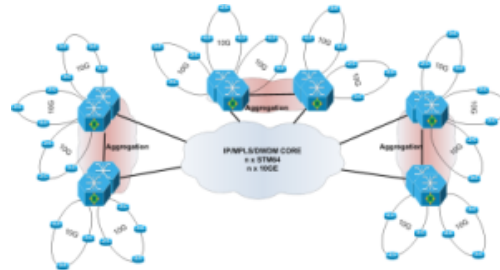


# Reinventing the Access Network



**Anastasios Chatzithomaoglou**  
([achatz@forthnet.gr](mailto:achatz@forthnet.gr))

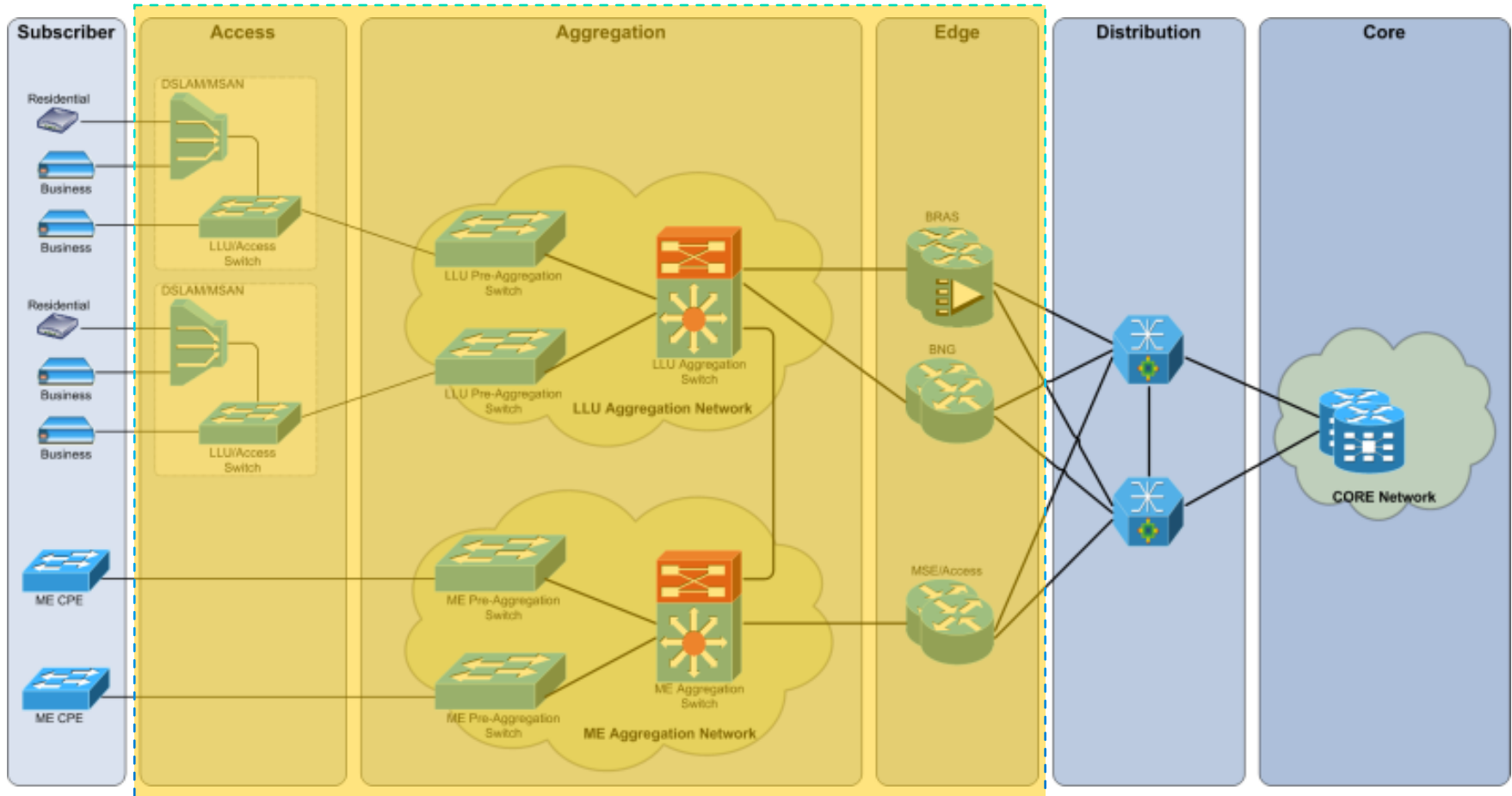
**IP Engineering – Forthnet**

17-10-2013

# Reinventing the Access Network

## High Level Design

This presentation focuses on this part of the network



### The usual jargon

#### ☐ **LE (Local Exchange )**

- ☐ The place where subscriber lines are terminated per geo-area

#### ☐ **POP (Point of Presence)**

- ☐ The place where Aggregation & Edge Routers can be installed

#### ☐ **Access Router**

- ☐ The L2/L3 device that connects all L2 access devices to the rest of the network

