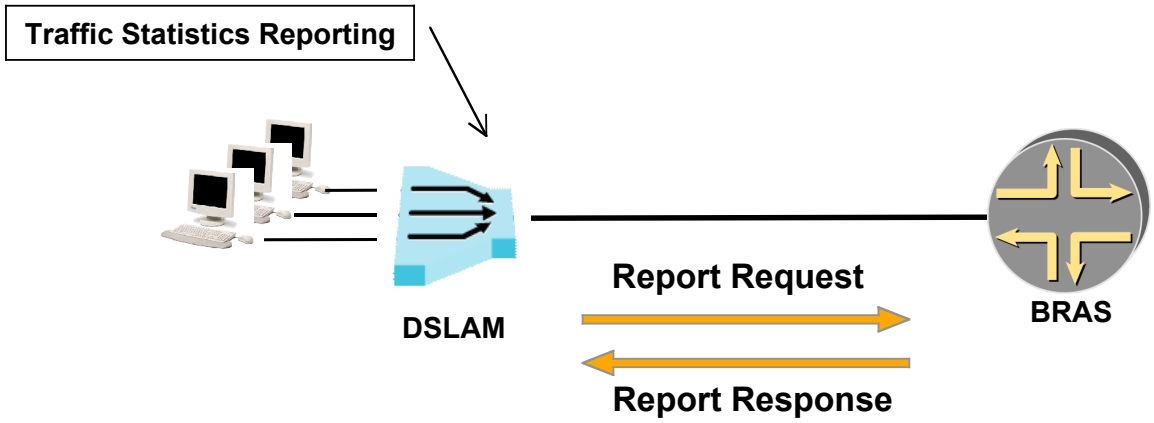


Figure 11: Traffic Statistics Reporting

The following message flow shows the upload of per subscriber traffic statistics from DSLAM to BRAS.

→ DSLAM sends Report Request (Subscriber-ID, Statistics Data)
← BRAS sends Report Response

4.2.5 OAM Connectivity Test

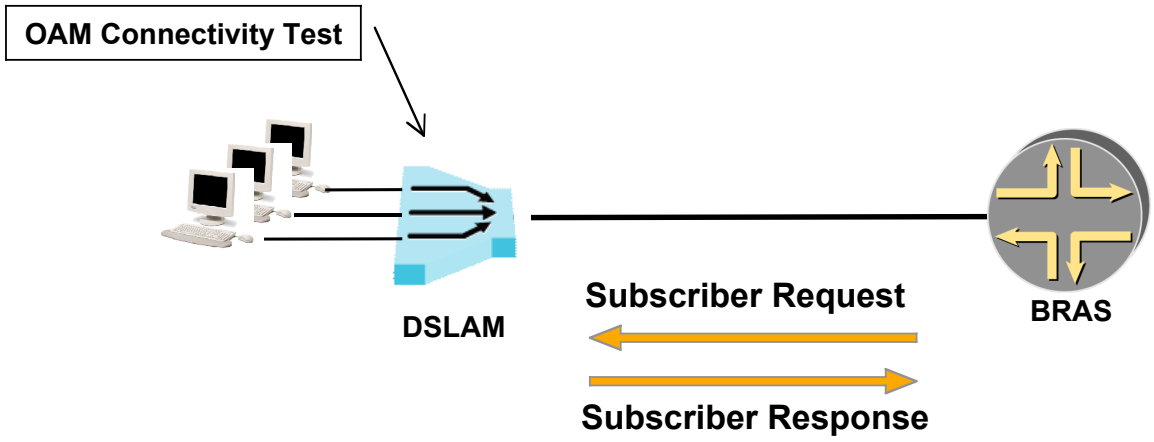


Figure 12: OAM Connectivity Test

The following message flow sequence shows the BRAS initiating a connectivity test in the DSL line, with the DSLAM reporting the test results to the BRAS.

← BRAS sends Subscriber Request (Subscriber-ID,OAM-Request, Message Authentication)
→ DSLAM sends Subscriber Response()

4.2.6 Multicast with ATM based replication

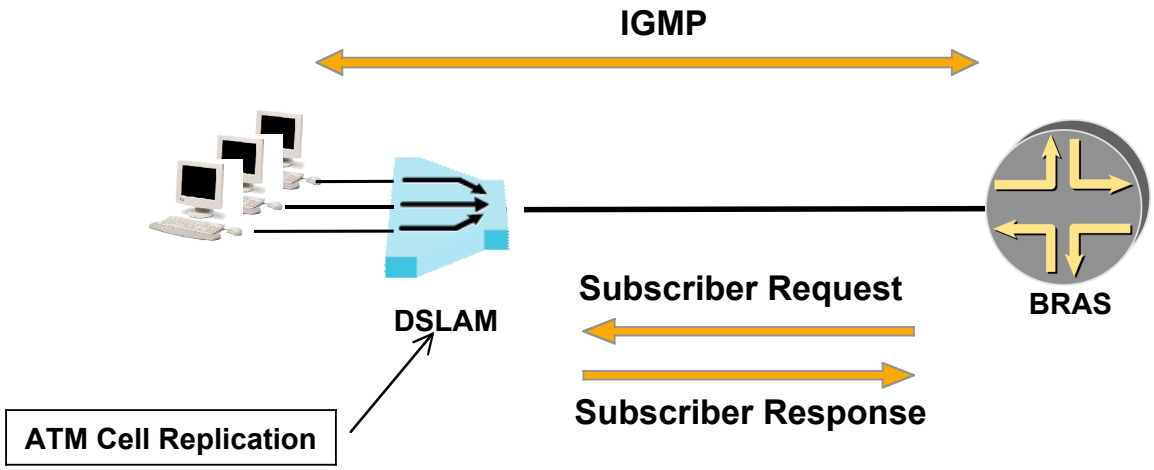


Figure 13: ATM Multicast

The following message flow description outlines the expected sequence of IGMP messages and Layer 2 control messages, to allow an ATM-based DSLAM multicast replication with IGMP and AAA functions residing the BRAS.

