IPTV/Video over Broadband

SwiNOG-12
04.05.2006

Thomas Kernen, thkernen@cisco.com

SP NGN IPTV Consulting Engineer

Agenda

- Introduction
- IPTV and Standards
- IPTV Building blocks
- IPTV Services
- Video over Broadband Architecture
- Design concepts
- Redundancy
- IPTV Quality of Experience (QoE)
- Q & A

Introduction

Focus

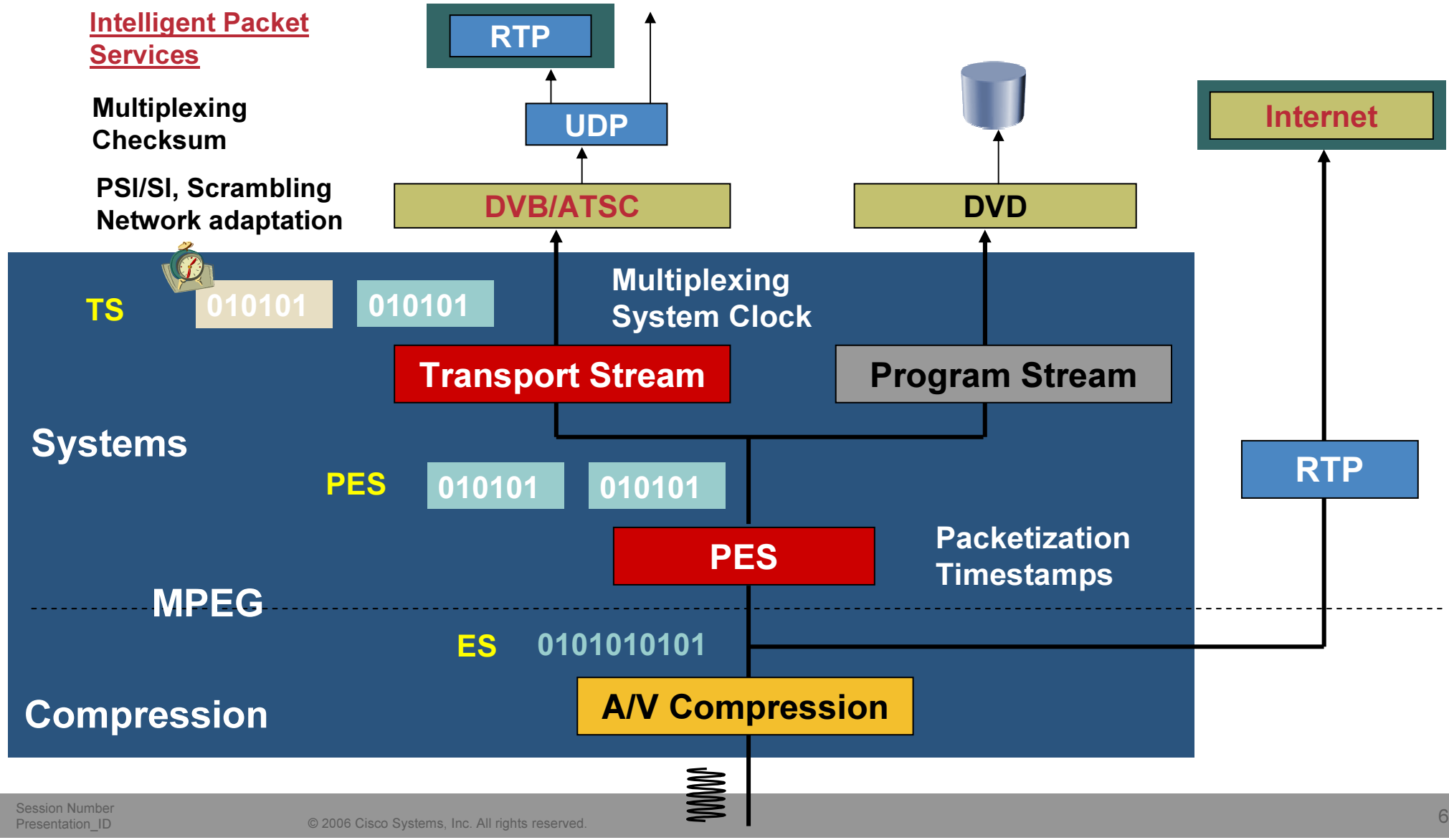
This presentation is about transporting broadcast video and video on demand services an over IP network infrastructure including a broadband last mile.

Due to time constraints, details about video encoding, middleware interaction, set top box designs, home networking, user interfaces, etc. are not covered in this session.

Introduction

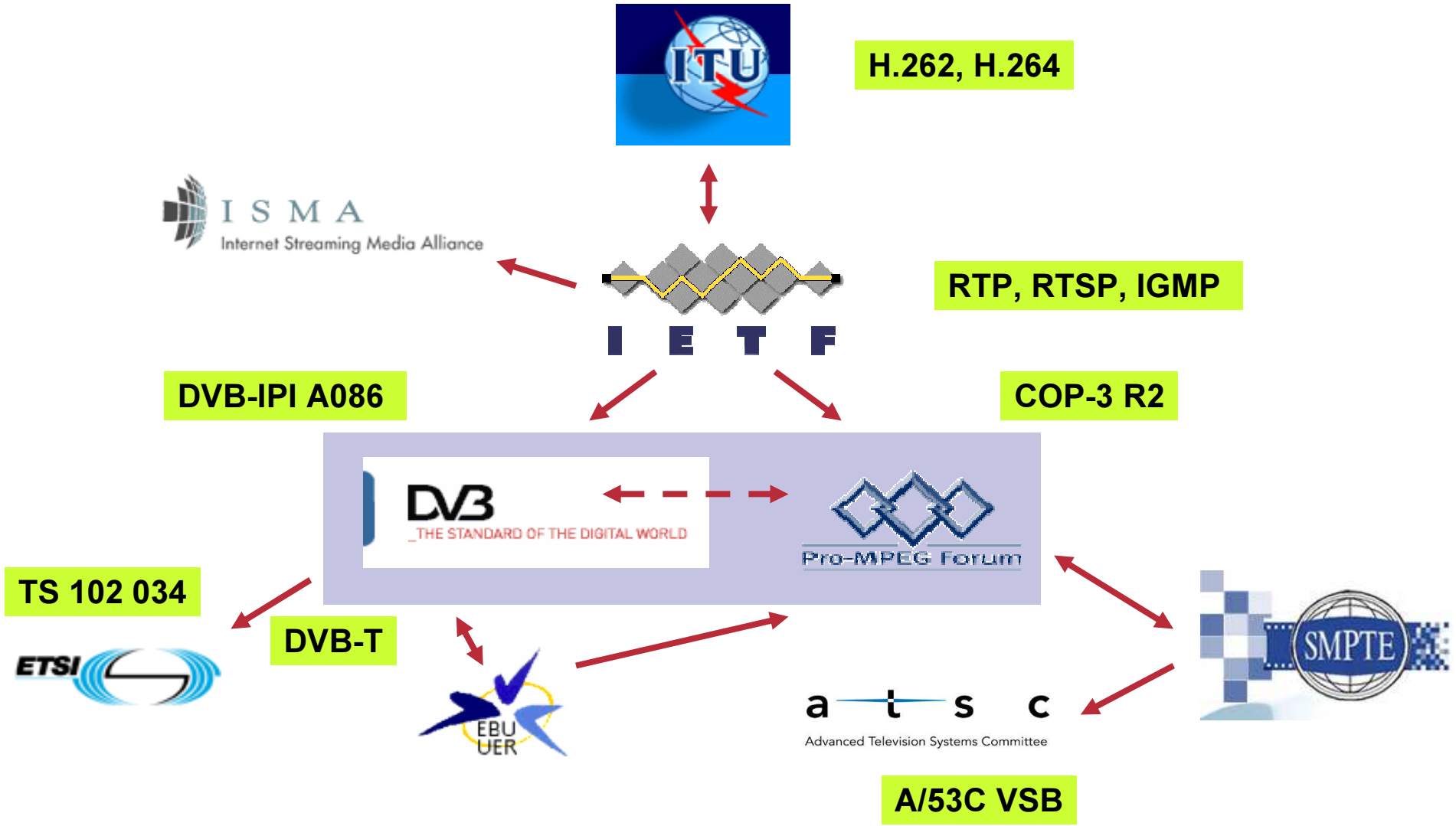
- **IPTV is happening: 4M subscribers at the end of 2005**
  - **Enhanced customer experience driven by IP**
- **IPTV services made possible by Broadband access leveraging on Carrier Ethernet architecture**
  - **Foundation for 3play services**
  - **Uses NGN architecture for service scalability**
- **IPTV is an innovative technology that borrows many things from years of DVB/ATSC experience**
  - **MPEG digital video compression and transport for compatibility**
  - **IP for service flexibility**

IPTV in perspective



IPTV and standards

IPTV Standardization



IETF (signaling)

- **Multicast-based services**

  RFC-3376 IGMPv3

- **On-demand services**

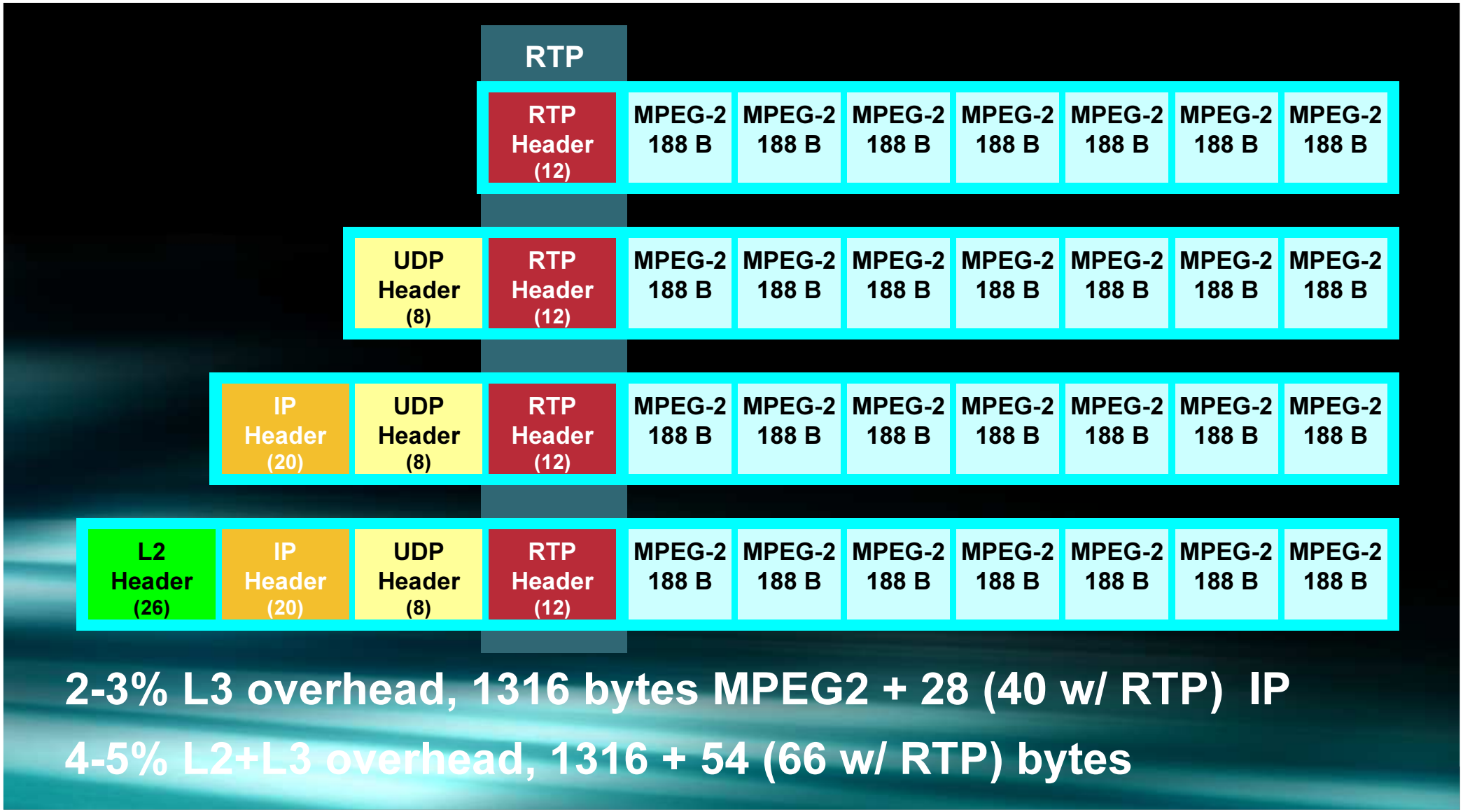
  RFC-2326 RTSP

- **Other IETF standards relevant: DHCP, DNS, …**
- **However, IGMPv2 is by far the most common signaling protocol used today**

IETF (transport)

- **Transport Service provided jointly by**
  - **UDP: checksum and multiplexing**
  - **RTP: sequencing and timestamping**
- **Defined in RFC-2250 "RTP Payload for MPEG-1/2 streams"**
  - **Section 2. MPEG TS encapsulation (MP2T payload type)**
  - **Section 3. MPEG ES encapsulation (MPV)**
- **However, industry keeps on using the "de-facto" MPEG-TS/UDP encapsulation**

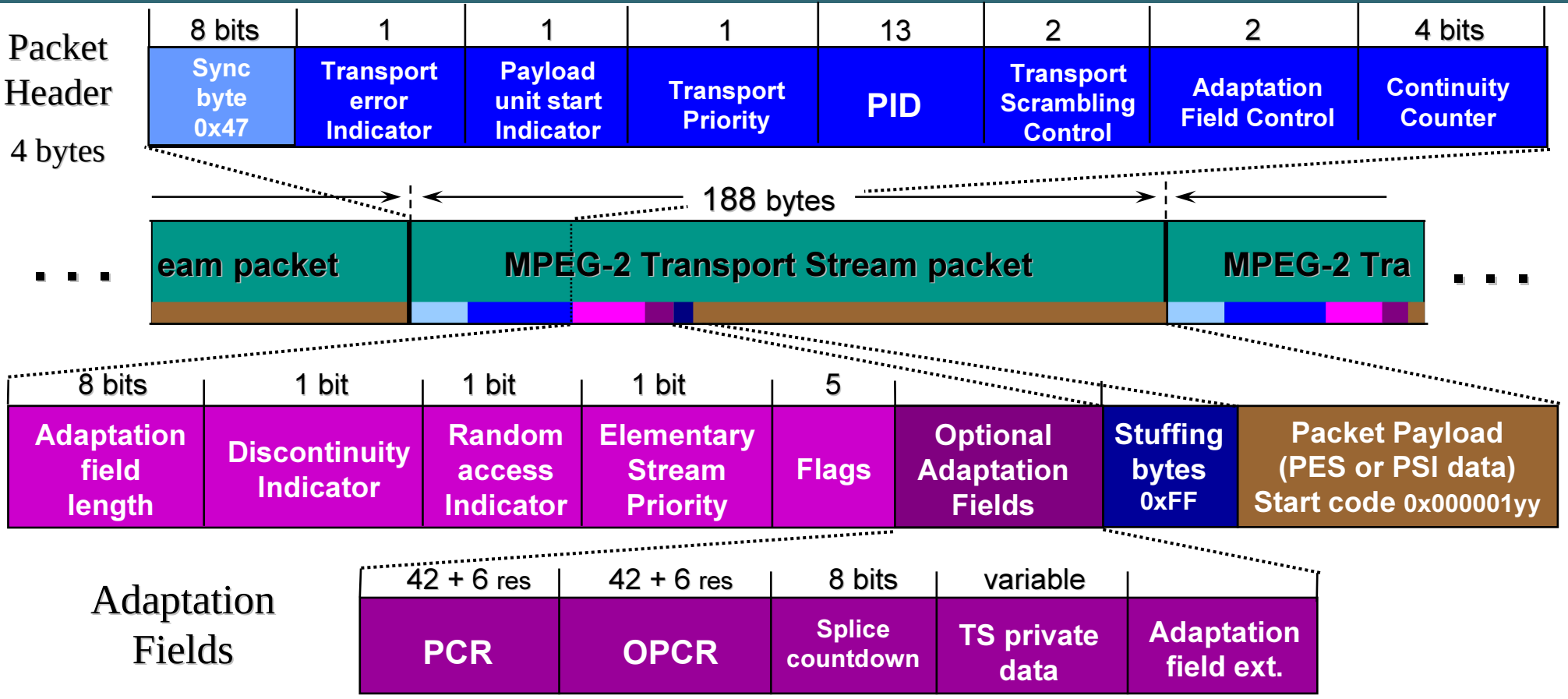
Encapsulation of MPEG-2 TS on IP



2-3% L3 overhead, 1316 bytes MPEG2 + 28 (40 w/ RTP) IP

4-5% L2+L3 overhead, 1316 + 54 (66 w/ RTP) bytes

MPEG-2 Transport Stream Details



- **Each 188 byte Transport Stream packet contains data from one elementary stream or PSI/SI data as defined by the 13 bit PID value**