#### ☐ **Aggregation Router**

- ☐ The L2/L3 device that connects multiple Access Routers to multiple Edge Routers
- ☐ Multiple levels of aggregation can exist

#### ☐ **Edge Router**

- ☐ The L3 device on the Edge Network that offers the final IP service to the subscriber (i.e. BRAS/BNG)

### Business

#### L2 Services

- ☐ EPL/EVPL
- ☐ VPN/ELAN
- ☐ NNI

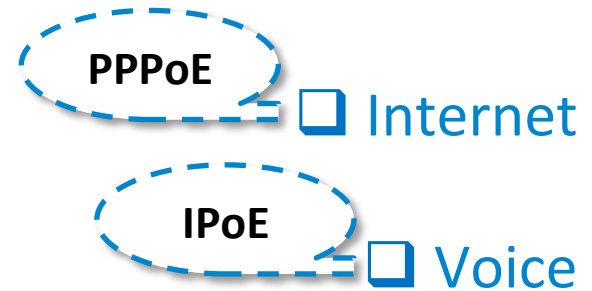
#### L3 Services

- ☐ Internet
- ☐ Voice
- ☐ VPN
- ☐ NNI

Services

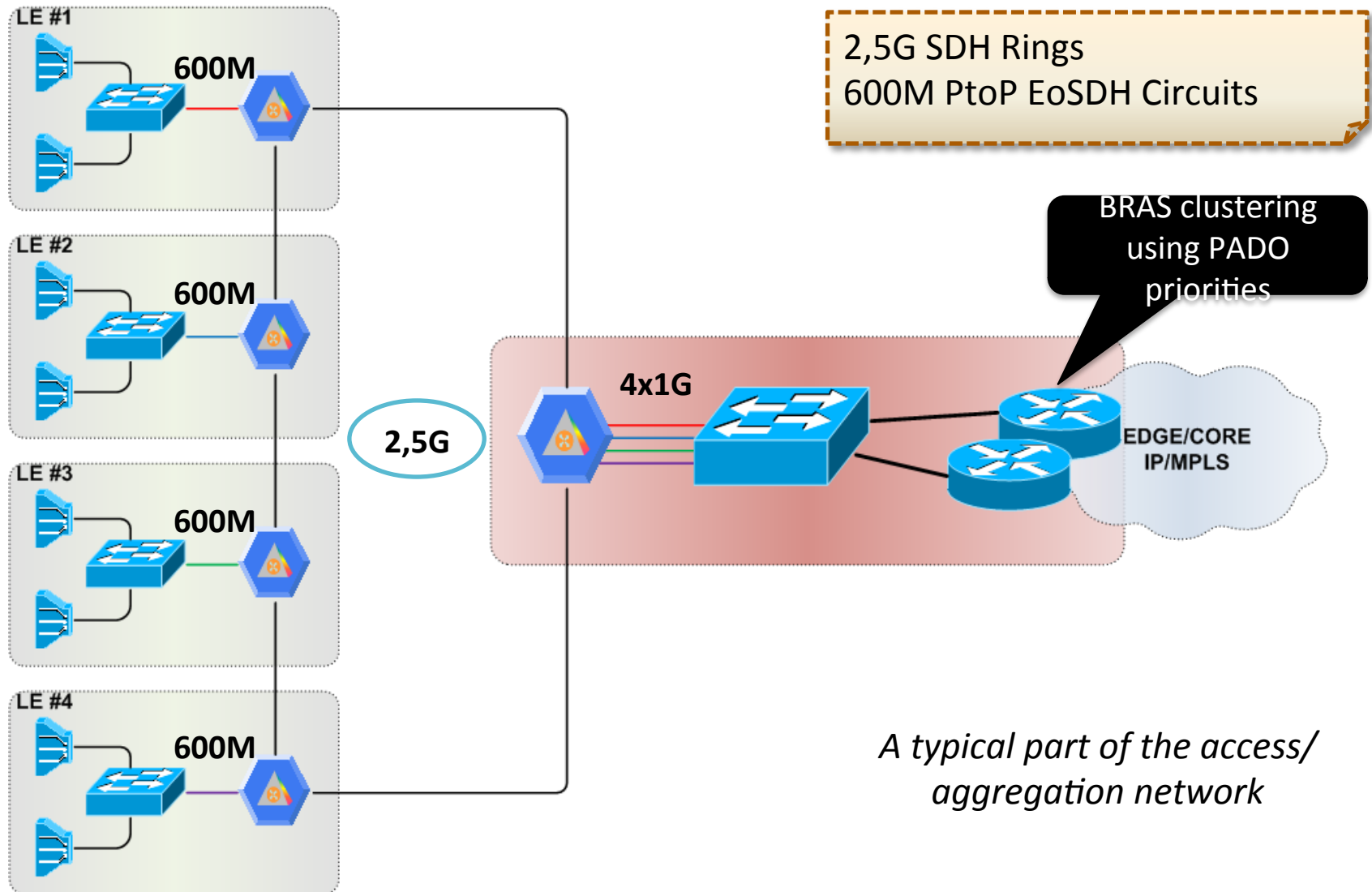