→ IGMP Membership Report message send from Host to BRAS
BRAS analyzes permissions of this subscriber locally (RADIUS or COPS) and
identifies appropriate Control VC ($VPI_{UNICAST}$ = $VPI_{CONTROL}$)
identifies appropriate Multicast PVC by analyzing it's multicast routing table
← BRAS sends Subscriber Request(Subscriber-ID, Multicast Group.. .)
→ DSLAM sets up cross connect for multicast according to
user port is evaluated by using the $VCI_{UNICAST-UNI}$ received from BRAS
any free PVC with $VCI_{MULTICAST-UNI}$ is added as leaf to $VCI_{MULTICAST-NNI(DSLAM)}$ = $VCI_{MULTICAST-UNI}$ (BRAS)
→ DSLAM sends Subscriber Respond with parameters (e.g. failure)
←→ BRAS queries the host according to IGMP to maintain subscription database
→BRAS receives IGMP Leave from Host
← BRAS sends Subscriber request (Subscriber-ID, Multicast Group…)
→DSLAM send Subscriber response (Subscriber-ID, Multicast Group…)
DSLAM deletes cross connect for multicast according to the parameter received from the BRAS

4.2.7 Multicast with Ethernet based replication

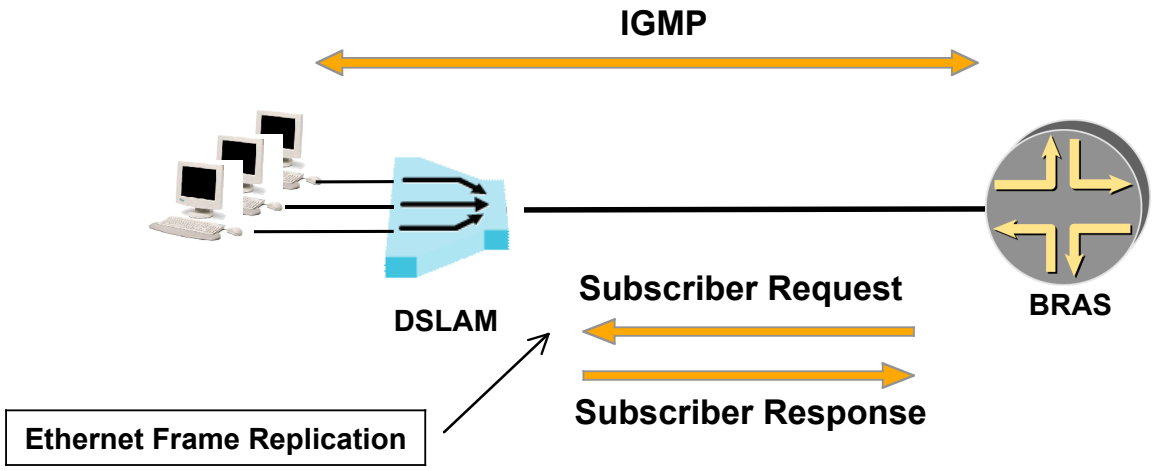


Figure 14: Ethernet Multicast

The following message flow description outlines the expected sequence of IGMP messages and Layer 2 control messages, to allow an Ethernet-based DSLAM multicast replication with IGMP and AAA functions residing the BRAS.

→ IGMP Membership Report message send from Host to BRAS. BRAS will analyze permissions of this subscriber locally (RADIUS or COPS).
← BRAS sends Subscriber Request(Subscriber-ID, Multicast Group.. .)
→ DSLAM sets up multicast replication table according to Subscriber-ID, Group Information
→ DSLAM sends Subscriber Respond with parameters (e.g. failure)
←→ BRAS queries the host according to IGMP to maintain subscription database
→BRAS receives IGMP Leave from Host
← BRAS sends Subscriber request (Subscriber ID, Multicast Group…)
→DSLAM send Subscriber response (Subscriber ID, Multicast Group…) and deletes multicast replication table entry.

4.2.8 Multicast with IGMP-Proxy in the DSLAM

The following message flow description outlines the expected sequence of IGMP messages and Layer 2 control messages, to allow the distribution of subscriber multicast entitlement data from the BRAS to an IGMP function that resides in the DSLAM.

A slightly different message flow would occur depending on the exact IGMP processing performed by the DSLAM (e.g. true IGMP-Proxy, IGMP router, transparent IGMP snooping).

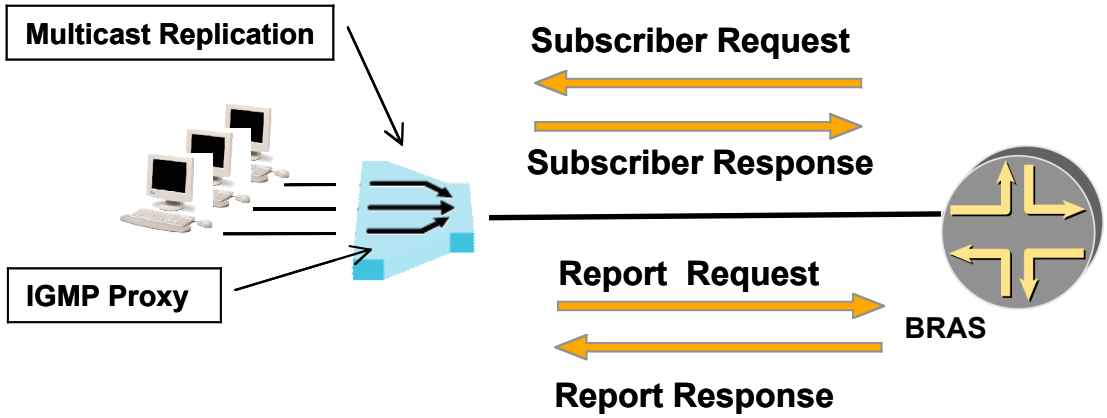


Figure 15: IGMP-Proxy

← BRAS sends Subscriber Request (Subscriber-ID, Multicast Entitlement Data (Group 1, Group 2…), Message Authentication)
→ DSLAM sends Subscriber Response()

→ IGMP Membership Report message sent from Host to DSLAM. DSLAM will analyze permission of the subscriber membership to that group, and if authorized will create a replication entry for (subscriber, group)
→ If membership is authorized, the DSLAM may send a Report Request message (Subscriber-ID, Data) when to indicate the change of available bandwidth for unicast traffic for this subscriber local loop. The DSLAM also forwards the IGMP message to the BRAS.
←BRAS will send a Report Response

→ DSLAM receives IGMP Leave from Host. DSLAM removes replication entry for (subscriber, group), then possibly forward the IGMP message to the BRAS.
→ If no more members of this multicast group are left, the DSLAM may send a Report Request message (Subscriber-ID, Data) when to indicate the change of available bandwidth for unicast traffic for this subscriber local loop. The DSLAM also forwards the IGMP message to the BRAS.
←BRAS will send a Report Response

5 Protocol Mapping

The use cases described within the document require the use of an inband control protocol between an Access Node (e.g. DSLAM) and a BRAS system.
This section will discuss possible protocol mappings for such layer 2 control messages. In order to facilitate a fair comparison a list of goals is used as already introduced by contribution dslf2004.034.
These goals are intentionally not described in a very formal way, but are more intended as a guidance to discuss particulars of a given protocol proposal.