DVB-IPI An architecture for the delivery of DVB services over IP networks

- DVB-IPI Bluebook A086 – ETSI TS 102 034 1.1.1 (2005-03)
- Service delivery
  - MPEG-2 TS encapsulation
  - Use of RTP and RTSP for VoD
- Service discovery
  - Using DVB PSI/SI information and DNS
- Service selection
  - Using IGMP and RTSP
- STB authorization (DHCP)
- Total jitter < +/- 20 msec
- Packet loss < 1 noticeable artifact per hour

http://webapp.etsi.org/workprogram/Report_WorkItem.asp?WKI_ID=15164

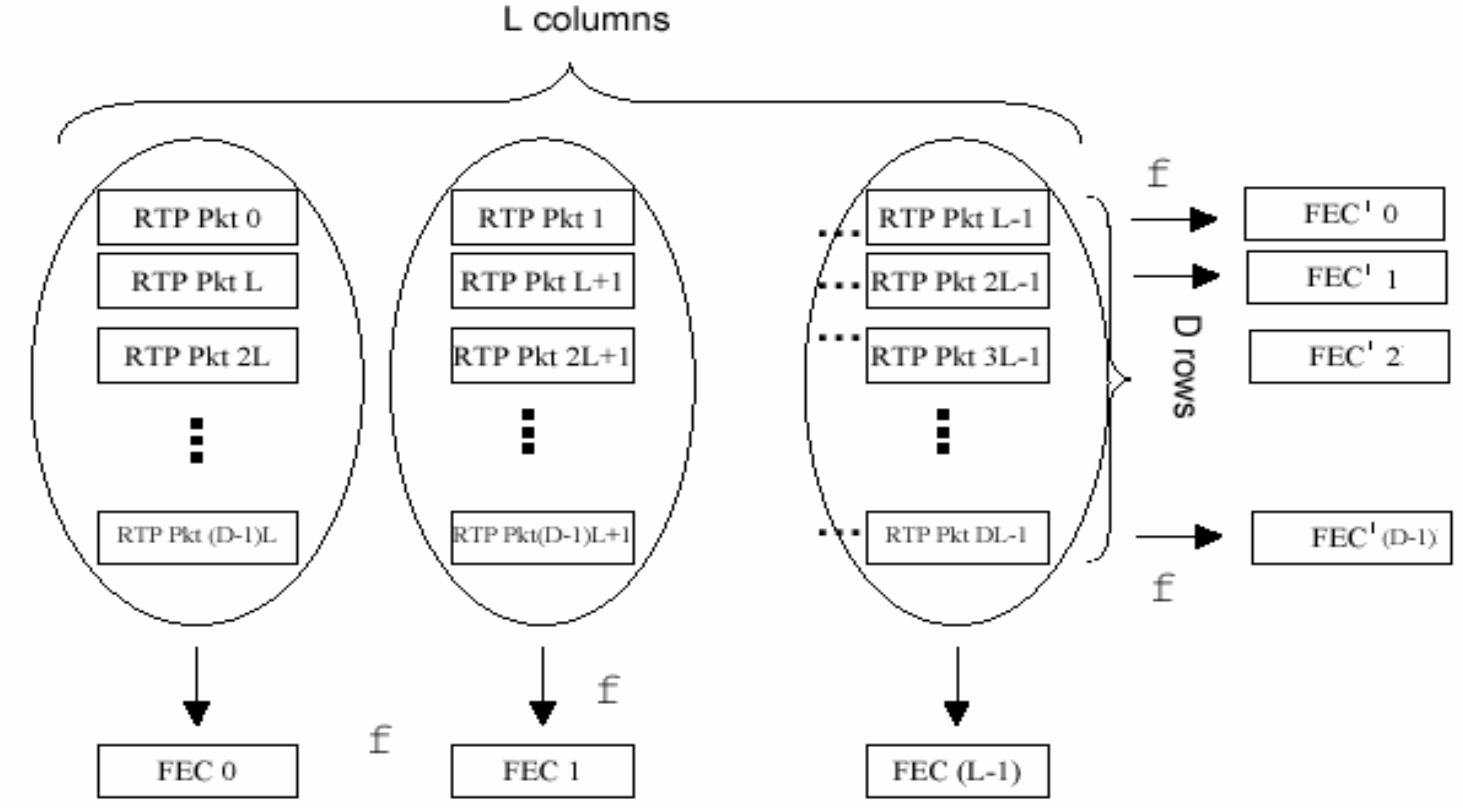
Pro-MPEG

- Professional Video Transport focus
- Transport of MPEG-2 (COP-3) and uncompressed SMPTE 292M (HDTV) (COP-4)
- Use of RTP
- Generally more strict in packet loss, less in jitter
- Jitter < +/- 60 msec
- FEC scheme for error recovery

http://www.pro-mpeg.org/publications/pdf/Vid-on-IP-CoP3-r2.pdf
http://www.pro-mpeg.org/publications/pdf/HBRSS-on-IP-CoP4-r1.pdf

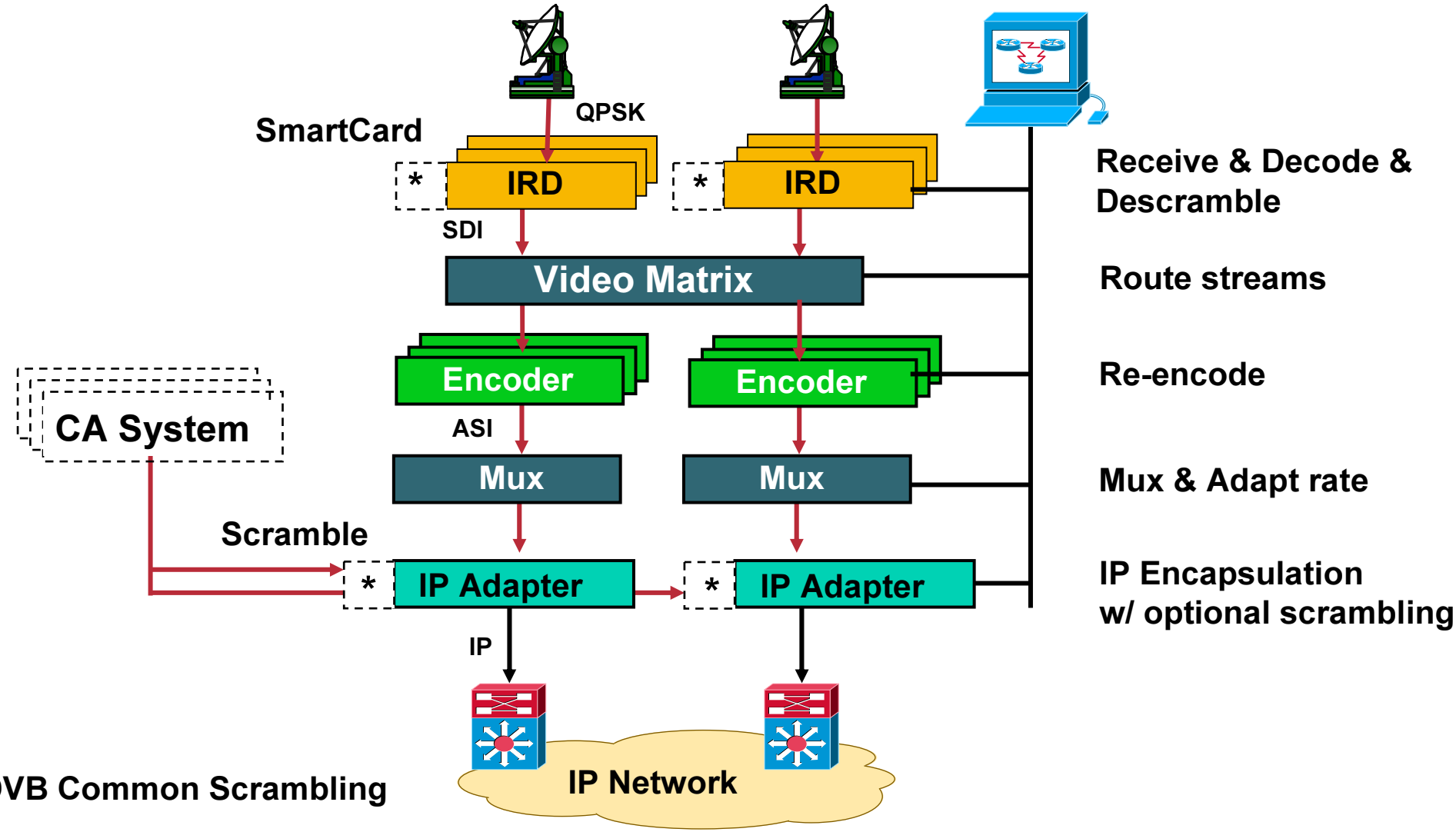
Pro-MPEG COP-3 2D-FEC scheme

- **Based on RFC2733, XOR FEC packets**
- **FEC streams sent on UDP ports RTP+2 (col), RTP+4 (row)**

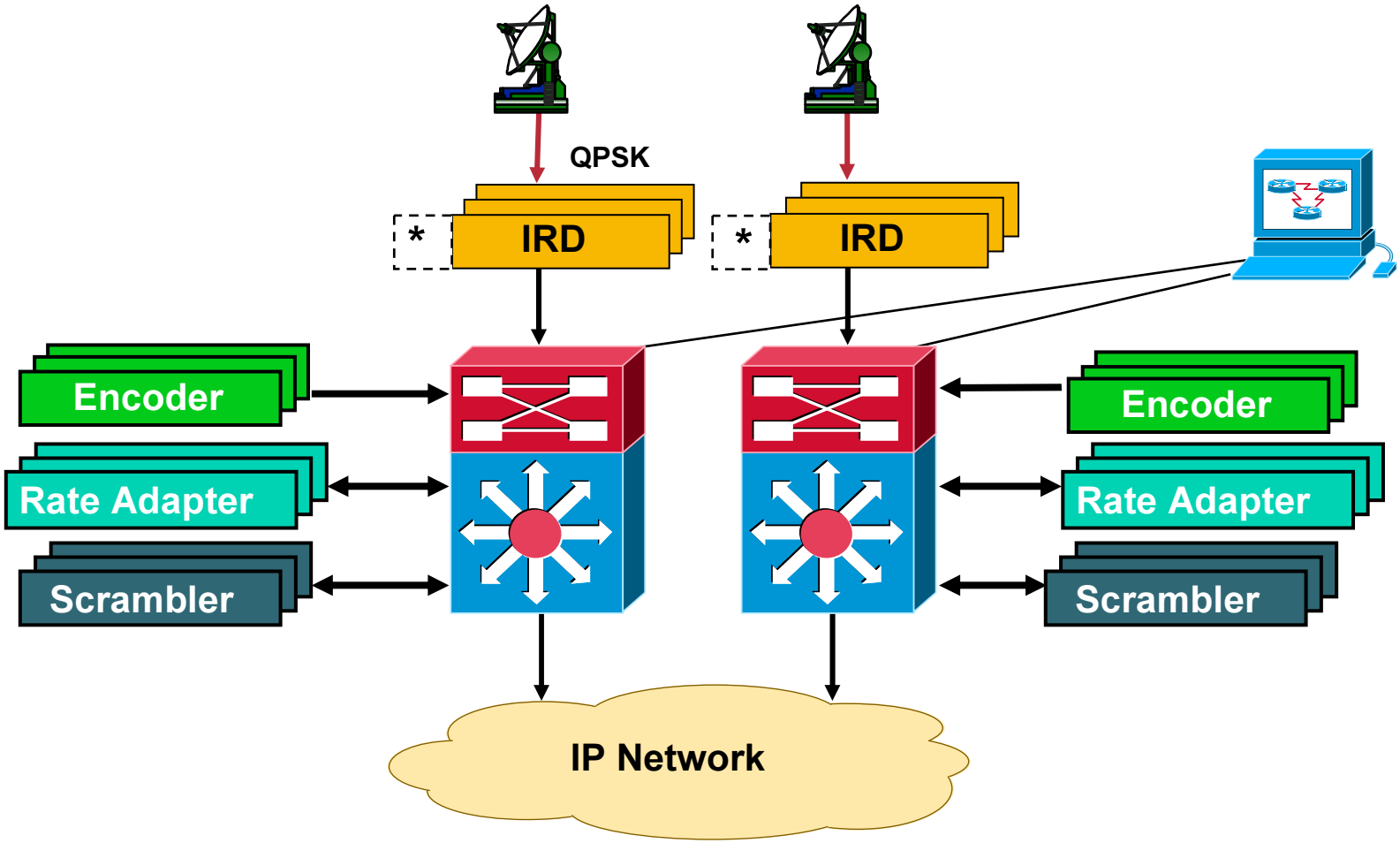


IPTV Building blocks

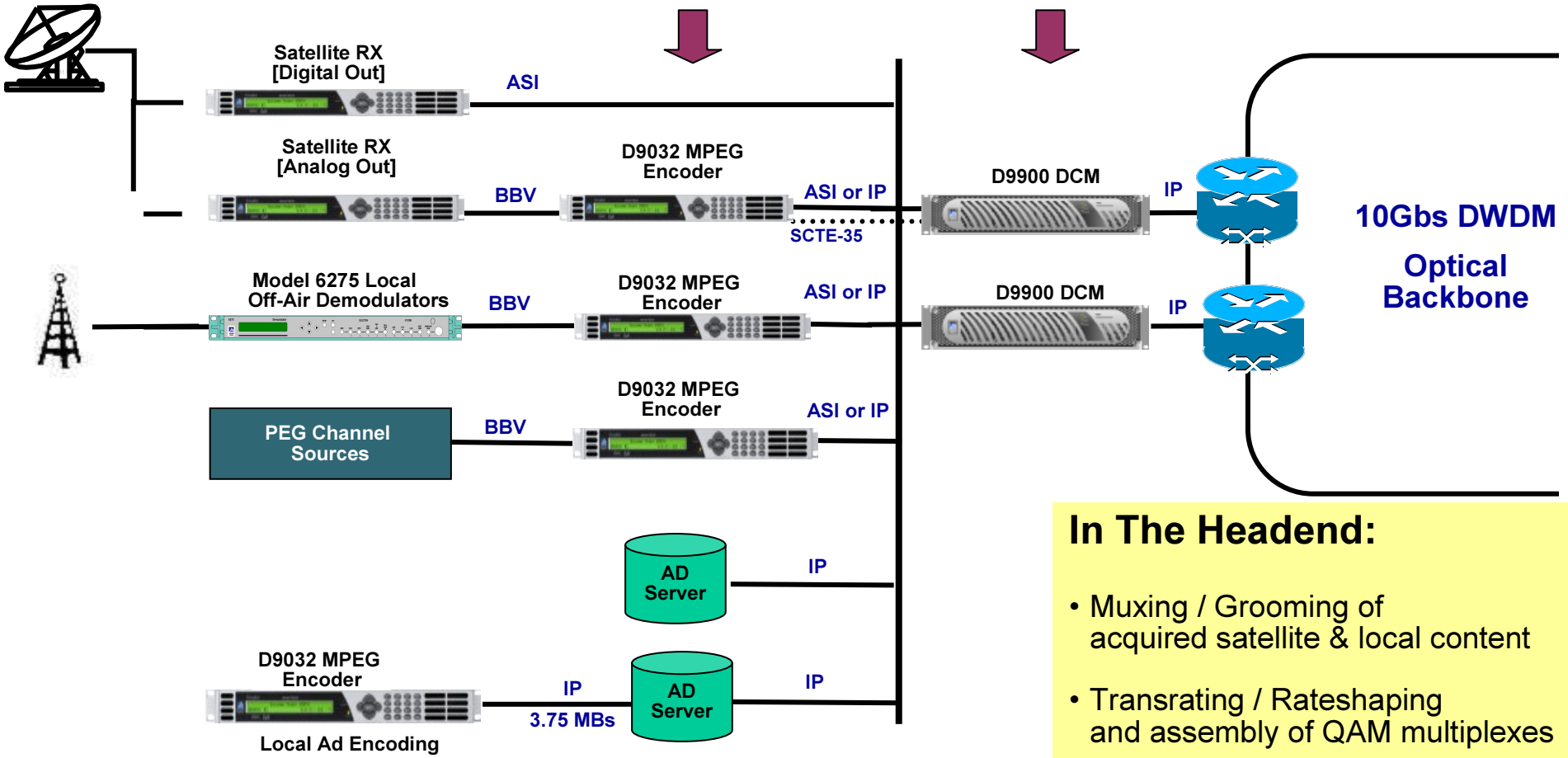
“Traditional” IPTV HE components



IP Centric HE



Master Head End - Key Components (SA example)



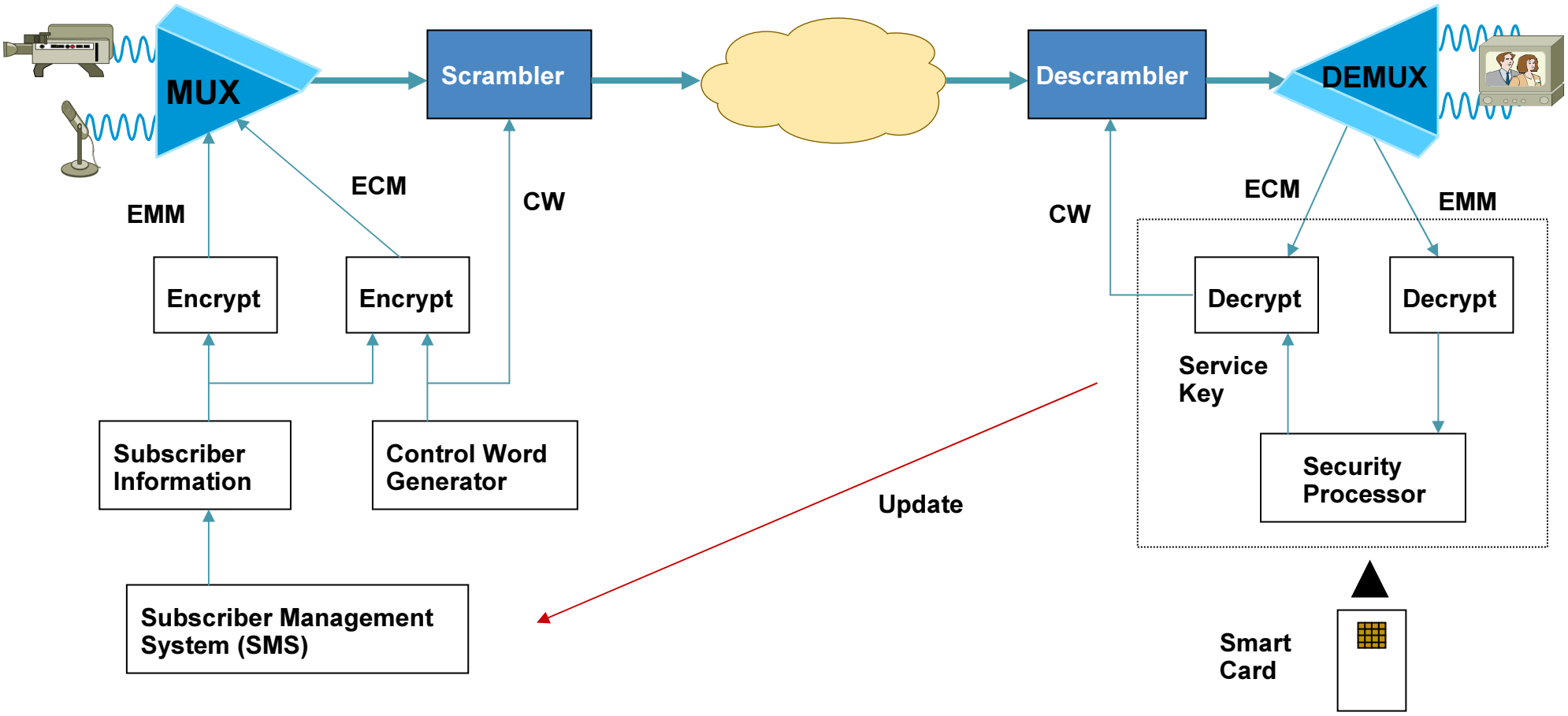
In The Headend:

- Muxing / Grooming of acquired satellite & local content
- Transrating / Rateshaping and assembly of QAM multiplexes
- Centralized DPI
- Switched Digital Video

Service/Content Protection: CAS and DRM

- **Conditional Access System (CAS) comes from Broadcasting world**
  - **Focus on protecting the service against theft**
  - **Protects content at the transport level**
  - **Moving from H/W based (SmartCard) to S/W based**
- **Digital Rights Management (DRM) comes from IT world**
  - **Focus on the contractual rights between the content owner and the consumer**
  - **When and how many times the content may be accessed and eventually recorded**
  - **Protects content at the application level**

Conditional Access. DVB-CSA



Middleware

- Service Discovery (EPG) and Selection
- Interface to Subscriber Management and Billing
- Abstraction layer that hides the access network to the STB
- Portal
  - Tandberg TV, Minerva
- Commercial
  - Myrio, Orca, NDS, Envivio, Motorola
- Home made
  - Fastweb, Imagenio (now Lucent)
- Vertical
  - Microsoft TV

Set-Top-Box

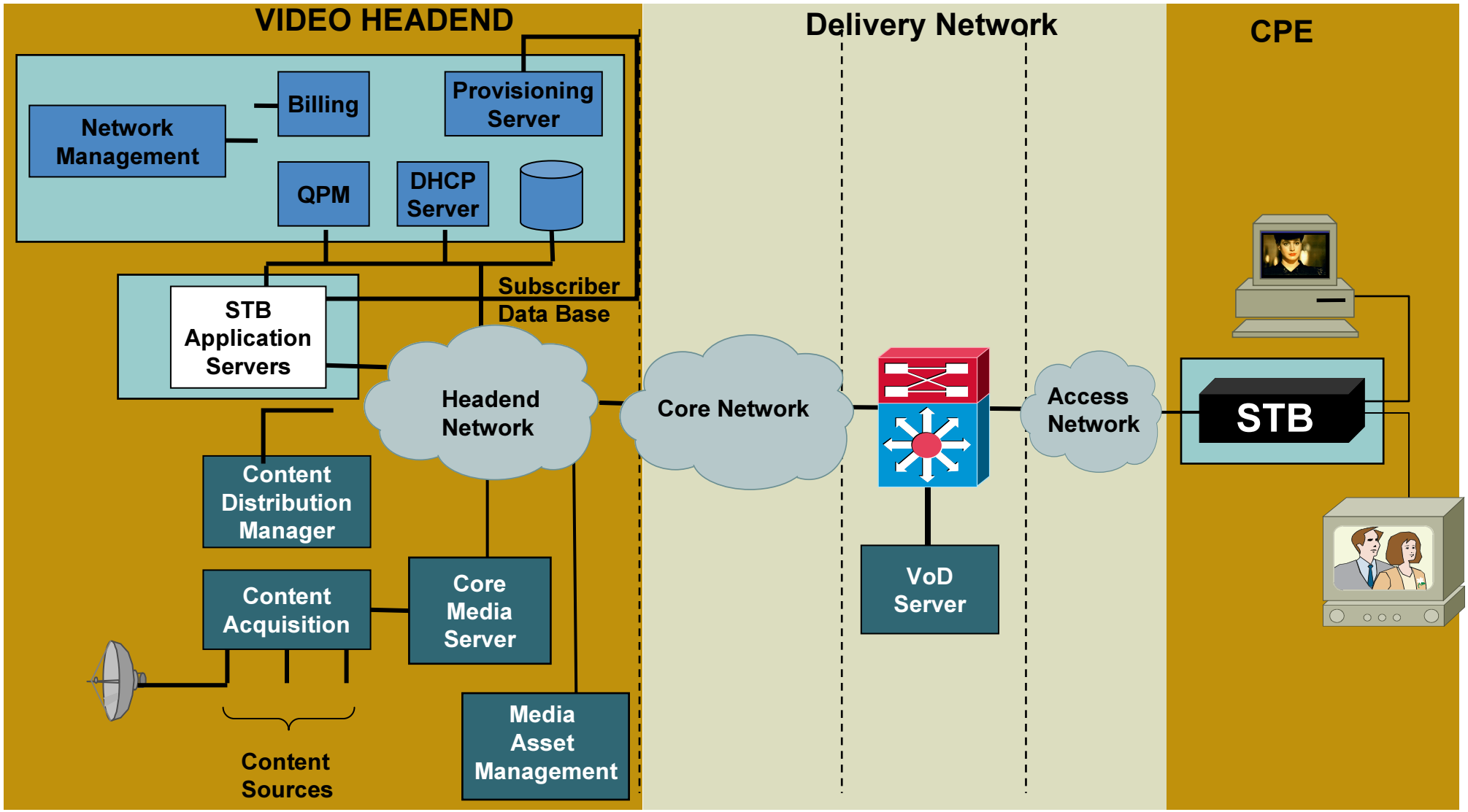
- Linux based (SA, Amino, Kreatel)
  - Portal Middleware
  - Commercial (Orca, Myrio)
  - Home made (Telefonica, Fastweb)
- MSTV (SA, Kiss)
- Recent trends to support
  - HD
  - AVC decoders
- Still no general support for
  - IGMPv3 with the SSM API in the STB

STB operation

TS → Look at the PCR

- **Transport buffer**
  - De-jitter and PCR synchronization (+/- 100 msec)
  - Look at the PID → PES
- **De-multiplex buffer**
  - Separate audio, video ES and other program components
  - Look at the DTS/PTS in the PES → ES
- **Decoding buffer**
  - Store reference frames
  - Reorder and decode frames
  - Display frames → TV

Video Servers



Publishing content

- **Asset = Content + Metadata describing the content**
- **Description uses XML language**
- **CableLabs specifications available**

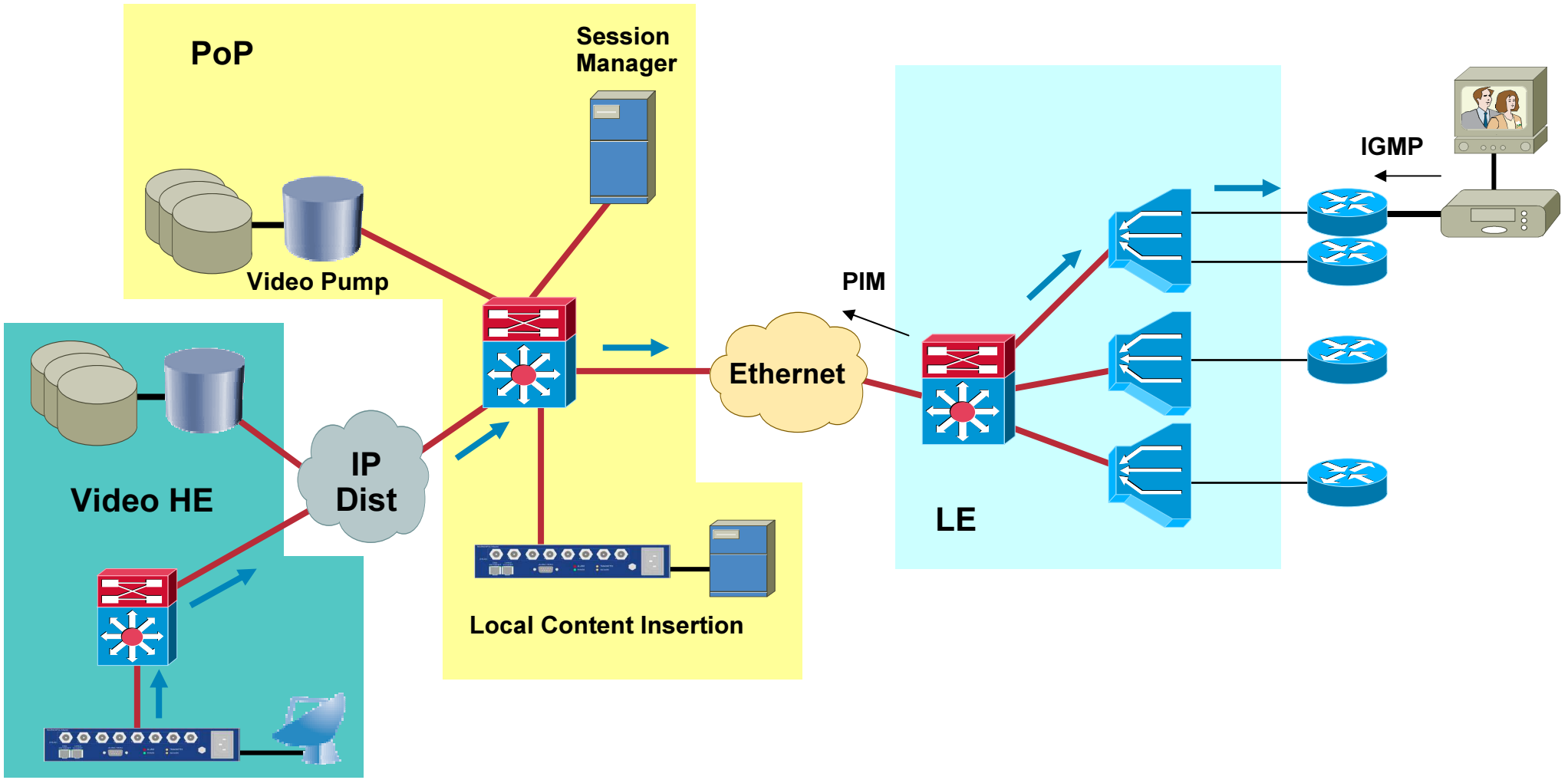
- **Building and associating Metadata with content**
- **Linking Metadata with EPG**
- **Asset Management**
  - **Pushing titles (media + metadata) to the edge servers (CDN)**
  - **Making titles available/unavailable (publishing)**

IPTV services

IPTV Broadcast

- Distributed using IP Multicast
- Dimensioned according to bitrate (CBR)
- Cable
  - Distribution only: From CHE to RHEs
  - Multiple programs per multicast group/channel. DVB SI for EPG
- DSL
  - End to end delivery: From Video HE to IP STB
  - Single program per multicast group/channel. Basic PSI (PAT/PMT)
  - Service selection with IGMP
- ETTH/FTTH
  - Less bandwidth constraints for VBR and HD
  - Multiple concurrent streams

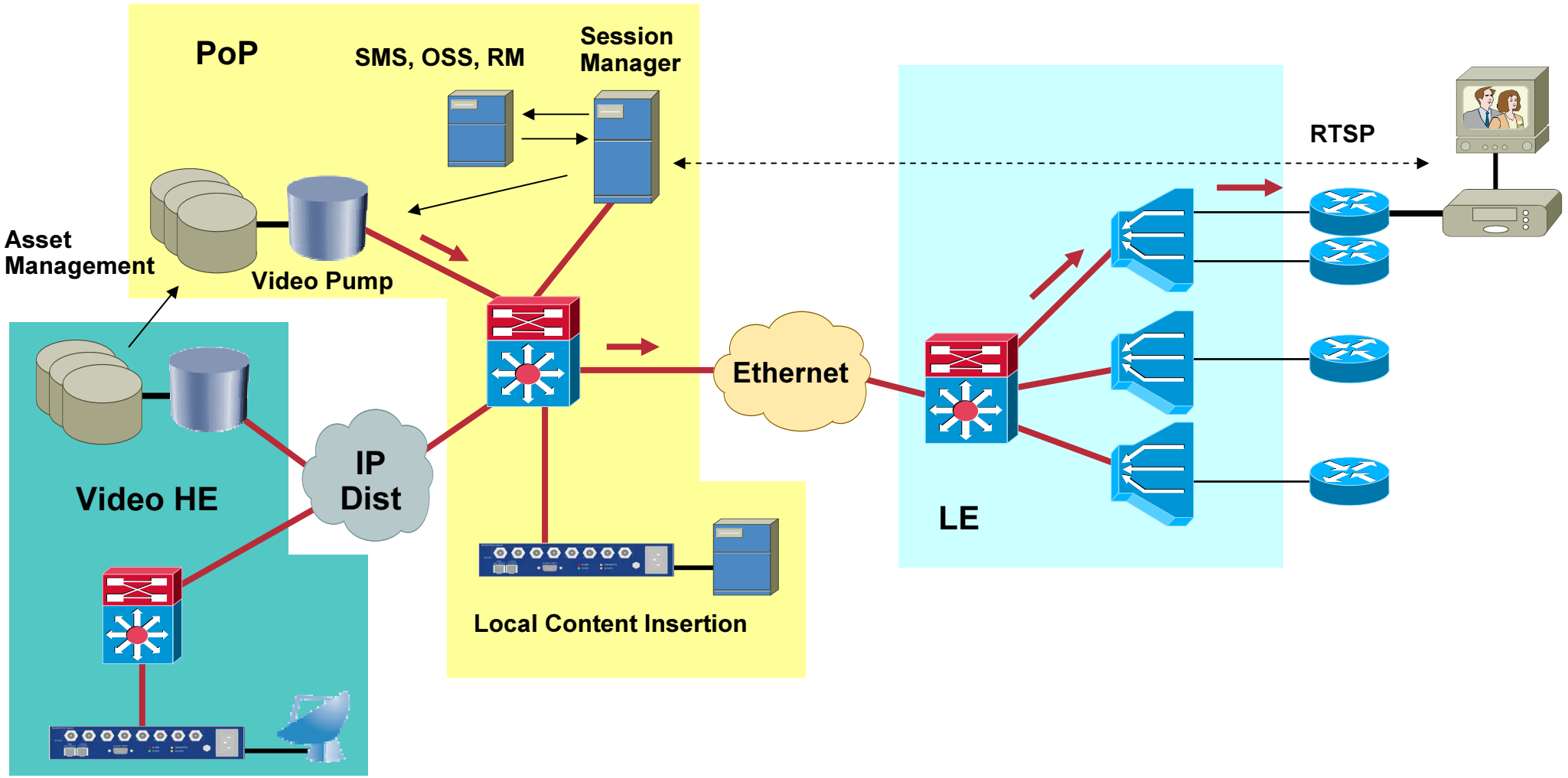
TV Broadcast in ADSL



IPTV Video on Demand

- Distributed using IP Unicast
- Dimensioned according to peak-hour concurrency
- Cable
  - 10GE RAN Distribution: From RHE to Hubs
  - Uses Out-of-band channels or DOCSIS for signaling
- DSL
  - End to end delivery: From City PoP to IP STB or distributed
  - Service selection with HTTP or RTSP
- ETTH/FTTH
  - No constraints for VBR and HD
  - Multiple concurrent streams

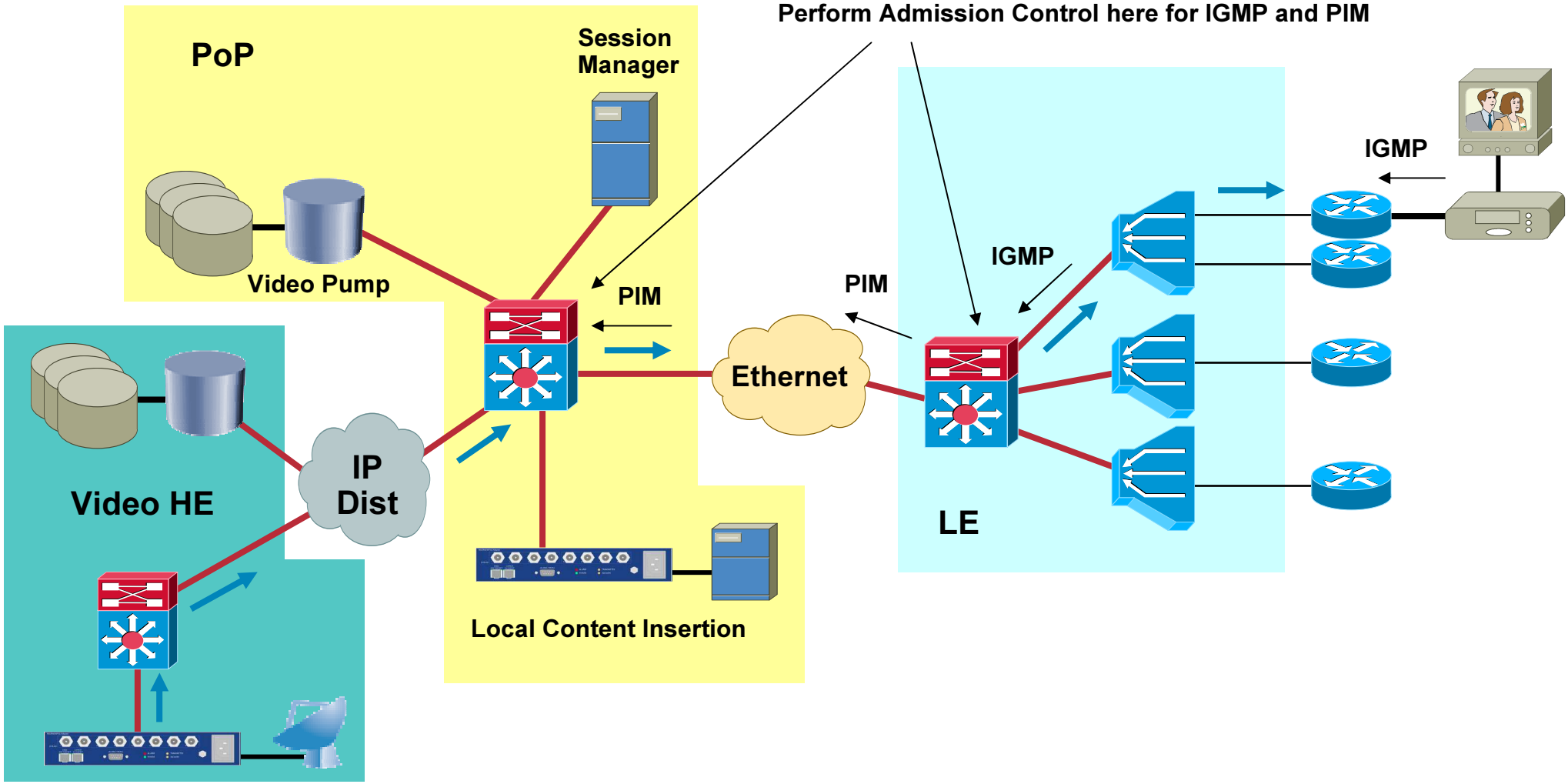
Video on demand



IPTV Switched Digital Video

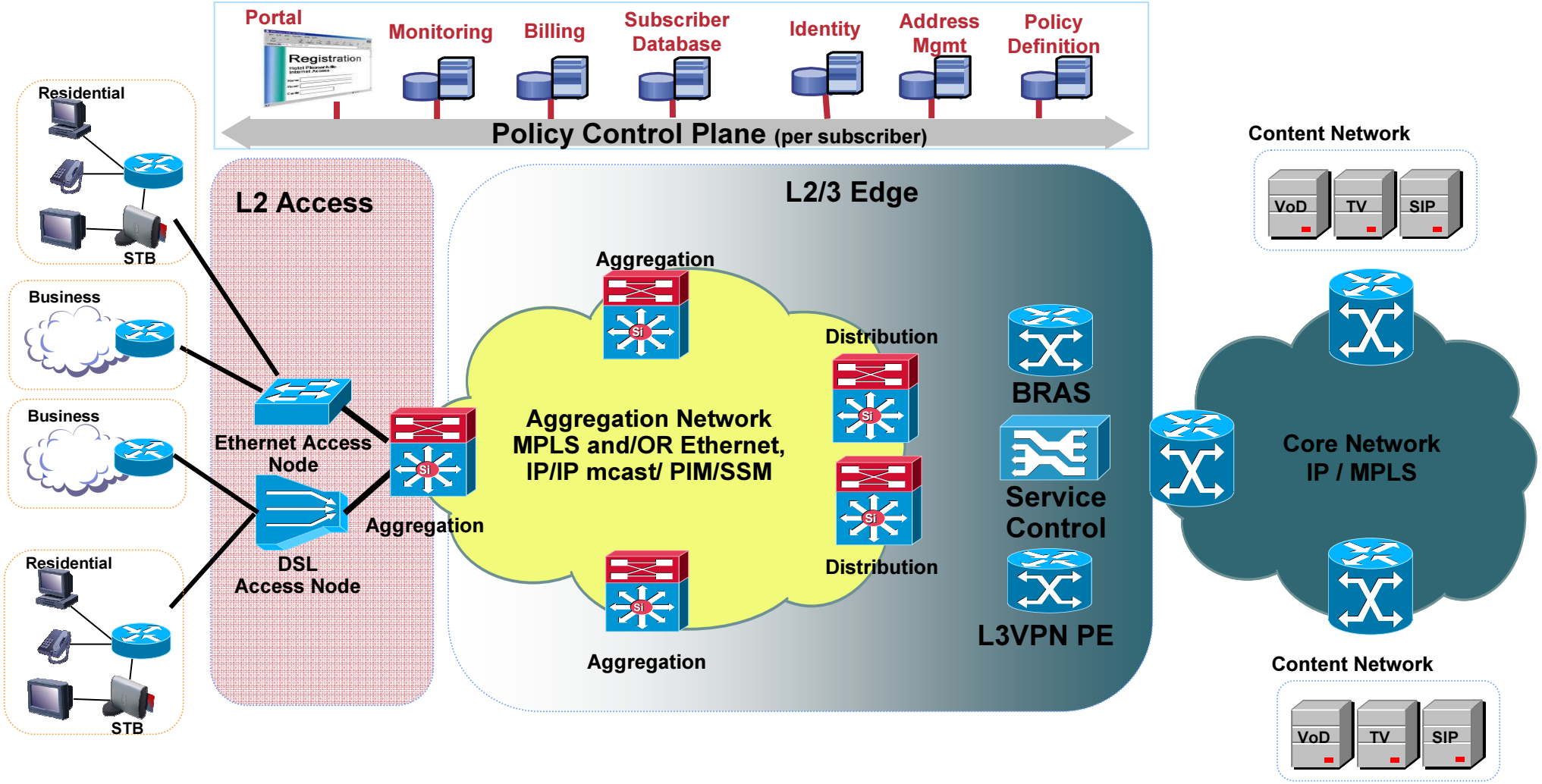
- Distributed using IP Multicast
- Service is oversubscribed for a certain group of programs
- Admission Control required
- Cable
  - Optimization of HFC bandwidth on a per Service Group basis
  - Only watched programs are sent
  - Service discovery through EPG
  - Admission Control based on available edge-QAM resources (RM)
- DSL/ETTH
  - Oversubscribed broadcast service
  - Admission Control based on IGMP (LHR) and/or PIM (aggregation points)