The full discussion of possible protocol candidates was the focus of contribution dslf2004.152, which was discussed during the Toronto meeting.

5.1 Goals

It should be noted that the goals listed below may be achieved by the control protocol itself, or by underlying protocols (e.g. a transport layer).

High-Level Functional Goals

FG#1: the control protocol must address all use cases described in this contribution, and be general-purpose and extensible enough to foresee additional use cases (including the use of other Access Nodes than DSLAM, e.g. OLT for Passive Optical Networks).

Rationale: self-explanatory.

FG#2: the control protocol must be flexible enough to accommodate the various technologies that can be used in an access network and in the Access Node.

Rationale: as an example, use of Ethernet VLANs vs. ATM PVCs. Or the use of Ethernet data-link bridging as a multicast technique as opposed to ATM point-to-multipoint multicast technique.

FG#3: the protocol must be an open protocol, either an existing protocol endorsed by an appropriate standard body (e.g. IETF) or a new protocol which will be submitted for standardization to an appropriate standard body. A protocol with a formal way of defining an information model (objects being conveyed) is preferable. In any case, it must be possible for the DSL Forum to define additional protocol information elements. Some protocol aspects or some protocol information elements may be technology-specific (e.g. ATM vs. Ethernet), but must be vendor-independent.

Rationale: typical carrier requirement about protocols.

FG#4: the protocol must be transaction-oriented, allowing to reliably share states between the BRAS and the Access Node, and recover from loss of synchronization (e.g. node or link failure). Transactions must be either fully completed, or rolled-back to the previous state.

Rationale: most of the use cases involve a form of reliable state sharing. Deterministic transactions will also help to make appropriate local decisions (e.g. BRAS performing multicast by itself if the Access Node can't create a new layer2 multicast rule).

FG#5: the protocol must be able to recover from access network connectivity disruption and automatically resynchronize. It must also be able to recover from message losses on the access network.

Rationale: self-explanatory.

FG#6: the protocol must allow fast-paced transactions, in the order of magnitude of tens of transactions per second between a given pair (Access Node, BRAS). The protocol must allow fast completion of a given operation, in the order of magnitude of tens of milliseconds. The protocol must be scalable enough to allow a given BRAS to control hundreds of Access Nodes.

Rationale: as an example, end users zapping between multicast channels require fast responsiveness. The overall response time (as perceived by the end user) should not exceed a couple of hundreds of milliseconds, including IGMP processing. As to the number of Access Nodes, the requirement envisions high-density Access Nodes (e.g. managing thousands of local loops) as well as low-density Access Nodes (e.g. managing tens of local loops.

FG#7: the protocol must be simple and lightweight enough to allow an implementation on Access Nodes with limited control plane resources (e.g. CPU and memory).

Rationale: typical Access Nodes (e.g. DSLAMs) have limited control plane resources. It would be undesirable to require a hardware upgrade (as opposed to a software upgrade) for such systems. Also "keep it simple" engineering principle must be kept in mind. Finally, time to market and ease of implementation are important considerations.

FG#8: security must be supported, mostly to ensure the authenticity of the message initiator and the integrity of the message. The main goal is to protect the systems (Access Nodes and BRAS) against attacks.

Rationale: there is a proper balance to reach between security and simplicity. Pragmatic solutions relying on layer 2 dedicated paths (instead of creating

complex security mechanisms in the protocol itself) should be explored. More complex security mechanisms like encryption do not seem needed.

FG#9: the protocol should minimize sources of configuration mismatch, help automation of the overall operation of the systems involved (Access Nodes and BRAS) and be easy to troubleshoot.

Rationale: OPEX reduction goal and reduction of error-prone situations. Note that is rather common to have DSLAMs operated by a separate team (and sometimes company) than the BRAS systems, which emphasizes the need for such goal.

FG#10: the implementation of the control protocol in the BRAS and Access Nodes must be manageable via an element management interface. This must allow to retrieve statistics and alarms (e.g. via SNMP) about the operation of the control protocol, as well as initiate OAM operations and retrieve corresponding results.

Rationale: self-explanatory.

FG#11: a BRAS supporting layer 2 control must correlate layer 2 configuration data with the RADIUS authorization process and related subscriber data. This also applies to multicast channels entitlement.

Rationale: most operators today use RADIUS for subscriber management. It is therefore strongly desirable to be able to leverage on such infrastructure to get the subscriber-related data conveyed via the control protocol to the Access Node.

Design Goals

Again, these goals are suggested as general guidance. There might be variations on the suggested approach which would achieve similar purposes.

DG#1: the protocol should have a “boot” sequence allowing to inform the peer about control capabilities supported by the two peers (Access Node, BRAS) and negotiate a common subset. This sequence should be such that a system supporting the control protocol would automatically recognize when its peer doesn’t support it at all.

Rationale: foresee future evolutions while keeping upward compatibility. Plan for a step-by-step transition from existing deployments.

DG#2: the protocol should include a “keep-alive” mechanism to automatically detect loss of connectivity on the access network or failure of the peer node.

Rationale: self-explanatory. Allows to reset states to a known basis when connectivity is lost then restored.

DG#3: the protocol should provide a “shutdown” sequence allowing to inform the peer that the system is gracefully shutting down.

Rationale: this allows minimizing temporary “black holes” while the keep-alive didn’t detect the lack of connectivity yet.

DG#4: the protocol should include a “request/response” transaction-oriented model for the BRAS to communicate control decisions or request information from the Access Node. If the response is negative, then the state of the Access Node must be unchanged (roll-back).

Rationale: self-explanatory. Most operations implied by the use cases will likely follow this request/response model, which provides an implicit flow-control model as a byproduct.

DG#5: the protocol should include a “report” model for the Access Node to spontaneously communicate to the BRAS changes of states.

Rationale: as an example, change of DSL sync rate. Reports including statistics may also be envisioned using this model, although this may create flow-control challenges.

DG#6: the protocol should be mapped on top of the IP network layer (possibly via a transport layer).

Rationale: to foresee any type of layer 2 connectivity between the Access Node and the BRAS. It is nevertheless intended that the Access Node and the BRAS would be one IP “hop” away, no more.