Switched Digital Video



Video over Broadband architecture

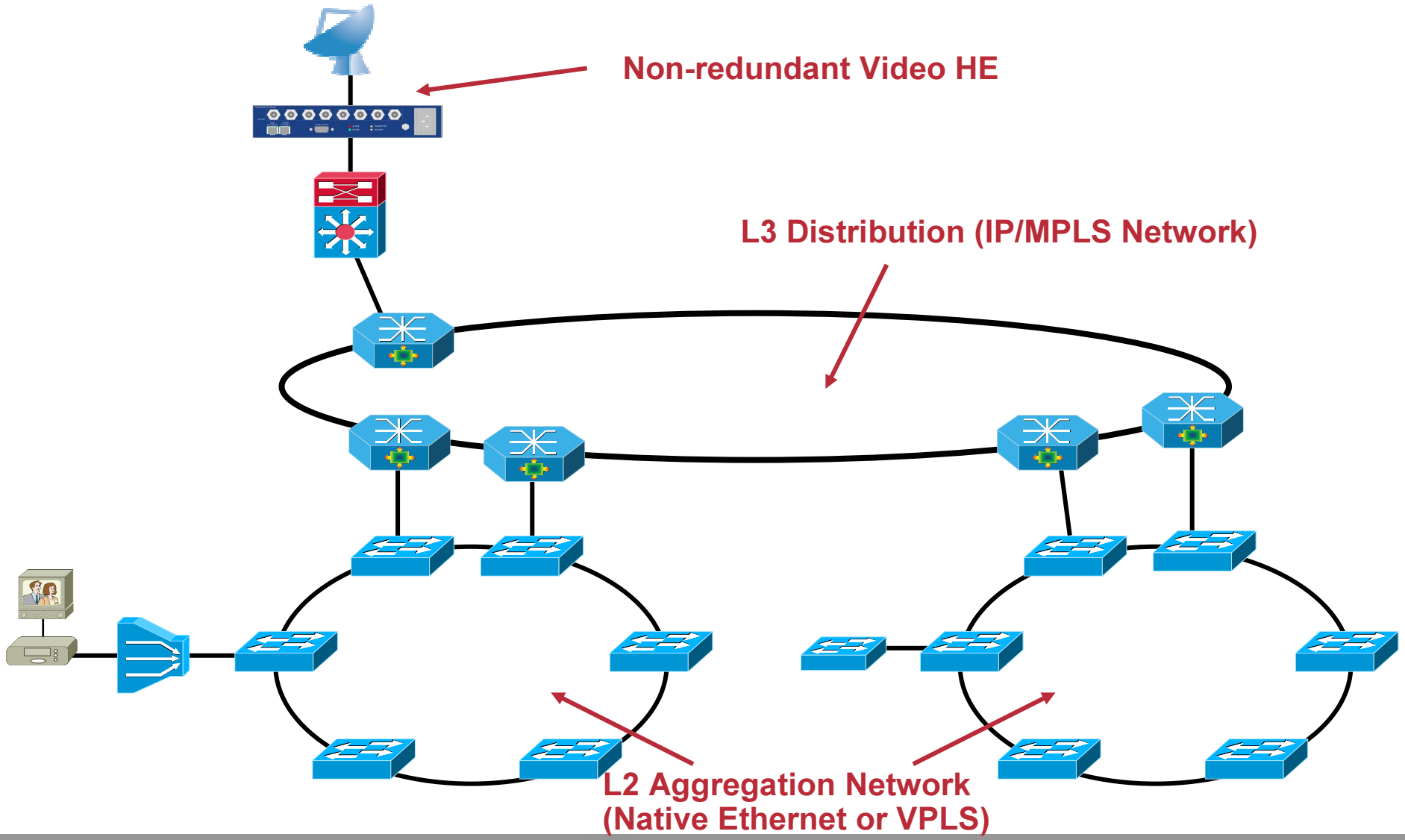
Next Generation Broadband Architecture



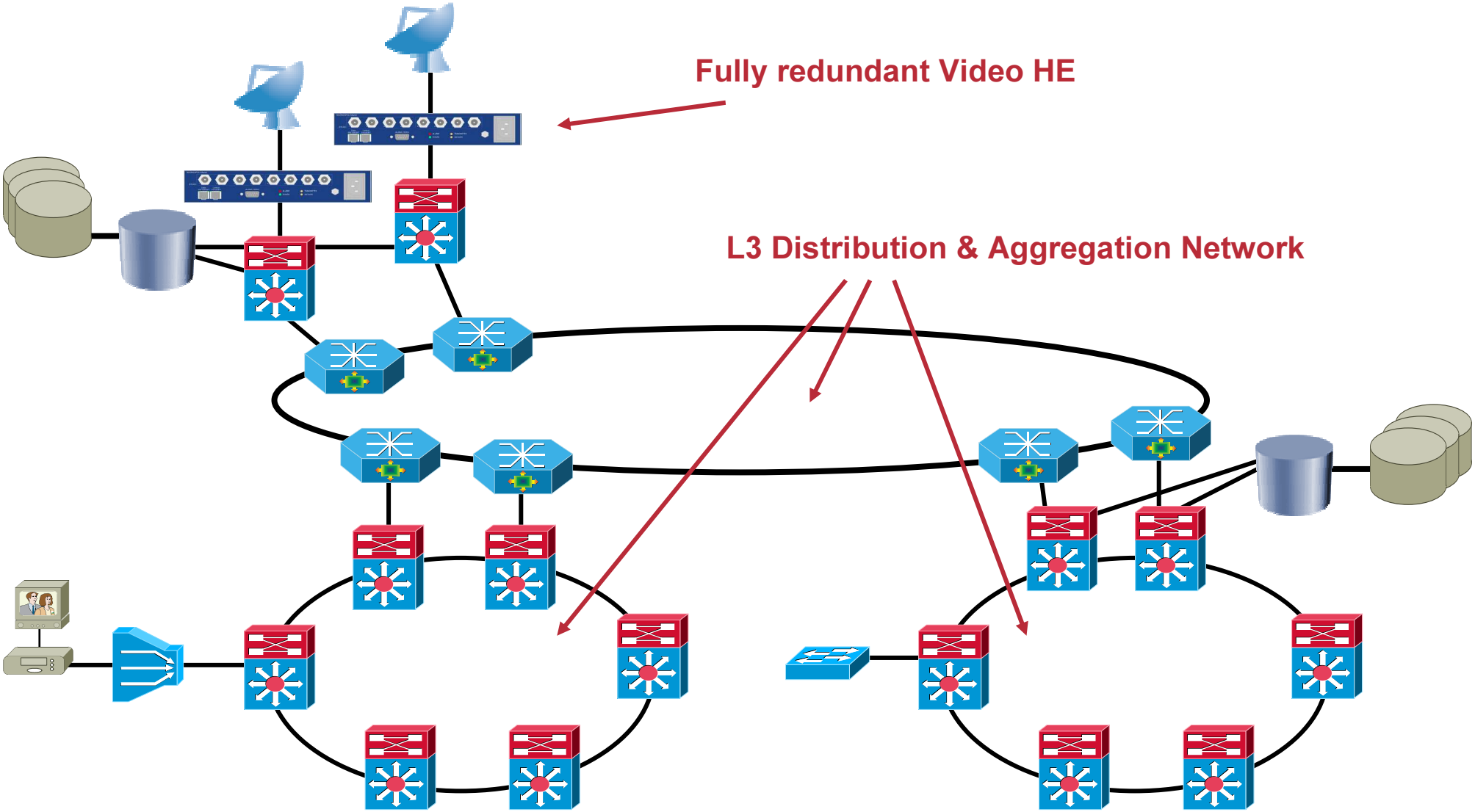
IPTV Architecture Highlights

- Per-service optimized infrastructure
  - Service-oriented vs. Transmission-oriented
  - L3 preferred to L2 for IPTV
- Distributed injection point model
  - Distributed vs. centralized
  - Video is an Application Service
  - VoD and Broadcast separated from B-RAS
- Homogeneous QoS model
- PIM SSM for multicast
- CAC for unicast and multicast
- Service separation/Per-service topologies
- Asymmetric networking

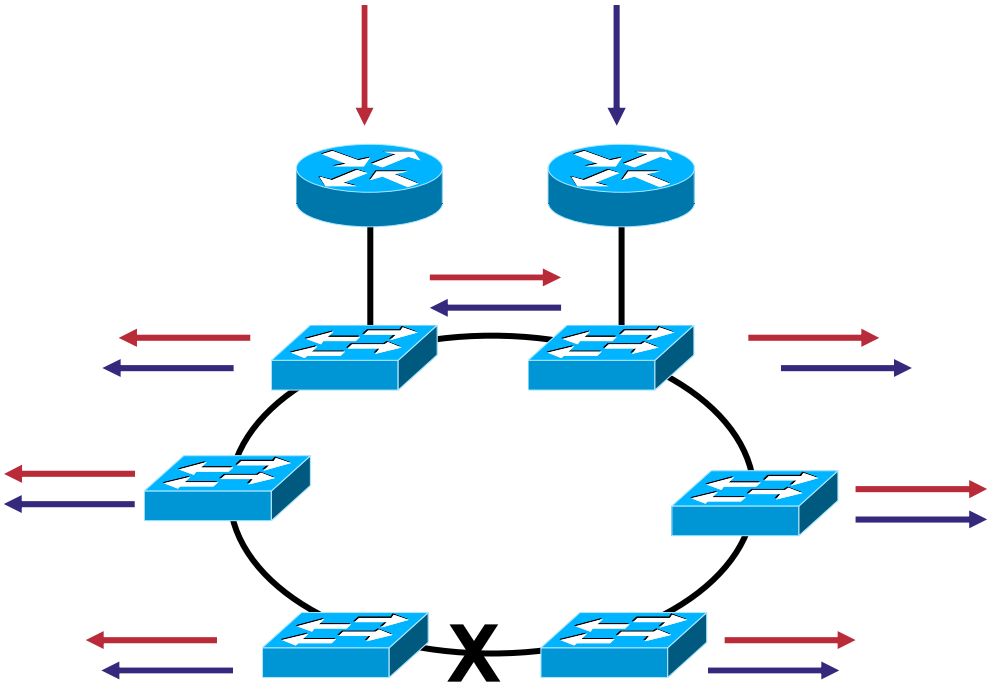
L2 IPTV Architecture model



Recommended IPTV Architecture

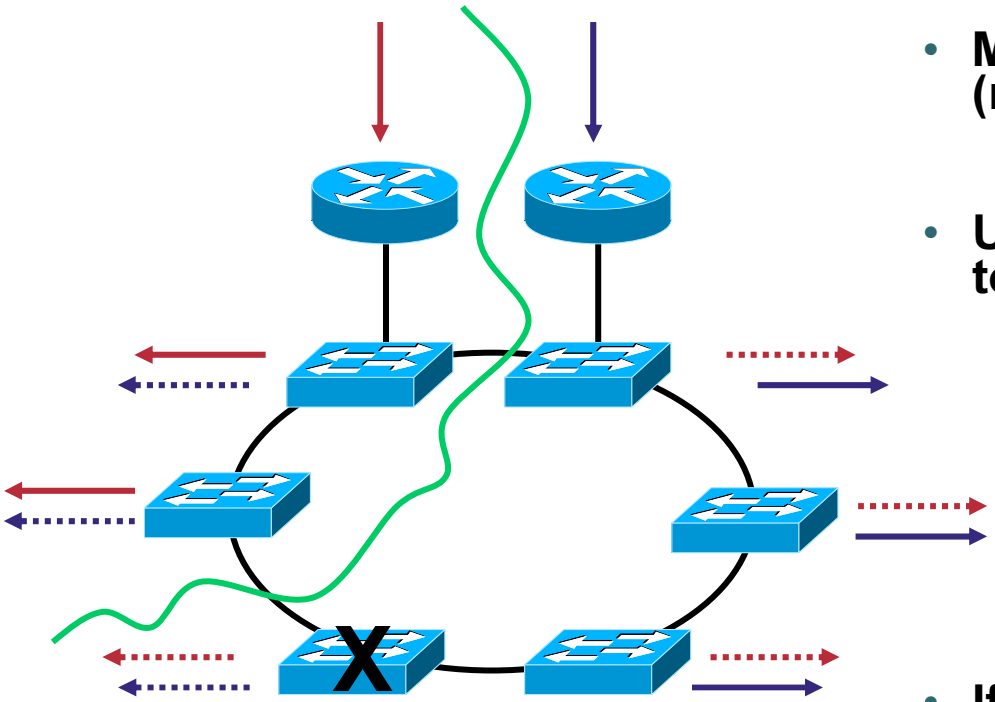


L2 forwarding in rings (native Ethernet)



- **Forwarding mechanism is constrained flooding**
  - **By MAC learning (Unicast)**
  - **By IGMP or PIM Snooping (Multicast)**
- **Link and node failures imply topology change**
  - **Convergence depends on STP**
- **L2 reverts to flooding when MAC forwarding tables are flushed**
  - **STP topology changes or MAC aging**
  - **Video is uni-dir. MACs are not learnt fast!**
  - **Requires blocking of unknown unicast traffic to work**

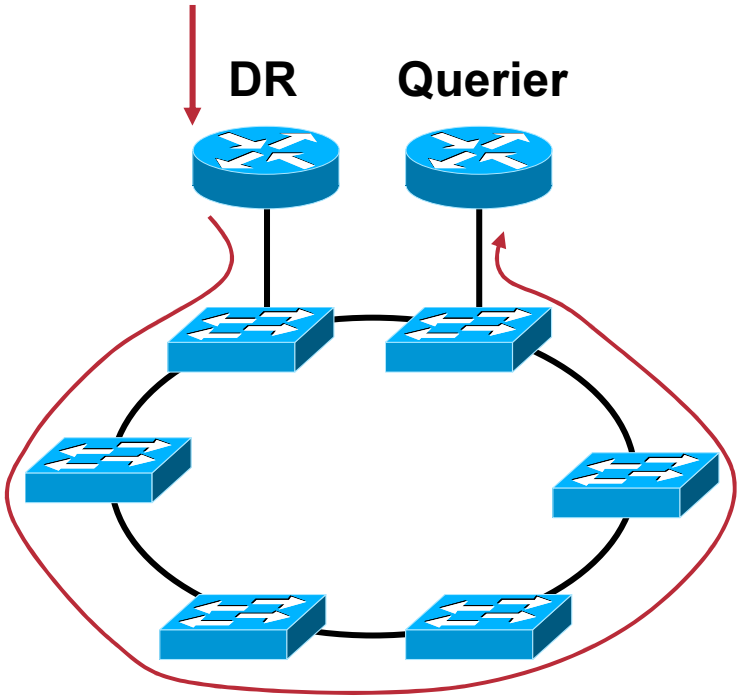
L2 forwarding in rings (MPLS)



- Multicast must be deployed in daisy-chain (ring-VPLS) topologies
  - Totally inefficient otherwise
- Unicast cannot be deployed in ring-VPLS topologies
  - Unless STP is used
  - Otherwise the ring can be split into two segments: Discontiguous subnet problem
  - Each segment will only receive what its attached router was injecting before
  - Massive traffic blackout
- If STP is run,
  - L2 reverts to flooding when MAC forwarding tables are flushed
    - Needs unknown-unicast-blocking!
    - Not recommended!

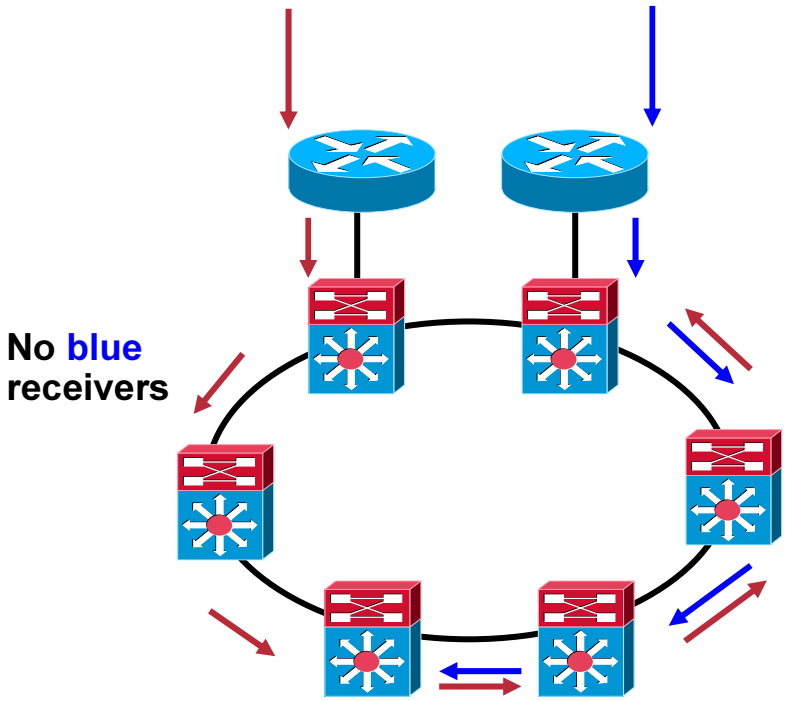
Incongruent unicast and multicast topologies!
Two different VPLS instances needed!

L2 multicast forwarding



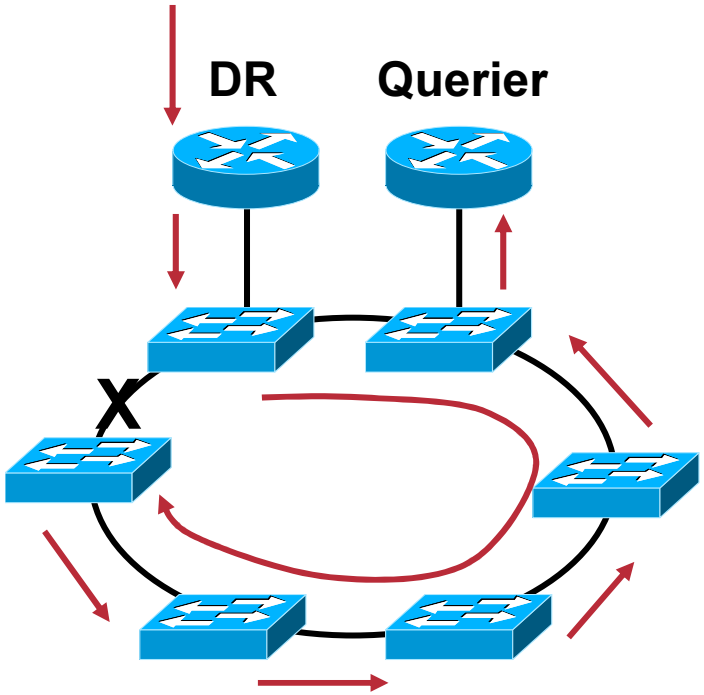
- Single injection point (DR)
  - DR failure affects all subscribers
- All content sent to Non-DR, where it is dropped
  - Non-optimal forwarding (broadcast)
- Pervasive user (IGMP) signaling in Aggregation Network
  - Security and scalability issues
  - Many subscribers depend on single Querier
- Inconsistent with L3 VoD Service
  - CAC requires L3 (segmented address allocation in subnets)

L3 forwarding



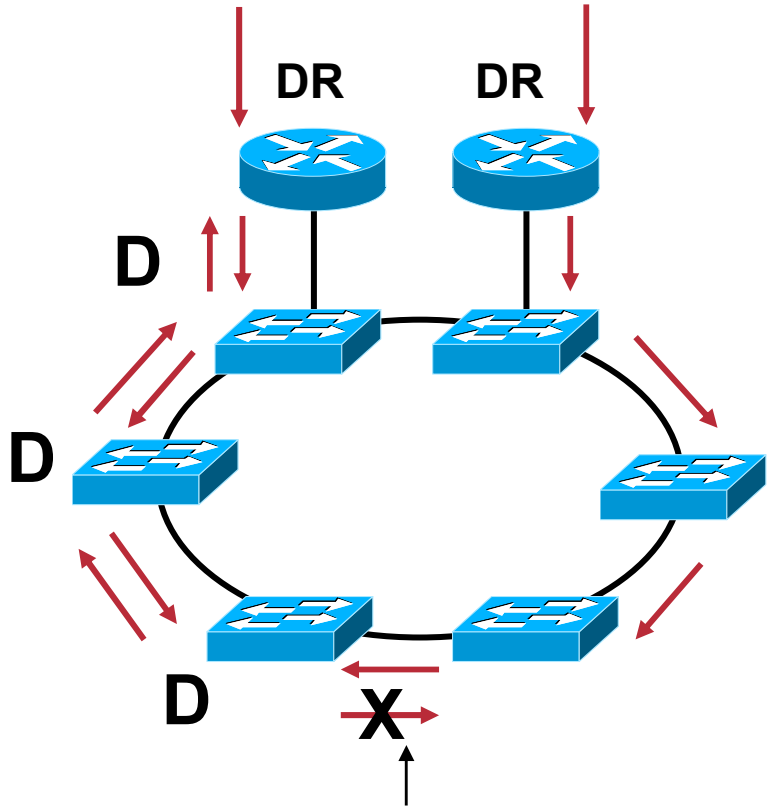
- **Multiple injection points**
  - **Failures do not affect all subscribers**
- **Content sent only to where is requested**
  - **Optimal forwarding (multicast)**
- **No user signaling in aggregation**
  - **More secure and scalable**
- **Fully consistent with L3 VoD**

Link protection



- LSP is protected by FRR
  - Carried pseudo-wires are rerouted
  - In theory, less than 50 msec outage
- However, topology is fixed
  - No L2 redundancy without STP
- Node failures must be dealt with at higher layers
  - DR failover is at L3 (PIM hellos)
  - Must be tuned very aggressively (many hellos) for fast failover
  - Missing hellos due to congestion can result in duplicates
- Same for DR and Querier failures
  - No fast messages in IGMP
  - Querier failover takes 20-120 seconds

PIM Fast Hellos and congested links



Congested link dropping control plane packets (PIM Hellos) in one direction

- Duplicates due to missing hellos
  - In case of congestion
  - Or unidirectional link
- Fast hellos increase the probability of this scenario
- Hellos should be used as L3 keepalives only, not for solving L2 deficiencies
- Hello interval should be long
  - Experience with other protocols (OSPF)

Impact of network failures

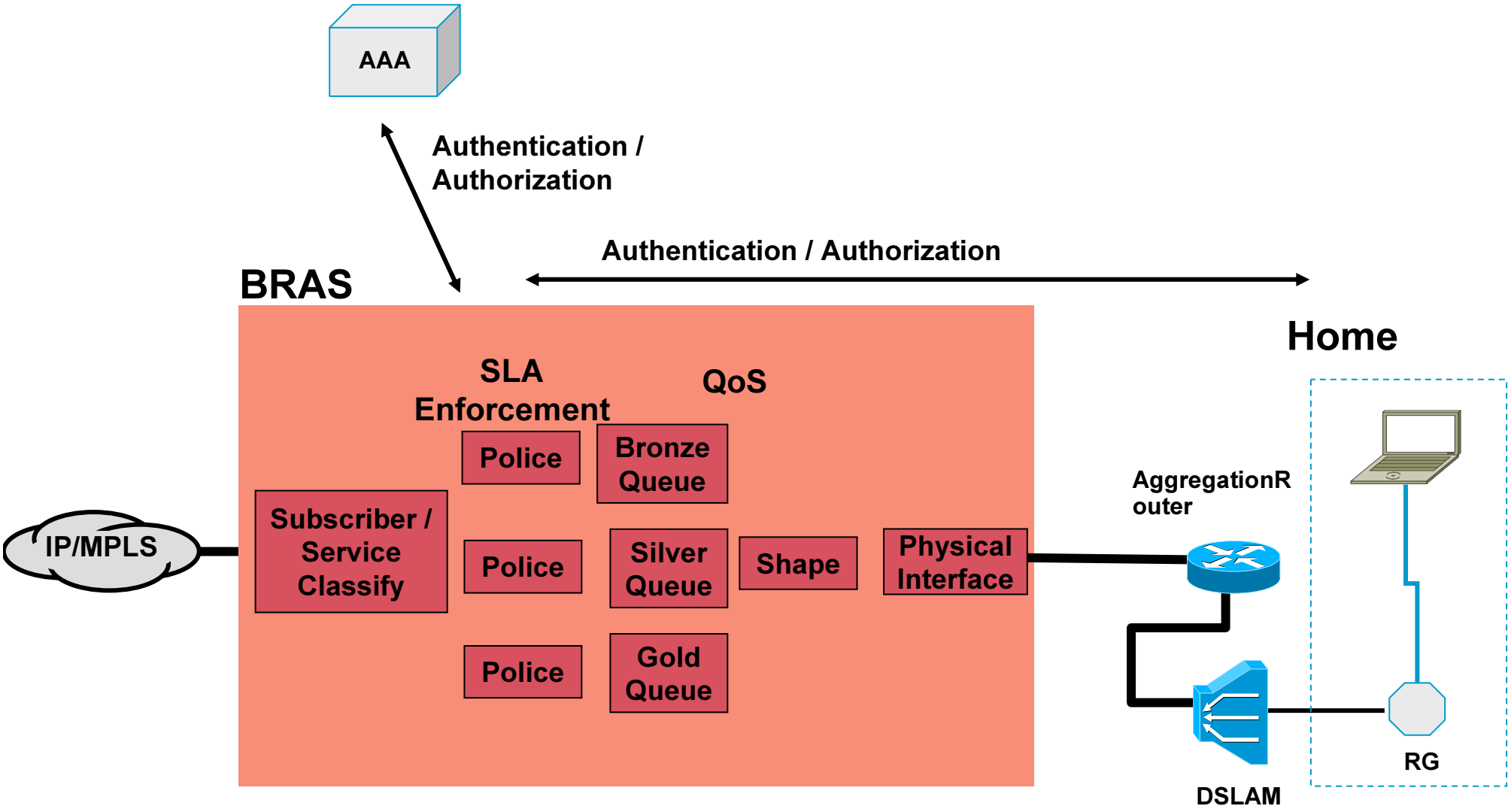
	L2 VPLS		L3 IP Multicast	
Link failure	MPLS-TE FRR	< 50 ms	Mcast fast convergence	< 300 ms
Node failure (redundant topology)	STP	3-60 secs	Mcast fast convergence	< 300 ms
Node failure (non-redundant, e.g. Ring-VPLS)	DR failover	3-60 secs (*)		
Source failure	DR failover	3-60 secs (*)	Mcast fast convergence	< 300 ms

(*) Need to activate L3 PIM Fast Hellos for sub-second

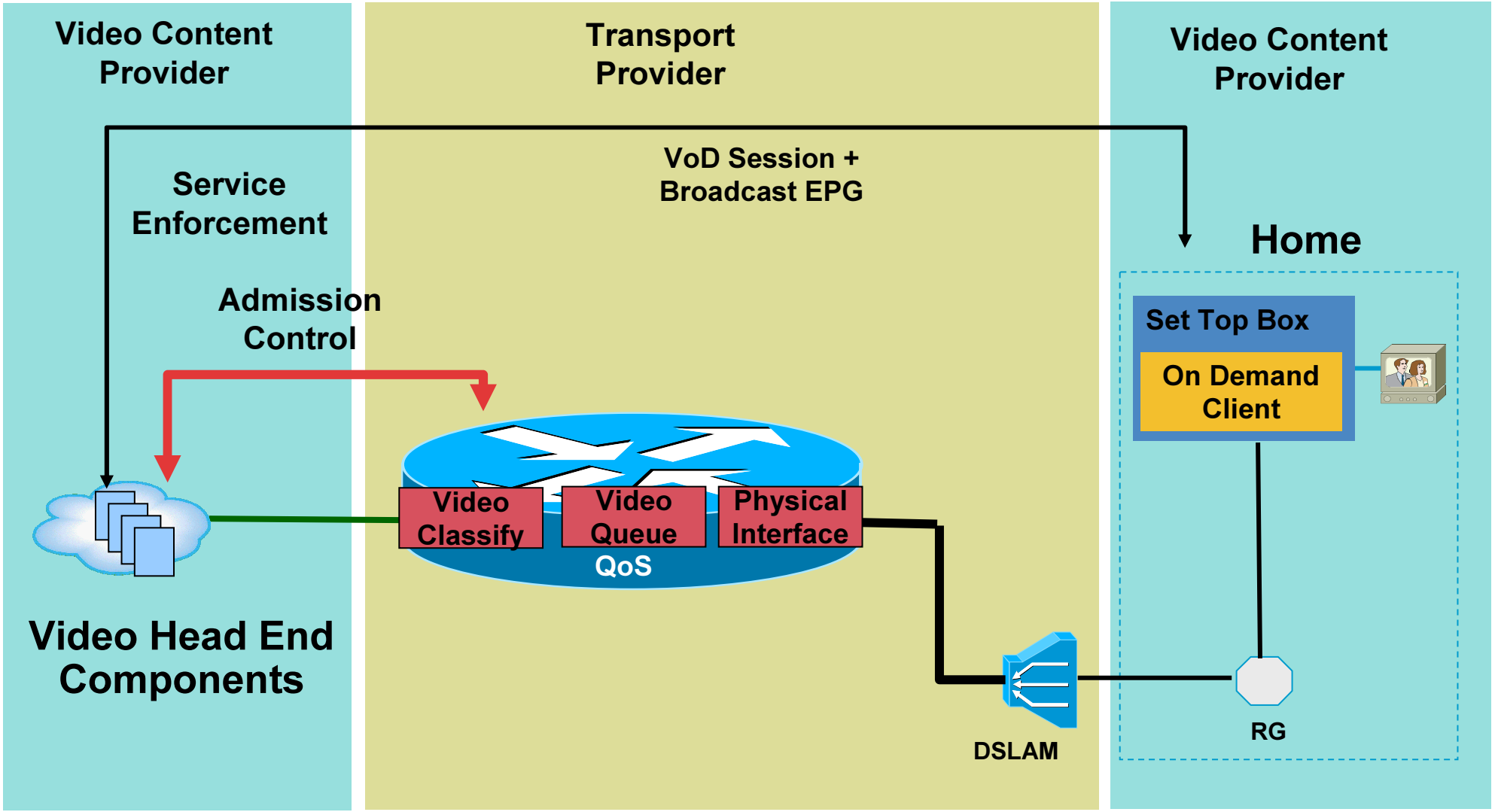
Application vs. Transport Services

Service Type	Transport Service	Managed Application Service
Service Level Agreement	Transport Parameters Bandwidth, Max Drop, Max Latency, etc.	Application SLA Video: # of Set Tops Basic vs. Premium Tier
Subscriber Authentication / Identification	Network Based PPPoE, 802.1X Per subscriber VLANs, DHCP Option 82	Application Based Video Middleware Network May Also Authenticate
SLA Enforcement	Network Based Per Subscriber Shaping / Policing	Application Based Based on App Signaling
QoS	Per Subscriber QoS Gold, Silver, Bronze Per subscriber classification, queuing	Aggregate Single Queue for Video Integration with Network CAC

Internet Access Transport



Video Application Service Transport



Multicast trees

Applications Spectrum

One to Many

Many to Many

IPTV

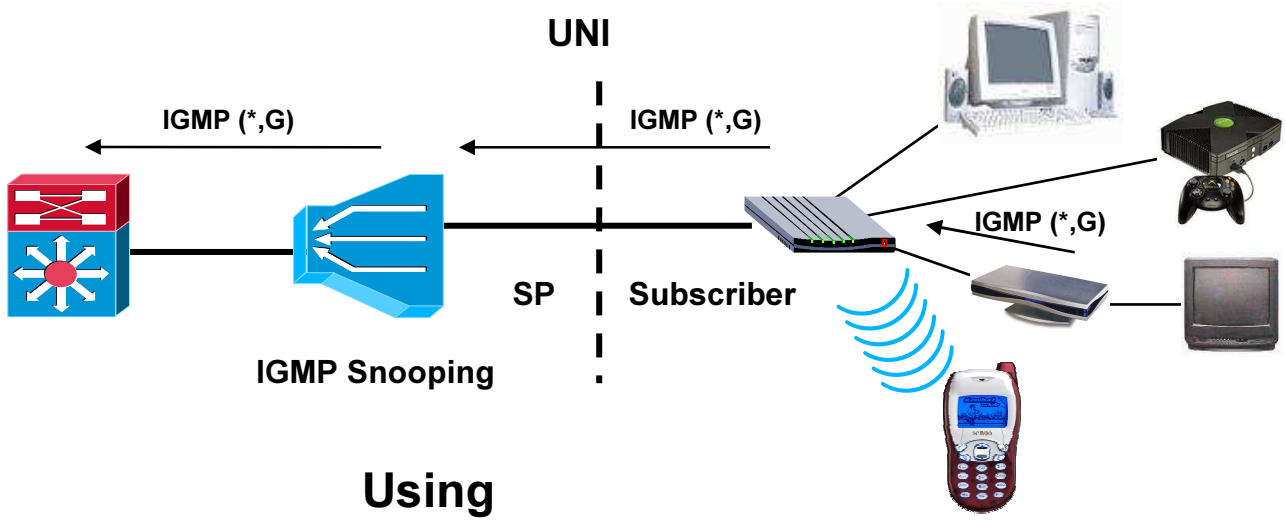
Video Conferencing

Advantages of SSM

- **Simpler**
  - Sources are known in advance
  - No RP
  - Single (Shortest Path or SPT) tree
- **More secure**
  - Only one source can send to SSM channel
  - Prevents DoS by malicious attacks or misconfigurations
- **More scalable**
  - Better use of address space

L3 tools for securing the UNI

- **Data plane**
  - **Filter upstream multicast injection**
- **Control plane**
  - **PIM**
  - **IGMP**



Prevent

PIM Adjacencies (hellos)

PIM Registers

PIM Assert election

DR election

Unauthorized SA messages

Unauthorized Access

Flooding of control messages (DoS)

Using

Neighbor filters

MSDP MD5 Authentication

Administative Boundaries

Mroute limits

Message Authorization

Access Control Lists

Filtering

Throttling

Rate limit

QoS Guidelines for Video

- **Network SLAs**

  Delay: not critical. Most applications unaffected

  Jitter: not critical. STBs can buffer 200 msec

  Packet-loss: critical. Packet loss rate < 10^{-6} (one noticeable artifact per hour of streaming @ 4Mbps)

- **Packet loss due to queue drops by bursts at aggregation points from multiple sources (also number of hops, link occupation)**
- **Size the queue depth using probability analysis so packet loss rate (e.g. 10^{-6}) is below target**
- **If drop needed, drop VoD packets first (threshold)**

Service Oversubscription

1) What needs to be oversubscribed

	Not conditioned	With Admission Control	With Per-subscriber QoS
Non oversubscribed	Broadcast Video		
Oversubscribed		Voice VoD Switched Digital Video	Internet Access