5.2 Possible Protocol Mappings

After comparison between possible protocol mappings for Layer 2 control messages (dslf2004.152), one outcome of the Toronto meeting was to focus the analysis based on the two protocol candidates best suited for the task: COPS and GSMP.

Further work needs to occur to identify the incremental work required on these protocols for their use in the context of all use cases described in this document.

5.2.1 COPS Mapping

COPS is typically used as a policy management protocol, to execute fast-paced transactions between a Policy Decision Point (PDP; would be the BRAS in our case) and a Policy Enforcement Point (PEP; would be the Access Node in our case). Since most L2 control operations could be viewed as policy management operations, it is

therefore legitimate to ponder if COPS (in its COPS-PR incarnation) is a viable mapping.

COPS means **C**ommon **O**pen **P**olicy **S**ervice. COPS-PR is a COPS variant geared at PRovisioning operations (actually a bit of a misnomer, it's more about structuring namespaces and policy objects).

FUNCTIONAL GoalS	SHORT Analysis	Score
FG#1: General-purpose, extensible	COPS can be used for many purposes and is easily extensible via the definition of PIBs	PASS
FG#2: flexible for multiple access technologies	COPS can be easily conveyed across any access technology. COPS could be used to manage any type of policy.	PASS
FG#3: open protocol; information-model oriented	COPS is a fully open protocol defined by IETF. The SPPI syntax (Structure Of Policy Provisioning Information) allows to define new objects in PIBs (Policy Information Base). The DSL-Forum could define such PIBs on its own.	PASS
FG#4: transaction-oriented, state sharing, roll-back	COPS has been designed from day 1 to be transaction-oriented and to reliably share states between the PDP side and the PEP side.	PASS
FG#5: recover from connectivity/network disruptions	COPS includes resynchronization support to allow such recovery behavior.	PASS
FG#6: fast-paced transactions; scale	COPS object-oriented allows to perform complex operations in one request/response. It has been demonstrated to support hundreds of transactions per second.	PASS
FG#7: simple & lightweight (notably for Access Node)	COPS, COPS-PR and PIBs are rather complex. Several freeware protocol engine allow to get started, but this still remains somewhat complicated.	ISSUE
FG#8: security (e.g. authentication)	COPS can be secured via IPSec (probably too heavy-duty in our case) or via TLS (seems a good match). There is no lighter (authentication-only) form of security though.	PASS
FG#9: help automation of operation; easy to troubleshoot	No intrinsic mechanism to help automation, it's more a matter of PIB design. The binary and ASN.1-oriented aspect of COPS SPPI makes it tough to troubleshoot, although the object-oriented facet should help. Not many tools available on the market.	ISSUE
FG#10: manageable protocol engines	It wouldn't be very difficult to define a MIB for managing the "layer 2 control engines" themselves. COPS has its own protocol MIB.	PASS

DESIGN GoalS	SHORT Analysis	Score
DG#1: boot sequence, negotiate capabilities	COPS has a boot sequence allowing to negotiate capabilities.	PASS
DG#2: keep-alive mechanism	COPS has a keep-alive mechanism which can be defined to run at a high-pace.	PASS
DG#3: graceful shutdown sequence	COPS has a graceful shutdown sequence.	PASS
DG#4: request/response transactional model	COPS model allows a request/response initiated by either side (PEP or PDP), as a transactional-oriented operation. Its TCP mapping improves the reliability of the exchange.	PASS
DG#5: spontaneous report model	COPS reports allow such model, with possibly extensive reports. The “feedback PIB” allows to further structure such reports, notably in the case of statistics reporting.	PASS
DG#6: mapping on top of IP network layer	COPS typical mapping is TCP.	PASS
DG#7: support for long messages	COPS messages are not limited in size.	PASS

Additional CONSIDERATIONS	SHORT Analysis	Score
AC#1: standardization	Although endorsed by IETF as a set of RFCs, some of them (e.g. PIBs) are only informational and the debate between the COPS proponents and opponents has been raging for years. The cable bodies (PacketCable, Docsis) and the 3G mobile bodies (3GPP) have endorsed COPS much more enthusiastically.	PASS
AC#2: view from the trenches	COPS implementations are still relatively uncommon. Successful implementations do exist and are deployed in large-scale carrier environments, but COPS technology is clearly still emerging.	ISSUE
AC#3: time to market	As far as we know, no Access Node (e.g. DSLAMs) has a COPS interface. We’ve heard rumors of such plans, but nothing public. COPS toolkits are available to jump start development. Having the DSL-Forum define a new COPS PIB should not take very long (6 months?).	PASS

Overall, COPS/COPS-PR is an excellent match from a technical standpoint. Its relative complexity, the not-entirely-clear endorsement by IETF and the emerging nature of this protocol are points of concerns, but definitely not showstoppers.

5.2.2 GSMP Mapping

GSMP is typically used as a switch management protocol, to execute operations between a control plane (implemented on a separate device) and the data plane of a switch. Since most L2 control operations could be viewed as switch management operations, it is therefore legitimate to ponder if GSMP (in its latest incarnation, GSMP v3, RFC3292) is a viable mapping.

GSMP means **G**eneral **S**witch **M**anagement **P**rotocol.

FUNCTIONAL GoalS	SHORT Analysis	Score
FG#1: General-purpose, extensible	GSMP is not defined as an extensible protocol by another body than IETF. Yet the protocol structure is such that new versions of the protocol can easily support new operations.	PASS
FG#2: flexible for multiple access technologies	GSMP can be easily conveyed across any access technology. GSMPv3 is specialized for ATM, FR and MPLS switches. GSMP doesn't include operations geared at Ethernet switches and bridges, which would be required for a L2 control mapping for many cases of access technologies. But such extension shouldn't be very difficult to add.	ISSUE
FG#3: open protocol; information-model oriented	GSMP is a fully open protocol defined by IETF. Yet it is not information-model oriented and extensions would have to go through IETF.	ISSUE
FG#4: transaction-oriented, state sharing, roll-back	GSMP has been designed from day 1 to be transaction-oriented and to reliably share states.	PASS
FG#5: recover from connectivity/network disruptions	GSMP does not explicitly include generic resynchronization support, yet has enough operations to query the status of the switch, so that the goal can be reasonably fulfilled.	PASS
FG#6: fast-paced transactions; scale	GSMP allows performing complex operations in one request/response. Its Type-Length-Value syntax is compact and very fast to parse.	PASS
FG#7: simple & lightweight (notably for Access Node)	GSMPv3 includes a large number of operations and not all of them appear relevant in the context of L2 control operations. Yet most of them are useful, and the protocol is quite straightforward and apparently pretty simple to implement, notably for engineers used to layer 2/3/4 protocols which are typically using a similar syntax.	PASS
FG#8: security (e.g. authentication)	GSMP can be secured via IPSec (probably too heavy-duty in our case). GSMP doesn't include security elements by itself. Although not	ISSUE

	standardized today, a TLS/TCP mapping shouldn't be difficult to add.	
FG#9: help automation of operation; easy to troubleshoot	GSMP explicitly includes discovery operations which should help reduce pre-configuration needs. Most network engineers and operators are used to TLV-oriented syntaxes when troubleshooting. Yet not many protocol analyzers or debugging tool exists today with GSMP support.	PASS
FG#10: manageable protocol engines	GSMP has its own protocol MIB, which includes full support for technology-specific operations.	PASS