2) Capacity planning

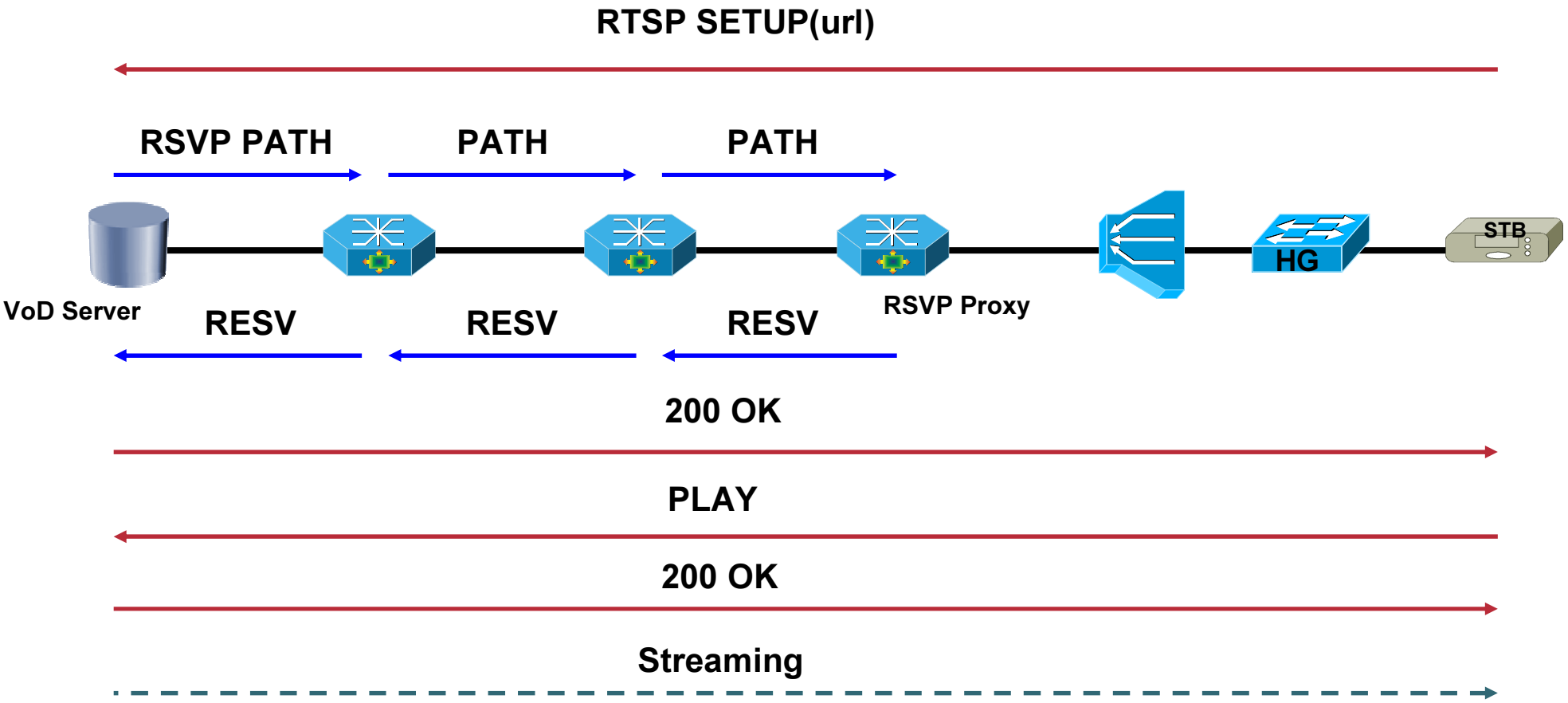
3) Define traffic classes and assign resources to them

4) Apply CAC (off-path geographical or on-path CAC)

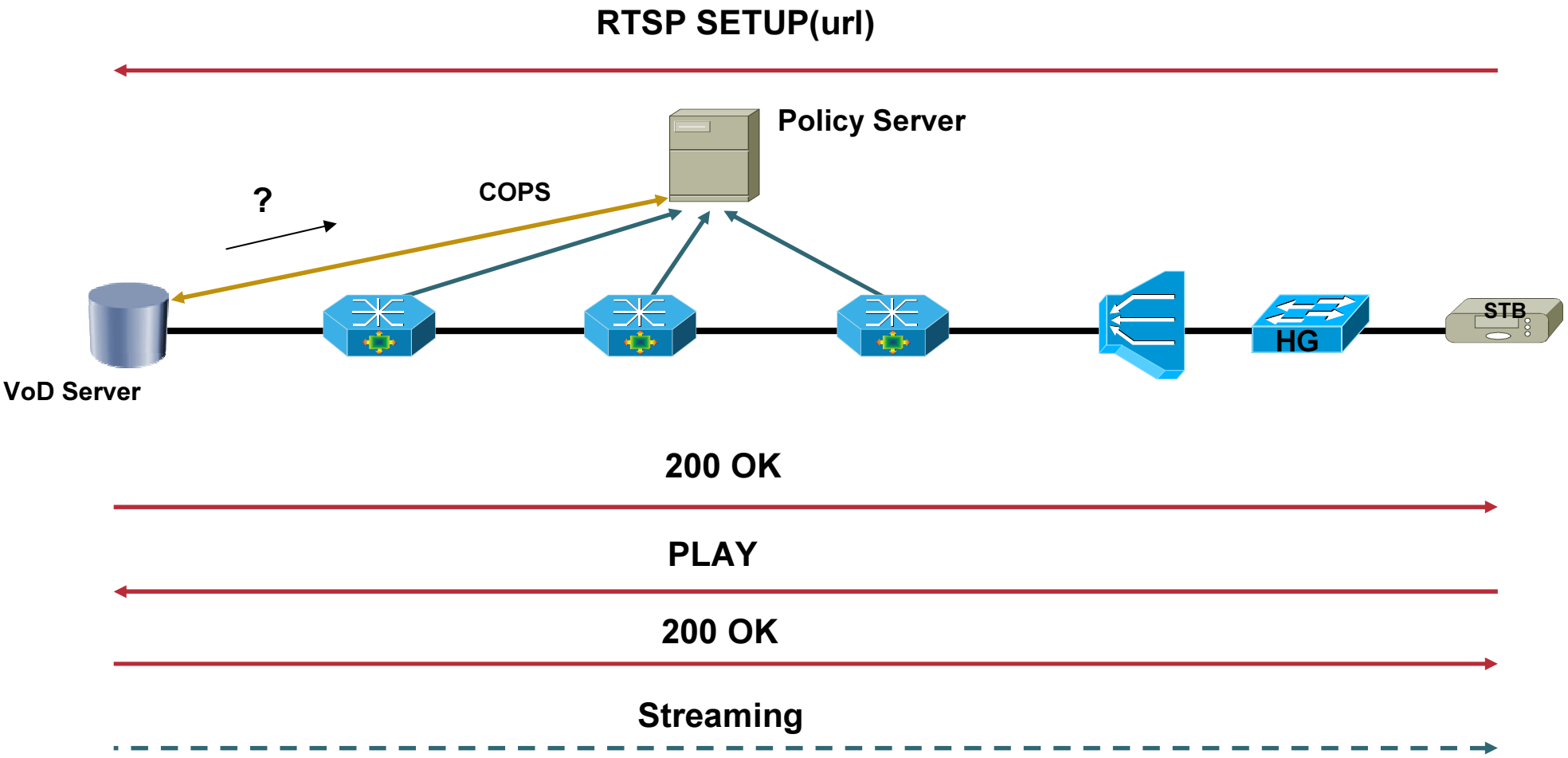
Traffic classes and network behavior

	DiffServ DSCP	Prec	Behavior	Queuing	Other
Voice and voice signaling	EF	5	Low latency, High priority	Weighted queue	Police on exceeding weight
Video Broadcast	AF41	4	Assured forwarding, very low drop	Weighted queue	
Video on Demand	AF42	1	Assured forwarding, low drop		Drop on exceeding threshold
Video Signaling	CS3	3	Non oversubscribed class	Weighted queue	Police on exceeding weight
Internet Access	BE	0	Best effort	Weighted queue	WRED

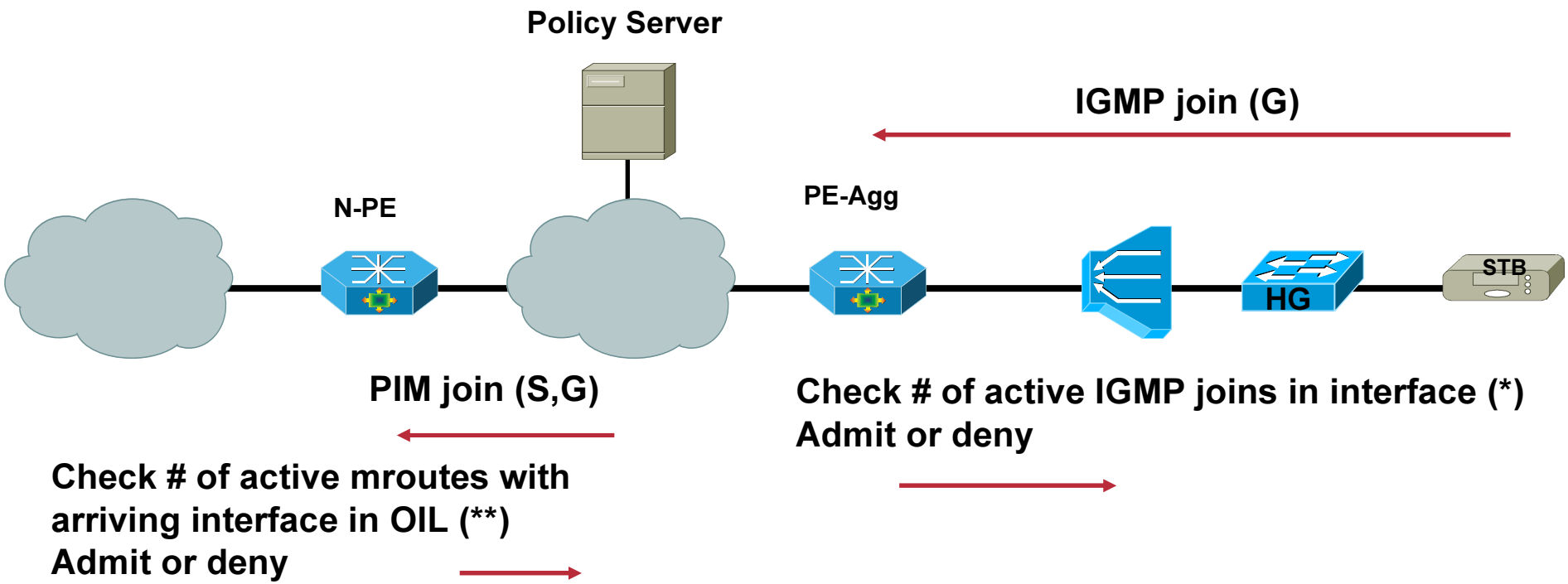
RSVP CAC for VoD



Policy Server CAC for VoD



CAC for Multicast (Switched Broadcast)

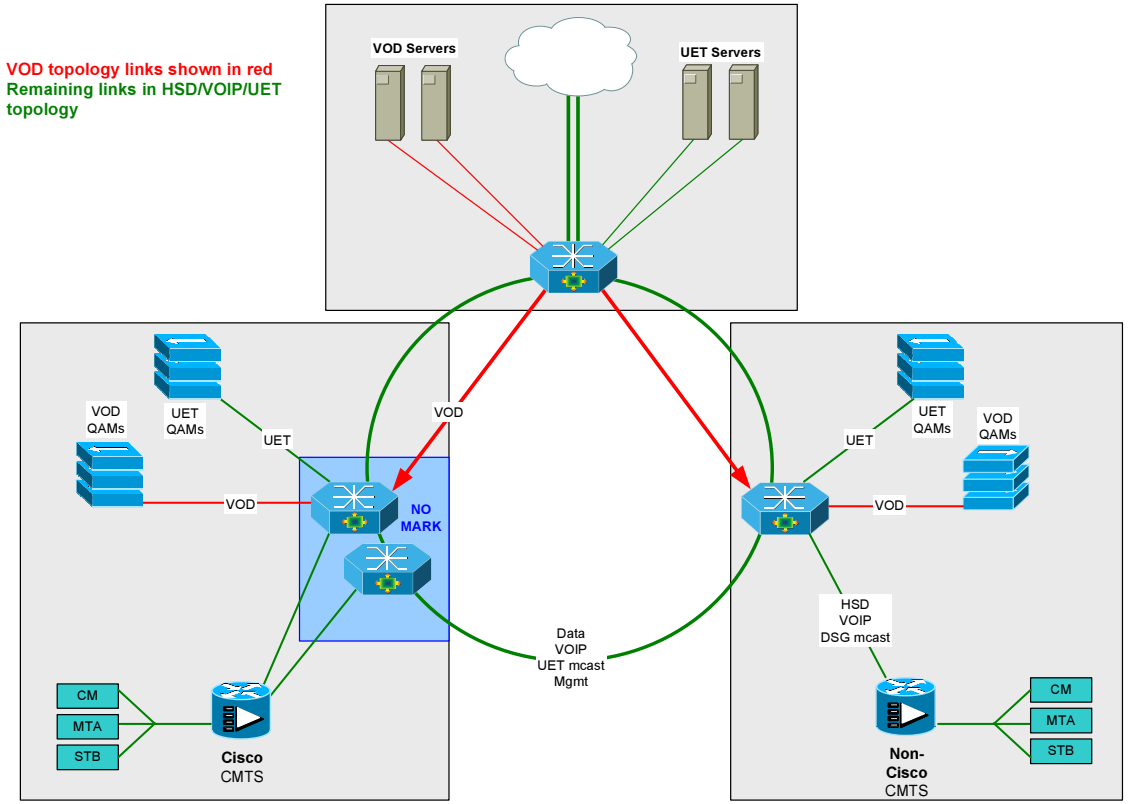



```
(*)  ip igmp limit <limit> [exclude <acl>]                12.2(15)T
(**) ip multicast limit out|rpf|connected <acl> <limit> 12.3(14)T
```

May use Policy Server in the future through Multicast AAA interface

Service Specific Routing (Multi-VRF)

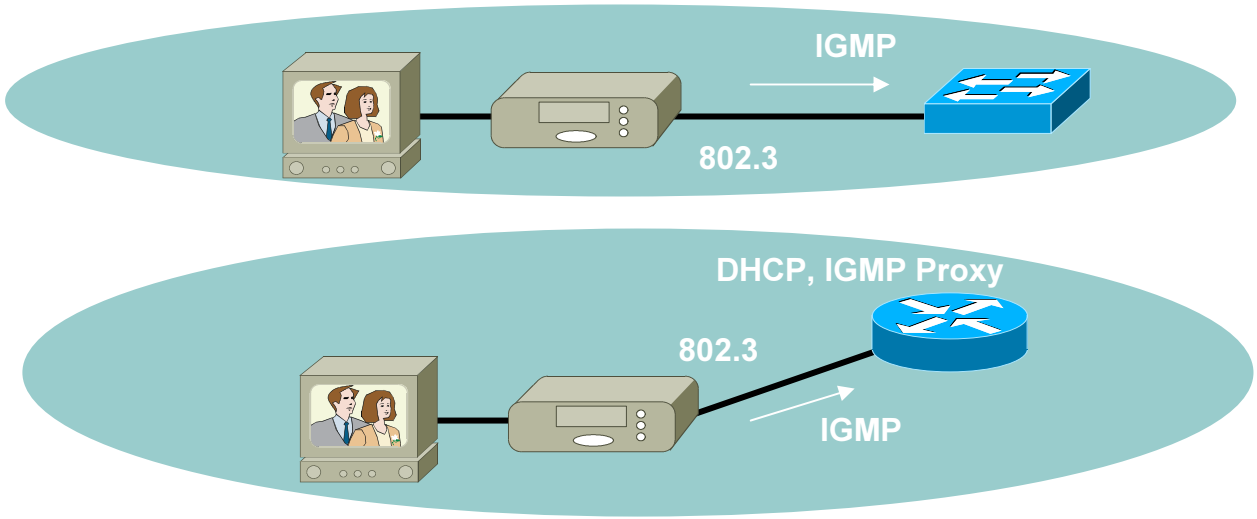
- **Multi-VRF allows for multiple routing topologies separated by interfaces**
- **Each interface (VLAN, subinterface) can belong to only *one* topology**
- **One IGP session *per topology***

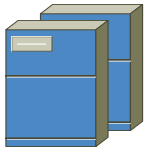


Design concepts

Home Network

- **Bridged Home Gateway option**
  - **STB obtains IP address from remote DHCP server**
- **Routed Home Gateway option**
  - **Local DHCP server with NAT for VoD or Remote DHCP server with DHCP relay**
  - **May use DHCP option 121 for route to L3 exit point towards servers**
  - **Always use IGMP Proxy on for Multicast traffic**
- **Video upstream traffic sent in non-oversubscribed class**

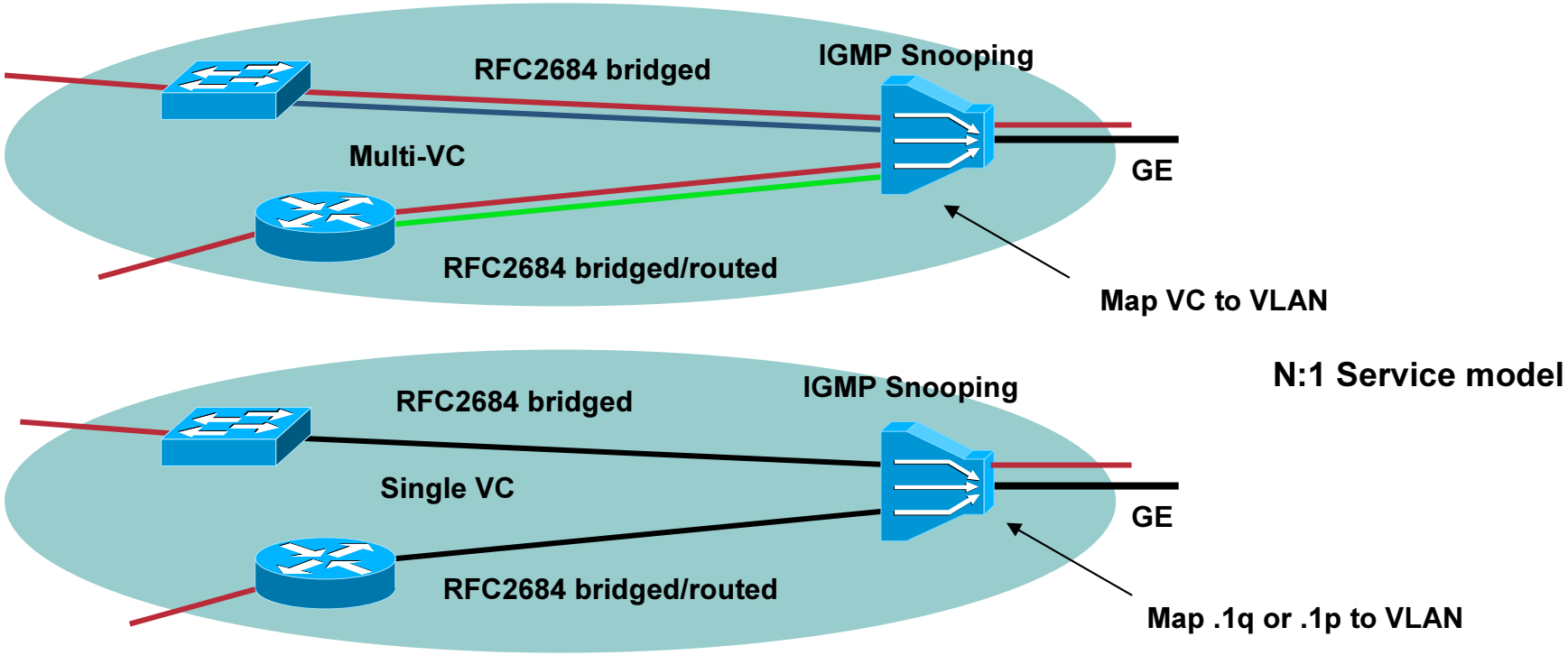




STB Boot Server
DHCP Server

Access Network

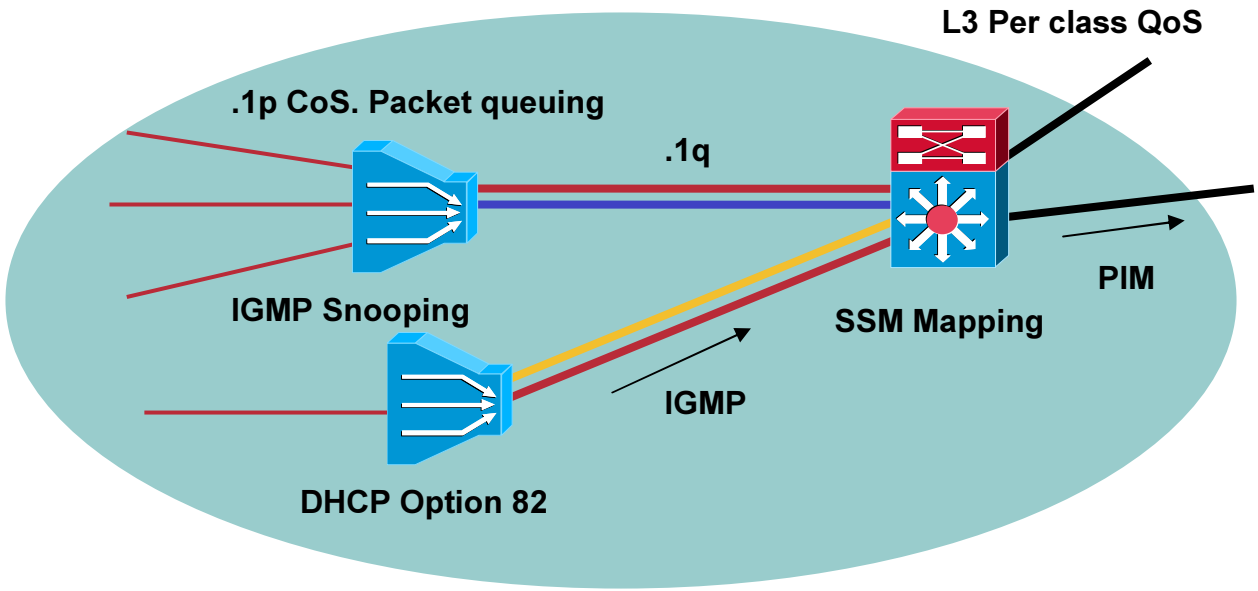
- N:1 Service model (DSL Forum TR-101)
  - Multi-VC mode with ATM CoS
  - Single-VC mode with .1p CoS, service separation with .1q (trunk UNI)
- Bridge (RFC2684-bridged) or Router (2684-bridged/routed) HG
- Native Ethernet encapsulation with VDSL