DESIGN GoalS	SHORT Analysis	Score
DG#1: boot sequence, negotiate capabilities	GSMP has a boot sequence allowing to negotiate capabilities.	PASS
DG#2: keep-alive mechanism	GSMP has a keep-alive mechanism which can be defined to run at a high-pace.	PASS
DG#3: graceful shutdown sequence	GSMP has a graceful shutdown sequence.	PASS
DG#4: request/response transactional model	GSMP model allows a request/response initiated by the controller side, as a transactional-oriented operation. Its TCP mapping improves the reliability of the exchange.	PASS
DG#5: spontaneous report model	GSMP enables a report model, notably for statistics or state changes (events).	PASS
DG#6: mapping on top of IP network layer	GSMP typical mapping is TCP/IP, although it can also be mapped on ATM or the Ethernet MAC layer.	PASS
DG#7: support for long messages	GSMP messages are not limited in size. A segmentation mechanism is included in the protocol itself.	PASS

Additional CONSIDERATIONS	SHORT Analysis	Score
AC#1: standardization	GSMPv3 is endorsed by IETF as a set of standard-track RFCs, but the working group is now dormant. As far as we know, no other standard body or industry forum has been endorsing GSMP, with the exception of the Multi-Service Forum (MSF).	ISSUE
AC#2: view from the trenches	GSMP implementations exist but are still relatively rare. Successful implementations do exist and are deployed in large-scale environments, but GSMP technology is admittedly still emerging.	ISSUE

AC#3: time to market	As far as we know, no Access Node (e.g. DSLAMs) has a GSMP interface. GSMP is clearly designed to be usable in an embedded environment, and should not require a 3rd-party protocol toolkit. Having the DSL-Forum and IETF work in a coordinated fashion on a GSMPv4 (e.g. adding Ethernet switching/bridging support) is likely to take a while (one year?), although GSMPv3 has a strong foundation for L2 control operations.	ISSUE

Additional considerations:

- GSMPv3 existing set of operations match very well the set of operations we've identified so far for L2 control messages. One can create point-to-point or point-to-multipoint cross-connections (cf. multicast on ATM), configure connections, retrieve the configuration of the switch connections, get statistics, notifications from state changes, boot sequence, keep-alive, shutdown sequence, etc. The main caveat is that Ethernet switching/bridging operations are not defined so far.
- GSMP was originally designed for MPLS switches, which may prove handy in the future if MPLS starts moving to the metro network (e.g. VPLS, etc).
- Concerns about directly manipulating VPI/VCI numbers (quantities with a local scope) will have to be addressed, probably by beefing up the discovery process. This is actually a rather general issue for any of the protocol mappings being proposed.

Overall, GSMP is an excellent match from a technical standpoint. The main caveat is the lack of Ethernet switching/bridging support but this could be easily addressed (assuming IETF would agree to revive the working group). The main point of concern is the lack of momentum of this protocol, and its limited support so far.

6 Abbreviations and Glossary

AAA	Administration, Authorization, Authentication
BRAS	Broadband Remote Access Server
CAPEX	Capital Expenditure
COPS	Common Open Policy Service
CPE	Customer Premises Equipment
DG	Design Goal
DHCP	Dynamic Host Control Protocol
DSL	Digital Subscriber Line
DSLAM	DSL Access Multiplexer
EFM	Ethernet in the First Mile
FG	Functional Goal
IGMP	Internet Group Management Protocol
MAC	Media Access Control
MPLS	Multi-Protocol Label Switching

OAM	Operation, Administration, Maintenance
OLT	Optical Line Termination
OPEX	Operational Expenditure
OSS	Operation Support System
PIM	Protocol Independent Multicast
PON	Passive Optical Network
POP	Point Of Presence
PPP	Point to Point Protocol
PPPoE	PPP over Ethernet
RADIUS	Remote Authentication Dial In User Service
RG	Routing Gateway
RT	Remote Terminal
SNMP	Simple Network Management Protocol
VLAN	Virtual LAN

5

Add/Delete Multicast Entry	Specify whether to add/delete a multicast entry
Device Type	BRAS, DSLAM etc
DSL Line ID	A physical identifier as defined in contribution dslf2004.071used in topology reports
DSLAM Uplink ID & Bandwidth	The uplink and bandwidth from DSLAM
Message Authentication	Authentication field for message
Multicast Capable	Device is capable of multicast services
OAM Test Results	Results of OAM test
Other capability	Other capabilities that may be defined in the future
RT/DSLAM Link ID & Bandwidth	Link between DSLAM and RT and its bandwidth
Sequence Number	Sequence number of message
Statistics	Statistical data
Statistics Capable	Device is capable of sending or receiving statistics
Subscriber ID	A subscriber identifier as perceived by the BRAS and used in any subscriber-related operation
Sync Rate	Sync Rate of DSL Line
Topology Capable	Device is capable of sending or receiving topology data
Topology Data	Network topology information

7 References

[1] DSLForum TR-058, Multi-Service Architecture & Framework Requirements, Mark Elias (SBC) and Sven Ooghe (Alcatel), 09/2003

[2] DSLForum TR-059, DSL Evolution - Architecture Requirements for the Support of QoS-Enabled IP Services, Tom Anschutz (BellSouth Telecommunications), 09/2003

[3] DSLForum2004.071.01, Use of DHCP Relay Agents and PPPoE Intermediate Agents for DSL line identification in Ethernet-based access networks, Jerome Moisand (Juniper), Thomas Gemmer (Siemens), Ole Helleberg Andersen (Ericsson), Amit Cohen (ECI), Sven Ooghe (Alcatel), Thomas Meehan (Redback) 02/2004

[4] DSLForum2003.367.00, Layer 2 Control Mechanism – Multicast, Jerome Moisand (Juniper), Norbert Voigt (Siemens), Ole Helleberg Andersen (Ericsson), 11/2003

[5] DSLForum2003.368.00, Layer 2 Control Mechanism – Topology Discovery and Line Configuration, Jerome Moisand (Juniper), Norbert Voigt (Siemens), 11/2003

[6] DSLForum2003.376.00, Aspects of Multicast Services, Thomas Haag (Deutsche Telekom), 11/2003

[7] DSLForum2003.434.00, DSLAM as a Service Proxy For Video Distribution, Ran Avital and Amit Cohen (both ECI), 11/2003

[8] DSLForum2003.441.00, A Subscriber Update Protocol, Thomas Meehan (Redback), Magnus Asbo (Nokia), 11/2003

[9] DSLForum2004.034.00, Layer 2 Control Mechanism – Use Cases, Jerome Moisand (Juniper), Norbert Voigt (Siemens), Thomas Haag (Deutsche Telekom), Ran Avital (ECI), Thomas Meehan (Redback), 03/2004

[10] DSLForum2004.071.00, Use of DHCP Relay Agents and PPPoE Intermediate Agents for DSL line identification in Ethernet-based access networks, Jerome Moisand (Juniper), Thomas Gemmer (Siemens), Ole Helleberg Andersen (Ericsson), Amit Cohen (ECI), Sven Ooghe (Alcatel), Thomas Meehan (Redback), 04/2004

[11] DSLForum2004.087.00, Layer 2 Control Mechanism – Message Flows, Thomas Meehan (Redback), Jerome Moisand (Juniper), Norbert Voigt (Siemens), Thomas Haag (Deutsche Telekom), Ran Avital (ECI), 03/2004

[12] ITU-T H.610 Full-Service VDSL - System architecture and customer premises equipment, 07/2003

[13] ITU-T Y.1730 Requirements for OAM functions in Ethernet-based networks and Ethernet services, 01/2004

[14] ITU-T Y.17ethoam OAM functions and mechanisms for Ethernet based networks, Draft 02/2004

Appendix A: DSL line rate resynchronization

Topology discovery (as described in section 3.1) is specifically important in case the net data rate the DSL line changes overtime. The DSL net data rate may first of all be different every time the DSL modem is turned on. Additionally, during the time the DSL modem is active, data rate changes can occur at two levels:

1. **Rate-adaptive at startup (applicable for ADSL1 & ADSL2)**
   This model is commonly deployed. When sudden large changes in line conditions would drop the SNR margin below the minimum required, the DSL line goes "out of sync" and will retrain to a lower value. Even if the cause of the re-initialization disappears later on, the trained rate will not be changed in showtime.

2. **Rate-adaptive in showtime (optionally applicable for ADSL2)**
   This model is called "Seamless Rate Adaptation" (SRA) and allows to dynamically change the DSL data rate, as an alternative to re-initialization due to sudden SNR margin decreases. SRA allows to change the data rate of the connection while in operation without any service interruption or bit errors (temporarily lower net data rate during interference).

Frequency of DSL net data rate changes

Depending on whether the DSL modem is "always on" or not, the frequency of DSL net data rate changes at modem boot time will be smaller or higher. Every time the DSL modem is turned on, the net data rate may potentially be different. This especially applies to USB modems.

In case of rate-adaptive behavior at startup, the DSL net data rate will change every time the SNR drops below the minimum required. The DSL line retrains to a new and possibly lower value. This takes between 15 and 40 seconds.

Numerical data collected in several operator networks shows that in practice, about 10% of the DSL lines has an unstable behavior. These lines will resynchronize between once every hour and once every four minutes. The other 90% of the DSL lines has a more stable behavior and will resynchronize once every 12 hours or more. The average time between two resynchronizations is therefore 0.9 * 12 + 0.1 * 1 hours = 10.9 hours.

There is a tradeoff between high planned bitrates and noise margin. If an operator decides to go for higher planned bitrates, the noise margin will be smaller and DSL lines will be more susceptible to the SNR deterioration factors described below. As a result the DSL line will become more instable and net data rate changes will occur more frequently.

In case of SRA the DSL net data rate may change even more frequently, since the granularity of SNR variations is smaller. At the same time, the size of the net data rate change will be smaller. Use of SRA may considerably impact the use of topology discovery (scaling wise). This topic may require future study.

Net data rate increases don't need to be sent to the BRAS immediately. These changes do not impact the overall behavior of the currently active services, but merely result in the BRAS being aware of additional bandwidth that could be used for the services later on. Therefore, a net data rate increase could be announced to the BRAS once the change is considered as "stable". This allows coping with a certain degree of temporary bit rate changes that would otherwise result in a number of bit rate announcements with limited applicability. One approach to determine the stability of a bit rate increase is to use a hold-off timer that would keep track of how long the net data rate hasn't changed. Once the hold-off timer would be reached, the DSLAM could send the updated bit rate to the BRAS.

In order for the hierarchical scheduling model to work correctly, all decreases of the net data rate should be announced to the BRAS in real-time. This ensures the BRAS is able to reduce the amount of best effort traffic in order to guarantee the quality of other services. Some temporary impact may still occur as the synchronization of the schedulers may not be immediate.

Sources of SNR deterioration

The probability of a line going out of sync or of a line initiating SRA depends on a number of factors. As a result of SNR variations, the attainable bit rate will vary heavily over time. As long as the attainable bit rate does not drop below the operator specified planned bit rate, the DSL line behavior is stable. Otherwise, the modem will re-synchronize to a new value or will initiate SRA. Factors impacting the attainable bit rate are listed below.

- **Cross-talk**
  Crosstalk causes by far the largest contribution to capacity limiting noise for DSL systems. The level of cross-talk depends on the number of active DSL lines in a binder, which depends on the number of copper pairs in the binder and the DSL subscription take rate. When a new DSL source (i.e. a crosstalker) is activated, the noise on other DSL lines can get significantly larger. One also has to take into account that twisted pairs are split into different lengths. Every part of the DSL line could potentially end up in a different binder.

  In case the line went out of showtime due to a new crosstalker, the data rate after re-synchronization may be significantly smaller than before. Assuming the drop would be 6 dB over the entire bandwidth, this results in a decrease of at least 1.5 Mbit/s.

Impulse noise

Impulse noise consists of intermittent, undesirable signals induced by external sources such as lightning, switching equipment, and heavy electrically operated machinery such as elevator motors and copying machines. It increases or decreases a circuit's signal level, which causes the receiving equipment to misinterpret the signal. Due to the fact that the interference is very temporary, the DSL net data rate after re-sync will typically be the same as before the occurrence of the interference.