N:1 Service model

Access Aggregation

- L3 PE-Agg as distributed service injection point for VoD & Broadcast
- Layer 3 QoS (per class) mapped into .1p CoS on .1q trunks
- SSM Mapping (static or DNS) function at PE-Agg
- IGMP Snooping and Report Suppression at the DSLAM
- DHCP option 82 at DSLAM

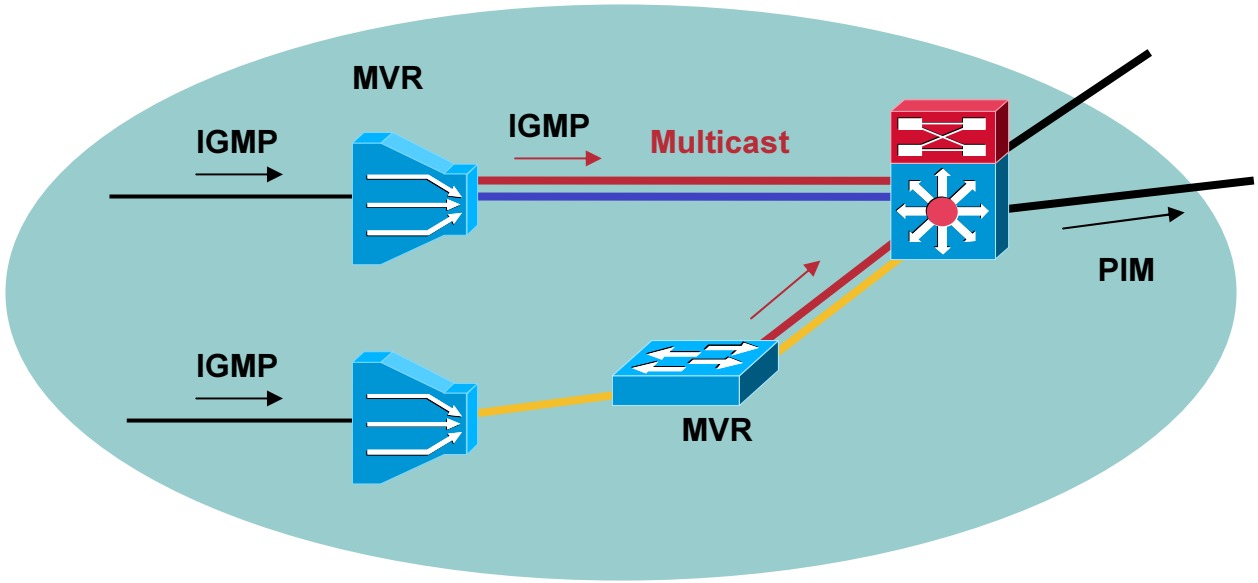


Merging Multicast VLAN

- Alternative UNI model with no service separation (e.g. HG NAT router)
- Service separation still required by Aggregation devices
- Merge Multicast VLAN using Cisco's MVR

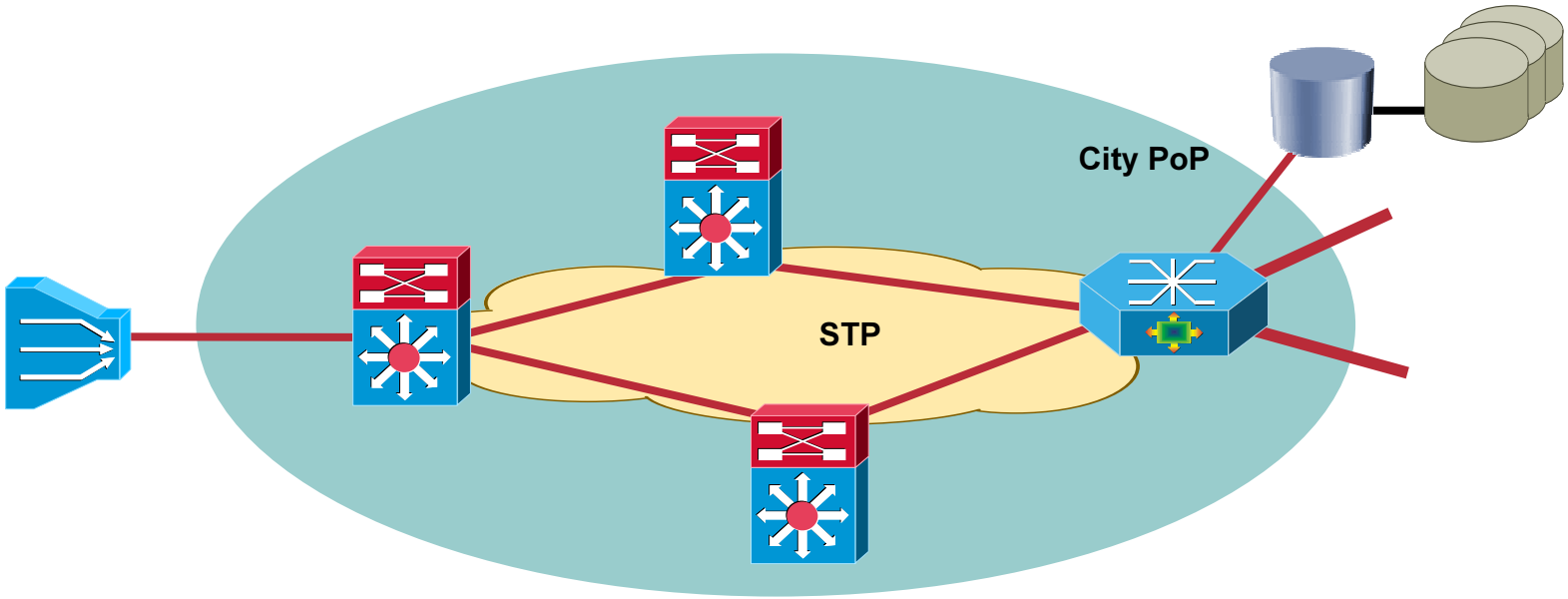
  At L2 aggregation switch

  At DSLAM



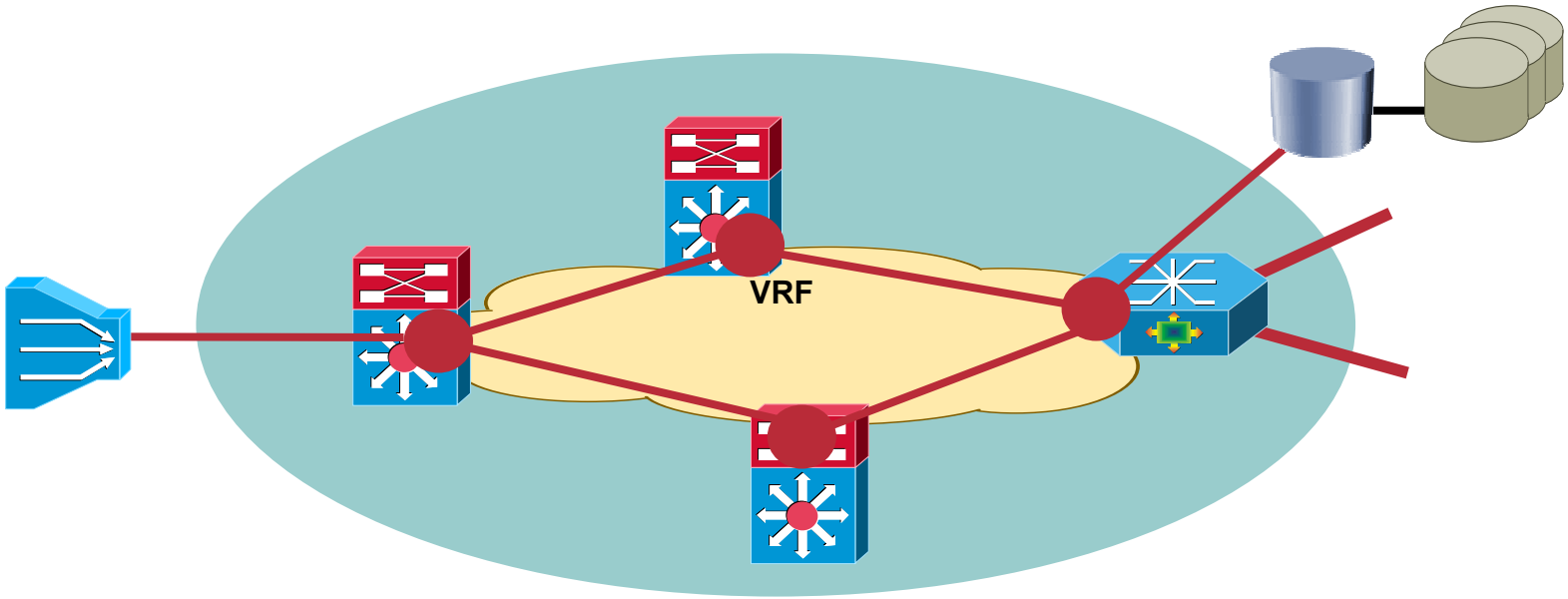
Aggregation Network

- **Ethernet based**
- **Alternative aggregation option (not preferred) using L2 technology**
- **IP/Ethernet or IP/Ethernet/PW/MPLS/Ethernet**
- **.1q CoS**
- **VLAN based service separation/virtualization**



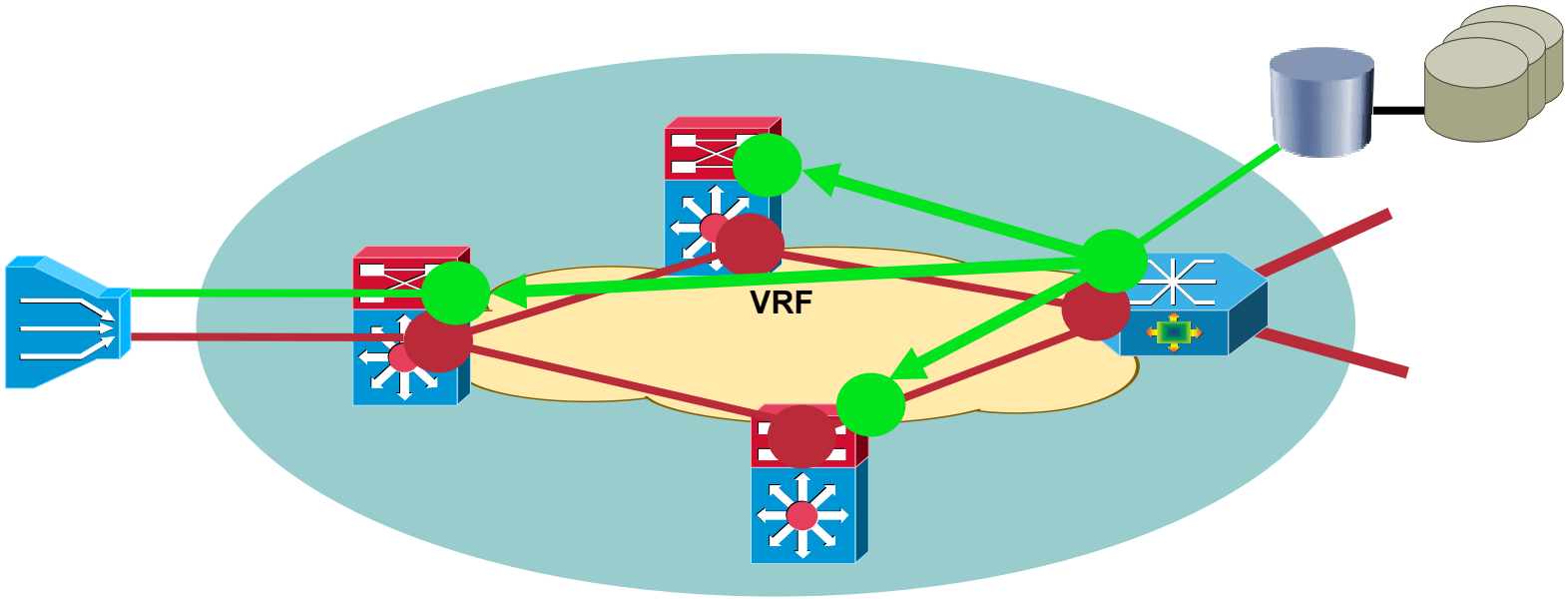
Aggregation Network (cont)

- **Ethernet based**
- **L3 option (preferred)**
- **IP/Ethernet**
- **Per-class QoS**
- **VRF based service separation/virtualization**



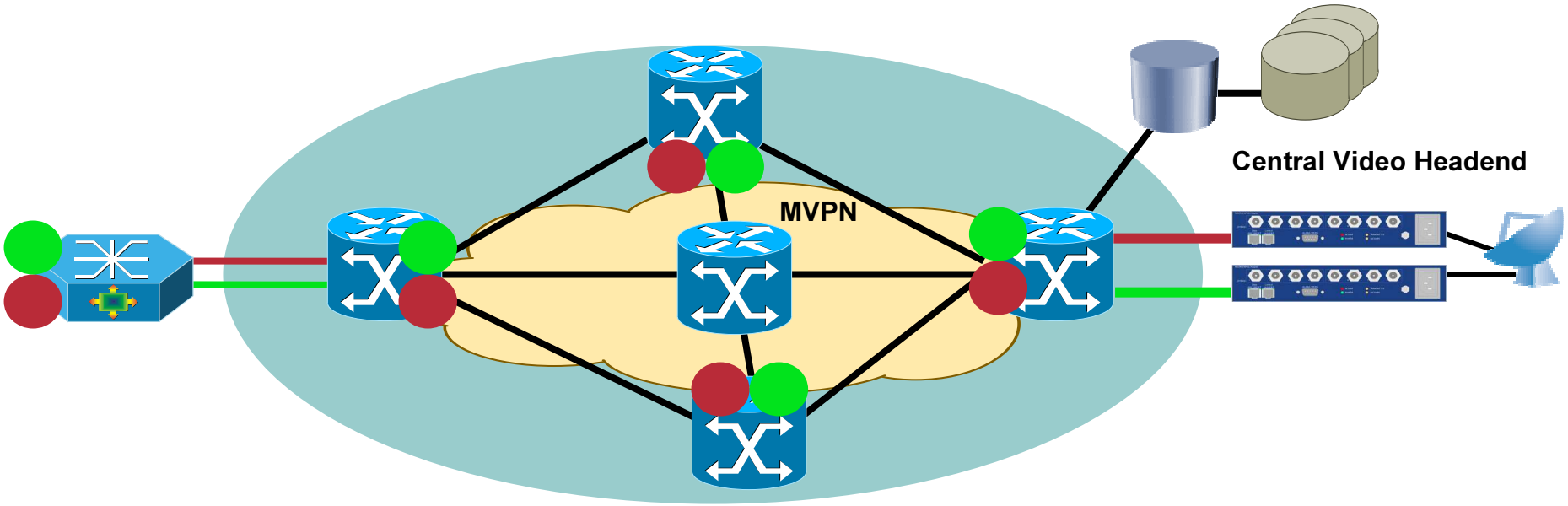
Aggregation Network (cont)

- **Per-VRF Topologies for further service separation (unicast, mcast)**
- **Optional Asymmetric Networking for cost reduction**
  - **Integrated DWDM optics**
  - **Receive-only GBIC and Xenpak**
  - **UDLR/RFC-3077 for return path**



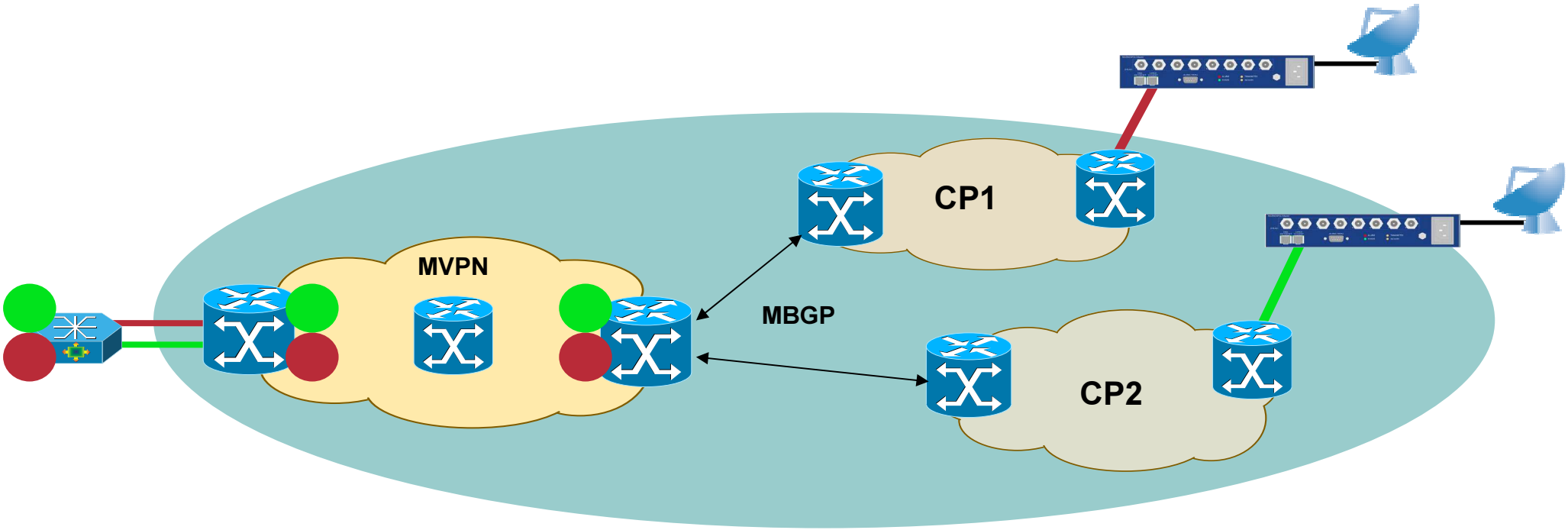
Distribution Network

- SP runs its own Video Headend
- MPLS core network
- BGP based MVPN for Multicast service separation/virtualization