Radio Frequency Interference (RFI)

This comprises interference caused by other sources that are sending out electric signals (e.g. an AM broadcasting station). These sources may considerably hamper the stability of the DSL line.

"Dimmer Noise"

This interference type is due to dimmers and electronic equipment in the home. Sometimes this is also seen as a form of impulse noise.

Appendix B: Topology discovery -- DSL parameters

The following table gives a useful extract of parameters that can be reported from the DSLAM to the BRAS with a L2-control Topology Discovery interaction. This table is an extract of ITU-T G.997, section 7.5 ("Test, diagnostic and status parameters").

Pos.	Message Type	Information	Reference	Priority
1	DSL Type: ADSL Transmission System	Which DSL type is connected (e.g. ADSL1, ADSL2, SDSL, VDSL,…) This parameter defines the transmission system in use	Not described ITU-T G.997 Section 7.5.1.1	mandatory
2	DSL Link State	Line up (showtime)/ line down (idle or silent)	(G.992. Annexe D / G.992.3 Annexe D)	mandatory
3	Actual data rate Up- and Downstream	Actual data rate upstream and downstream of a synchronized DSL link	ITU-T G.997 Section 7.5.2.1	mandatory
4	Attainable Data Rate Up- and Downstream	Maximum datarate which can be achieved.	ITU-T G.997 Section 7.5.1.12 and 7.5.1.13	mandatory
5	Minimum Data Rate	minimum data rate desired by the operator in bit/s.	ITU-T G.997 Section 7.3.1.1.1	optional
6	Maximum Data Rate	maximum data rate desired by the operator in bit/s.	ITU-T G.997 Section 7.3.2.1.3	optional
7	Minimum Data Rate in low power state	minimum data rate desired by the operator during the low power state (L1/L2).	ITU-T G.997 Section 7.3.2.1.5	optional
8	Maximum Interleaving Delay	maximum one-way interleaving delay	ITU-T G.997 Section 7.3.2.2	optional
9	Actual interleaving Delay	Value in milliseconds which corresponds to inter leaver setting.	ITU-T G.997 section 7.5.2.3	optional

Table 3 DSL diagnostic and status parameters

Appendix C: Topologies with Remote Terminals

Introduction

The following general reference architecture will be used in this document to classify various aggregation scenarios in the Access Network involving a form of Remote Terminals (including optical ONUs or ONTs). The reference architecture depicted below is generally based on the H.610 reference architecture yet complements it to include FTTP access in addition to DSL-based access.

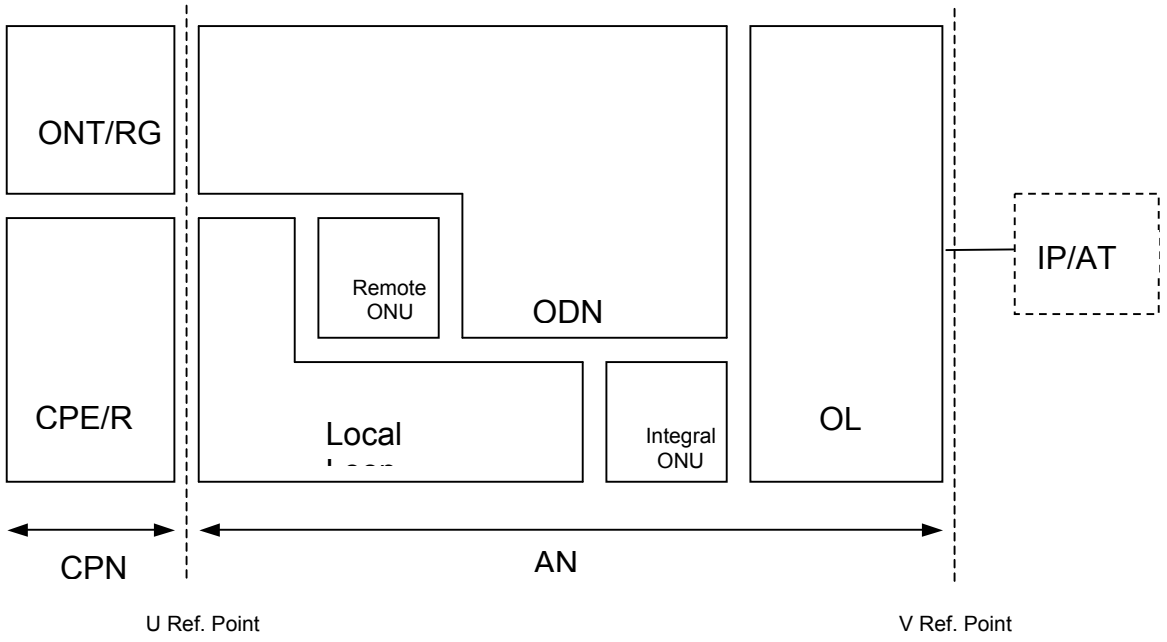


Figure 16: General reference architecture

The two extreme scenarios represented by Fig. 1 are: (1) CO-based DSL access, where the ONU function is in effect integrated with the OLT function in one platform as customary in the majority of ADSL deployments, and (2) All-fiber access, where an optical network termination is being done at the customer premises network. In the interim case, a remote ONU (i.e. DSLAM) is located closer to the end-user (e.g. at the cabinet), with last mile connectivity still being copper-based.

In all these scenarios, various aggregation schemes could exist, both between the OLT-ONU as well as internally among several OLTs or ONUs, in order to create a cluster of several units, which is optimized to the number of subscribers and the expected service penetration at the desired location.

Classification of aggregation topologies

For each of the scenarios illustrated in Fig. 1 here above, aggregation and clustering could be achieved through the use of various topologies and technologies. Topology choices include P2P, P2MP and ring architectures, while technology choices can span standard STM-i/OC-j interfaces, Gigabit Ethernet, PON, RPR, proprietary implementations and more.

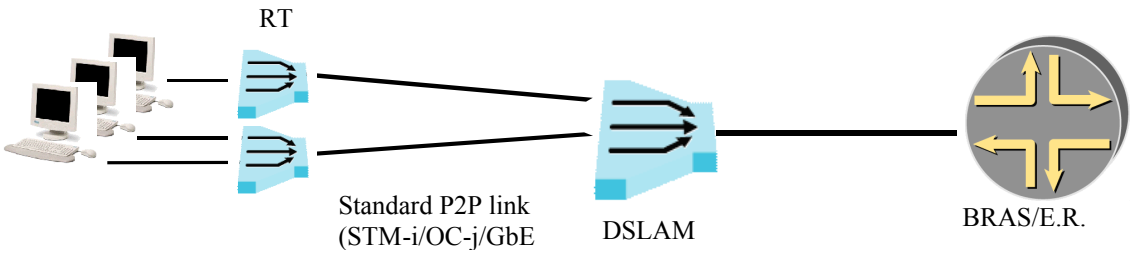
Table 1 below classifies typical topologies:

- **P2P (Point-to-Point) aggregation**: typically uses STM-i/OC-j interfaces or Gigabit Ethernet
- **P2MP (Point-to-Multipoint) aggregation**: various xPON technologies are possible, including BPON, GPON and EPON.
- **Ring aggregation**: could be done using access SDH/PDH type of solutions, Gigabit Ethernet, RPR etc.
- **Proprietary aggregation**: typically applicable to clustering among OLTs or ONUs by means of platform bus extension. However, a long haul extension of Remote ONUs to an OLT could also be based on this approach. Clustering is typically different in the fact that only the master unit - whether it is the ONU interfacing the ODN, the OLT interfacing the IP/ATM core or the OLT to which remote ONUs are connected – serves as a managed NE (i.e. SNMP agent).

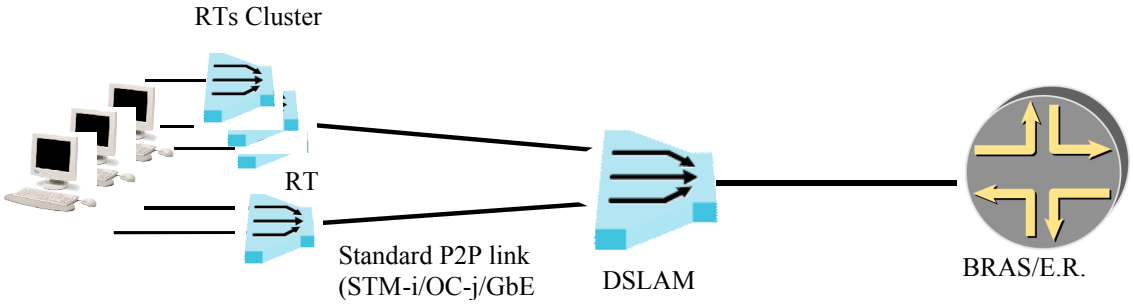
	Access Node(s)	**ODN**				**Clustering of OLTs or ONUs**	**Comments**
	OLT-ONU aggregation or clustering method	P2P	P2MP	Ring	Propriety	Propriety	
1.	OLT with integral ONU	N/A				+	
2.	OLT and remote ONU						
2.1.	- " -	+					
2.2.	- " -		+				
2.3.	- " -			+			
2.4.	- " -				+		
3.	Remote ONU						
3.1.							
3.2.							

Examples of Aggregation Scenarios

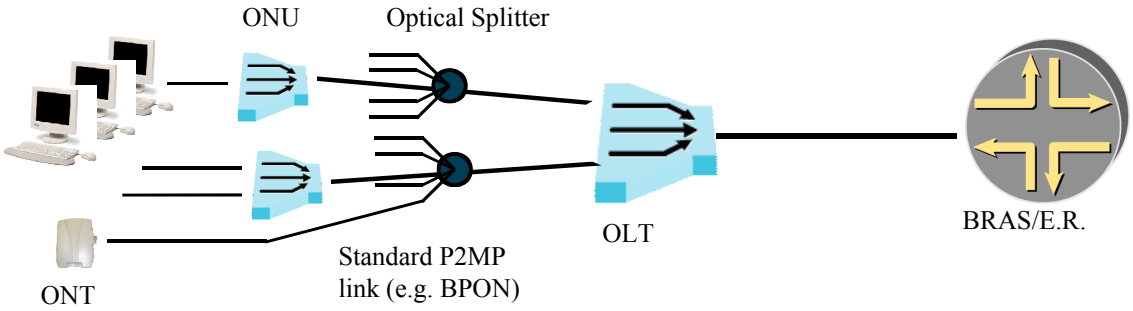
OLT, Remote ONUs, P2P aggregation



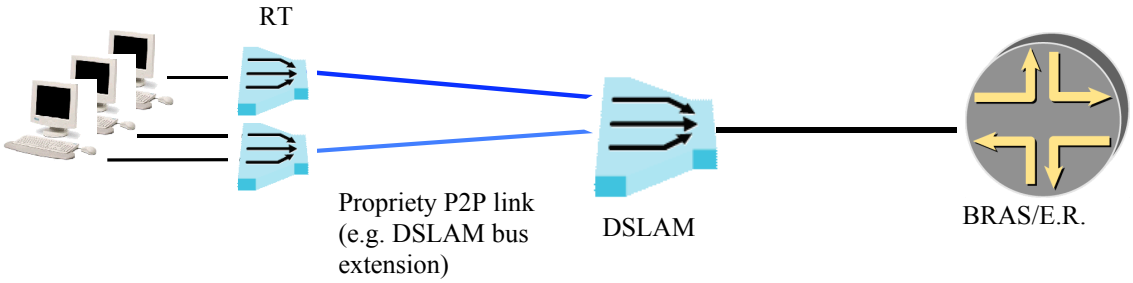
5 OLT, Remote ONUs or ONTs, P2P aggregation, with ONU clustering



OLT, Remote ONUs or ONTs, P2MP aggregation



OLT, Remote ONUs connected using OLT bus extension



Layer 2 control considerations

Most use cases described is the document will require a level of interaction between the (master) DSLAM/OLT and the Remote Terminals, in order to enable proper L2-control interactions between the (master) DSLAM/OLT and the BRAS.

Most typically, there is already an inband control scheme between the master DSLAM and the Remote Terminals, if only for such a group of devices to be managed as a cluster by OSS systems. Two main variations appear to be deployed:

a) In a PON scheme, an open cell-based protocol dubbed OMCI is defined in the ITU recommendations (G.983, G.984) to achieve this purpose.
b) Otherwise, no standard exists, but proprietary mechanisms have been developed by DSLAM vendors allowing such interaction.

As long such RT/DSLAM interaction exists (to report sync rates, to delegate the enforcement of a DSL configuration parameter, etc), it is possible for the master DSLAM (or OLT) to also act as a cluster when interacting with the BRAS via a form of Layer2-control.

Such distributed architecture is therefore compatible with all Layer2-control mechanisms described in the main document.