Inter-Domain Network

- Content provided by external CP or daughter company
- Inter-AS Multicast Distribution
- MBPG peering between SP and CP
- MVPN in Core Network



Redundancy

Redundancy models

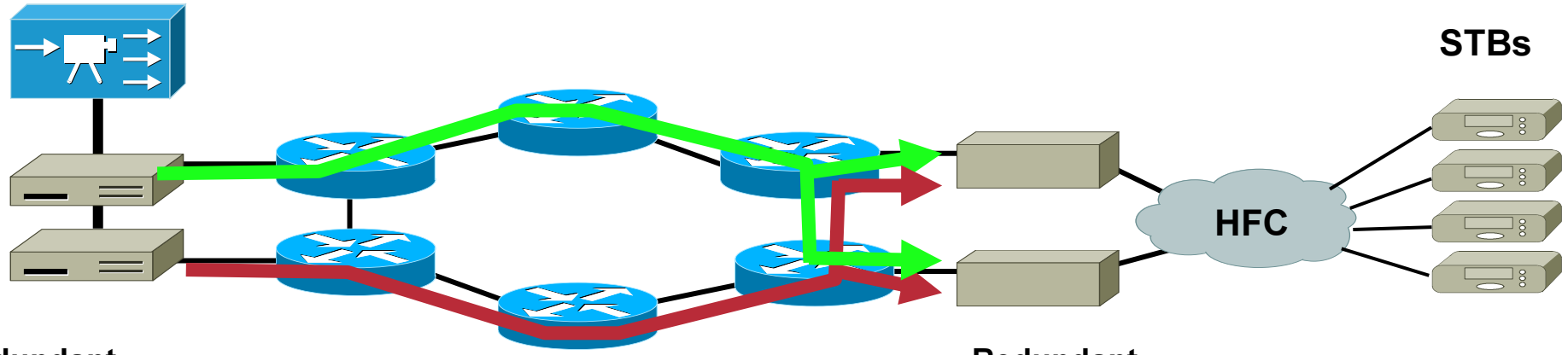
- **Dual streams (1+1 RTP sessions)**
  - **Let the receiver decide which one to take**
- **Heartbeat**
  - **Active sends periodic hello to standby (muted) source**
- **Receiver driven**
  - **Same group with two sources. STB decides which one to join using IGMPv3**
- **Anycast-Source**
  - **Two (or more) sources actively sending with same origin IP address**
  - **Network decides which one to use using its metrics**
  - **Disaster-recovery and redundant headend applications**

Stream redundancy with path separation

A solution that can guarantee 0 loss upon single network outages without adding latency

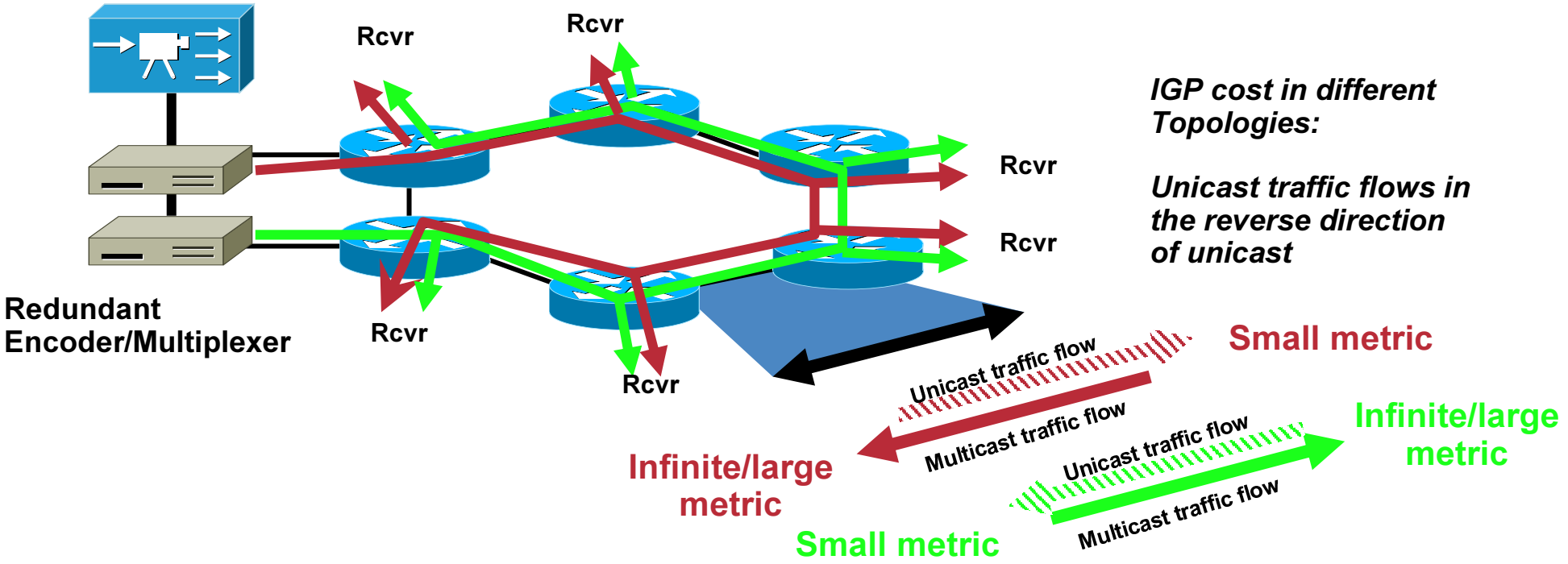
- **Duplicate copies of multicast data**
- **Long-time use in finance market data feeds**
  - **Source and receiver hosts handle creation and elimination of duplicates**
  - **Two networks built:**
  - **No single network failure will impact both flows**
- **Also starting to see interest in broadcast industry**

Stream redundancy with path separation
Candidate example from broadcast-TV in cable



- Encoder/Multiplexers generate two copies of IP multicast flows
- Network uses methods of path separation
  - Multiple IGP instances, topologies, two networks, VRF-lite, RSVP-TE, …
- Each receiver consumes both copies
  - Remove duplicates by sequence numbers (eg: MPEG timestamp).
  - Any single failure in network: 0 packet loss. 0 added latency
- Same bandwidth allocation needed as in traditional SDH rings, but solution even better: 0 loss instead of <= 50 msec.

Stream redundancy with path separation
Some IGP details



Can share links for two copies in rings !

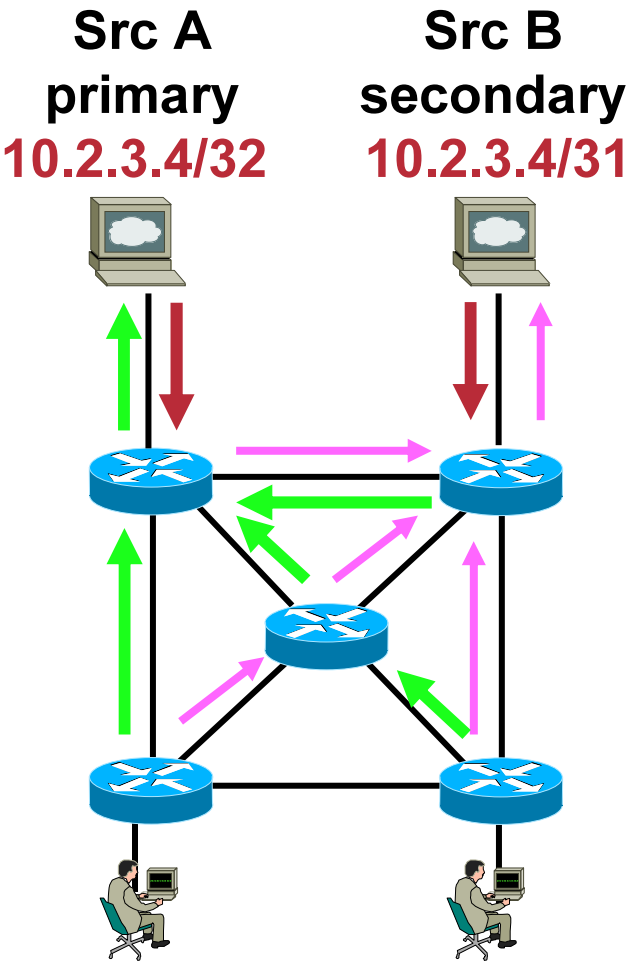
Use asymmetric metrics!

May need infinite metric if reconvergence is not wanted

Available in IS-IS, in draft only for OSPF

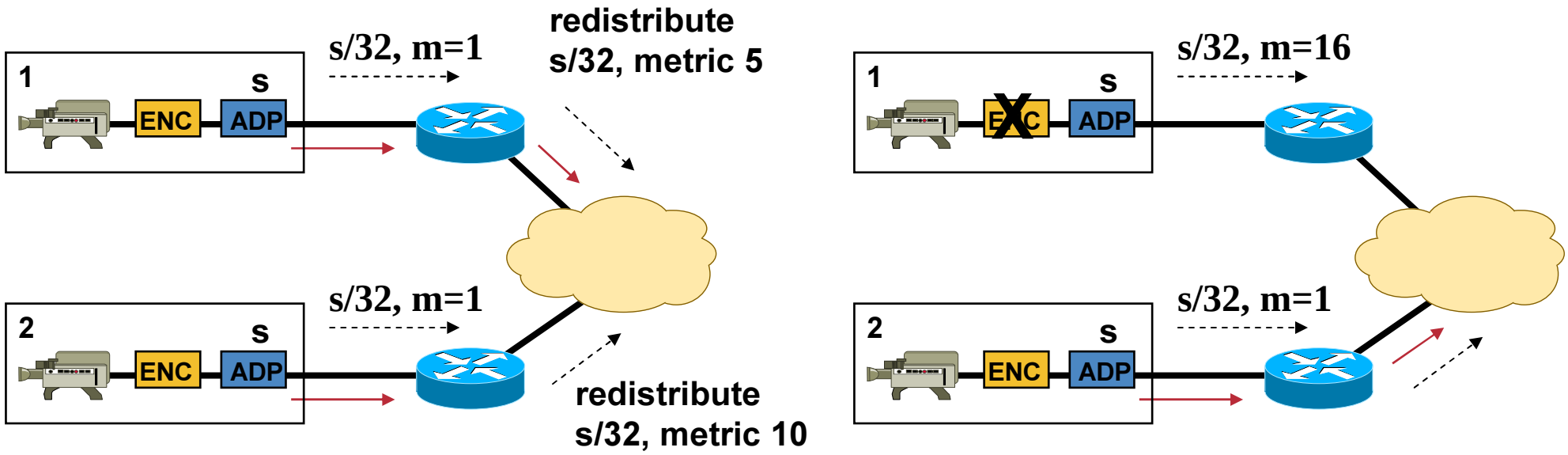
Source Redundancy
Anycast - Prioritycast policies

- **Policies**
  - **Anycast**: clients connect to the closest instance of redundant IP address
  - **Prioritycast**: clients connect to the highest-priority instance of the redundant IP address
- **Also used in other places**
  - Eg: PIM-SM and Bidir-PIM RP redundant, DNS
- **Policy simply determined by routing announcement and routing config**
  - Anycast well understood
  - Prioritycast: engineer metrics of announcements or use different prefix length.



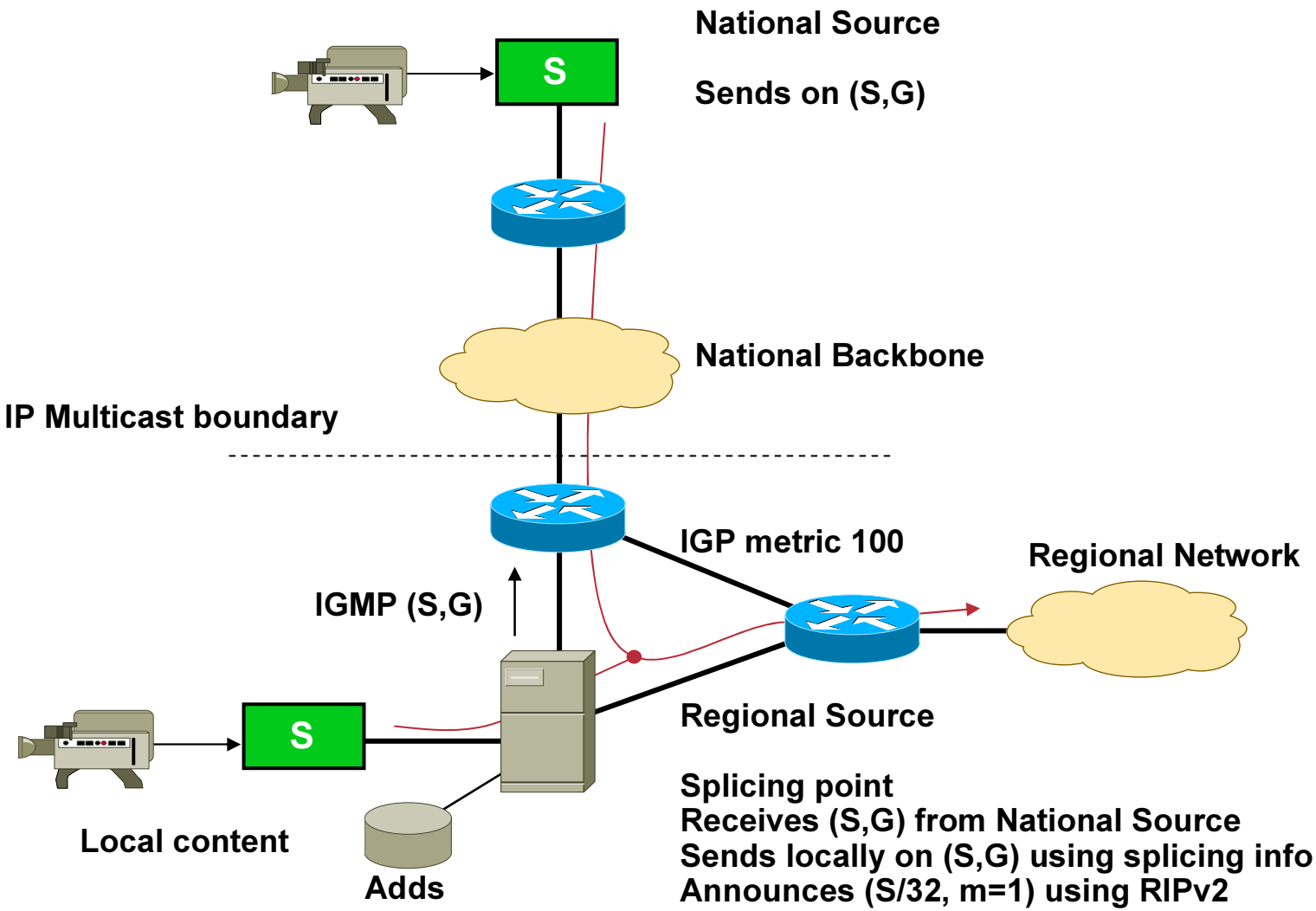
Example: Prioritycast with Prefix length announcement

Anycast-Source with RIPv2 Poison Reverse



- The two sources are active and sending
- s/32 routes are generated by both source using RIPv2 updates
- Host routes for anycast source are redistributed into IGP with variable metrics (optional)
- Network selects source (PIM join messages) based on metric
- Upon video failure, sources withdraw s/32 routes using Poison Reverse (infinite metric) updates

Local content insertion with splicing



IPTV Quality of Experience (QoE)

Quality of Experience (QoE)

- QoS relates to packet performance from a network perspective
  - SLAs based on packet metrics: delay, jitter, packet loss
- QoE relates to service performance from a user perspective
  - SLAs based on viewer perception: image quality, easiness of use
- QoE requirements must be mapped to network requirements at the different layers
  - QoE requirement: less than 1 visible artifact per hour of streaming
  - Equivalent to a Packet Loss Rate (PLR) < 10^{-6} for a 4 Mbps stream
  - Equivalent to a Bit Error Rate (BER) < 10^{-10}
- Defined by DSL Forum WT-126

QoE metrics

- **Responsiveness (control)**
  - **Channel change time for Broadcast TV**
  - **Trick mode delay for VoD**

  < 200 msec considered interactive, < 2 sec considered responsive
- **Resiliency (media)**
  - **Amount and Frequency of packet losses and its impact**

  From minor disturbance (pixels) to screen blackout
- **Other subjective aspects**

  Easiness of use of EPG

  Quality of content

  ...

Resilience to impairments

- Lost packets have a visible impact
  - No application-level concealment mechanisms for IPTV
- Packet losses are due to
  - Network failures (loss proportional to re-convergence time)
  - Excessive queuing at aggregation points (many sources, many hops, link/node congestion)
  - Link transmission errors (common in access lines, e.g. DSL. uncommon in core optical links)
- Visible impact of an individual packet loss is variable
  - Depends on where you hit the Video Sequence (I, P or B frame)
  - Depends on the structure of the Video Sequence (GOP size)

Link errors and concealment techniques

- DSL uses FEC and interleaving to correct from impulse and stationary noise
- DSL modems re-train when BER > 10^{-7}

  This results in PLR >> 10^{-4} or > 1 artifact per 2 min.
  Unacceptable. Video requires PLR < 10^{-6}
- Bit errors cause packet losses (UDP checksum)
- Need a mechanism to monitor packet losses
- Mechanisms to recover from packet loss

  Source Block repetition

  Packet FEC (Pro-MPEG Forum COP-3)

  RTP retransmission
- All these require RTP transport

Summary

IPTV is about

- **IPTV is about BB subscriber services**
  - **Broadcast TV, VoD, Switched Digital Video (today)**
  - **Future interactive video services**

... over a DVB/ATSC application (MPEG) layer providing

- **Compression, Packetization, Multiplexing, Scrambling, ...**

... over an RTP/UDP session layer providing

- **Sequencing, Timestamping, Multiplexing, Checksum**

... over an IP packet network layer providing

- **Intelligent packet operations, QoS, CAC, FEC, retransmission, ICC, ...**

IPTV is not

- **IPTV is NOT about transmission**

  It is IPTV, not VPLS-TV

- **Efficient transmission is always desirable and provides for fast convergence but does not cure packet loss**

  One single packet loss (5 msec impairment) may result in a visible effect lasting for more than 500 msec

  OTOH, good L3 queuing and FEC techniques prevent and recover from packet losses

- **The transmission layer does not provide intelligent services to TV. Those are provided at the IP layer and above**

  Packet FEC, packet retransmission, block transmission may actually recover from packet losses

- **IPTV Service SLAs are packet based (drop, jitter, delay)**

  Analyze and Monitor packets. Do not just Snoop them

Q and A