### Residential

#### L3 Services



# Reinventing the Access Network

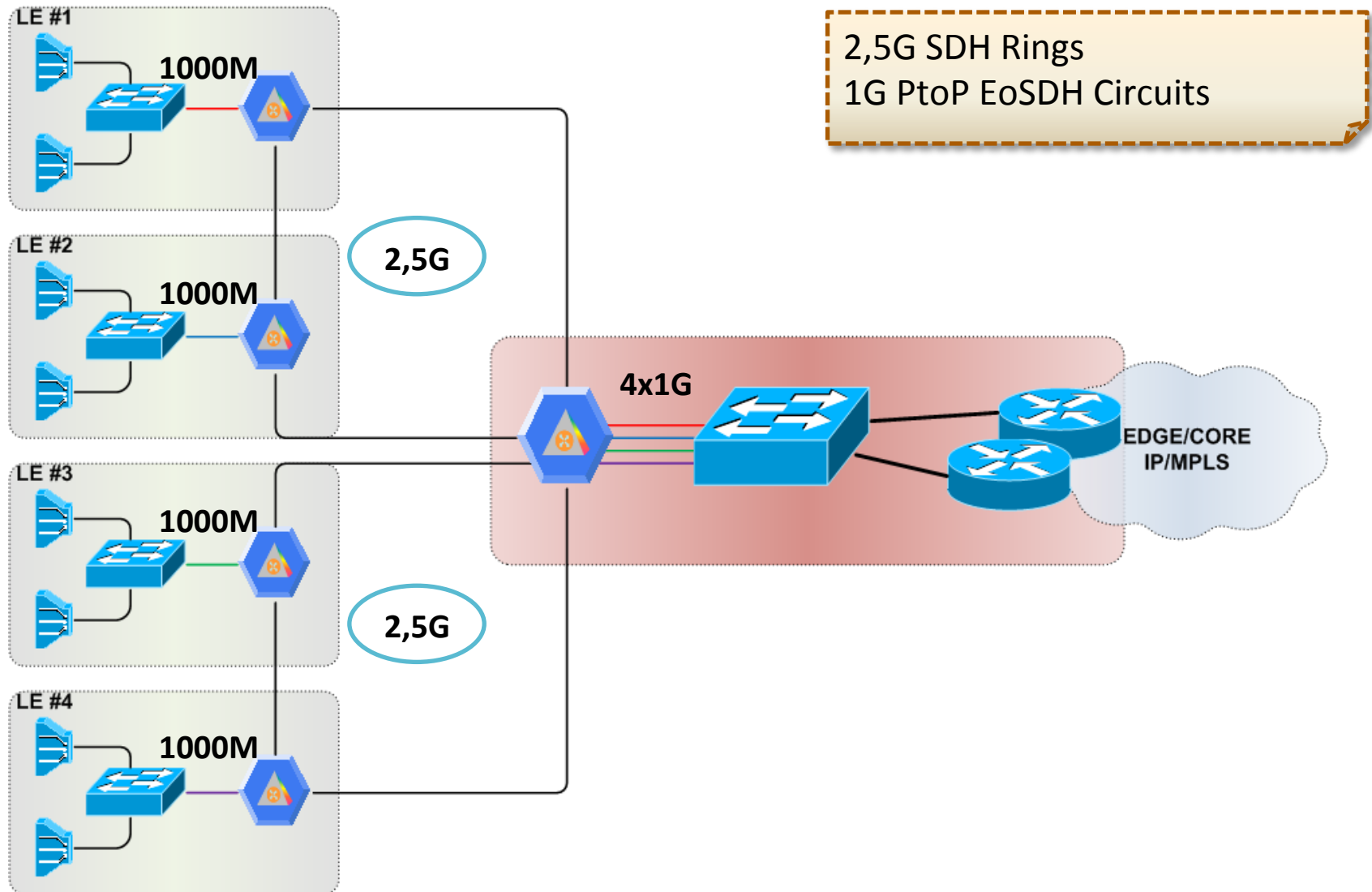
## Access Network in 2006



*A typical part of the access/  
aggregation network*

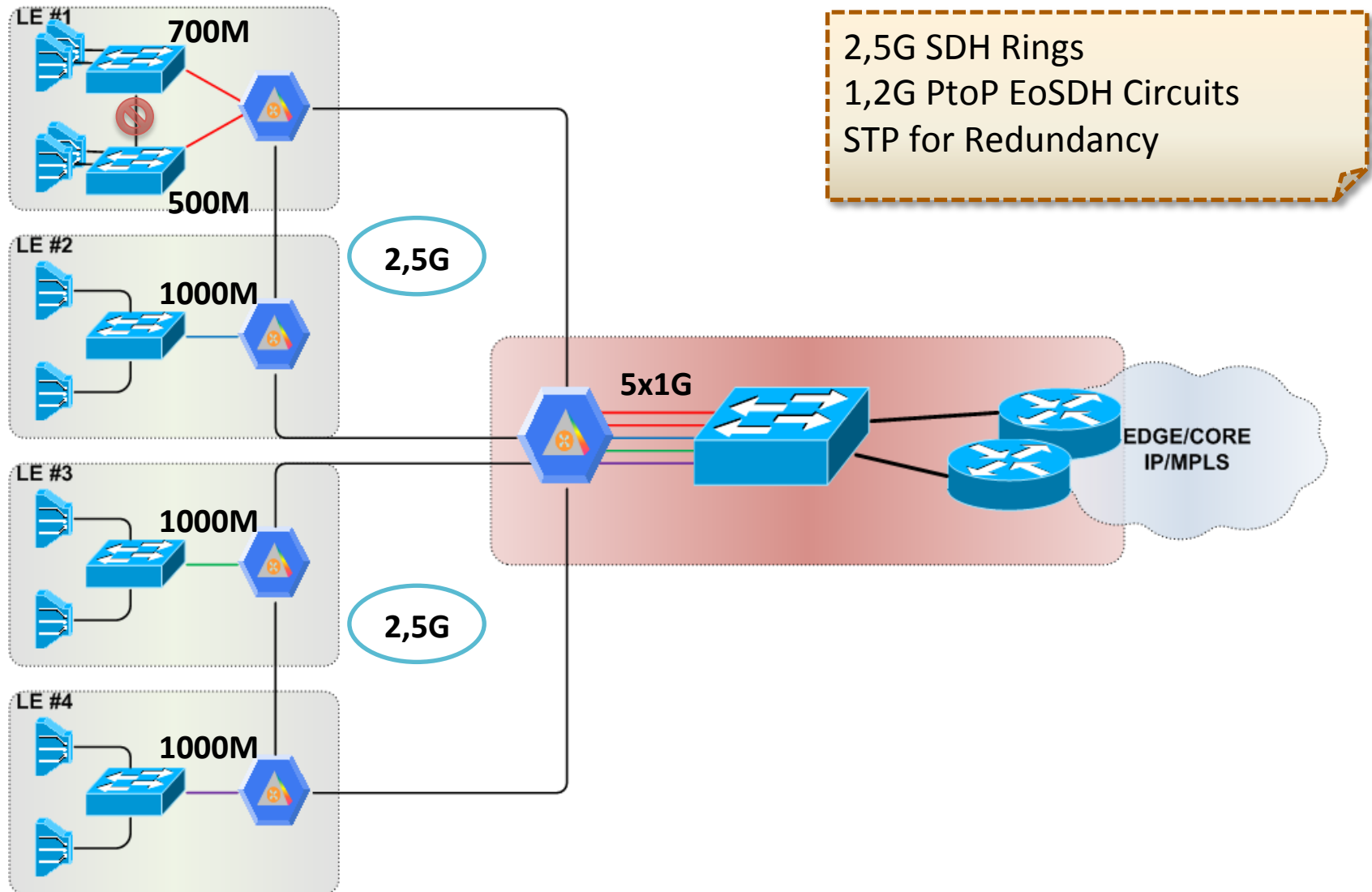
# Reinventing the Access Network

## Access Network in 2008



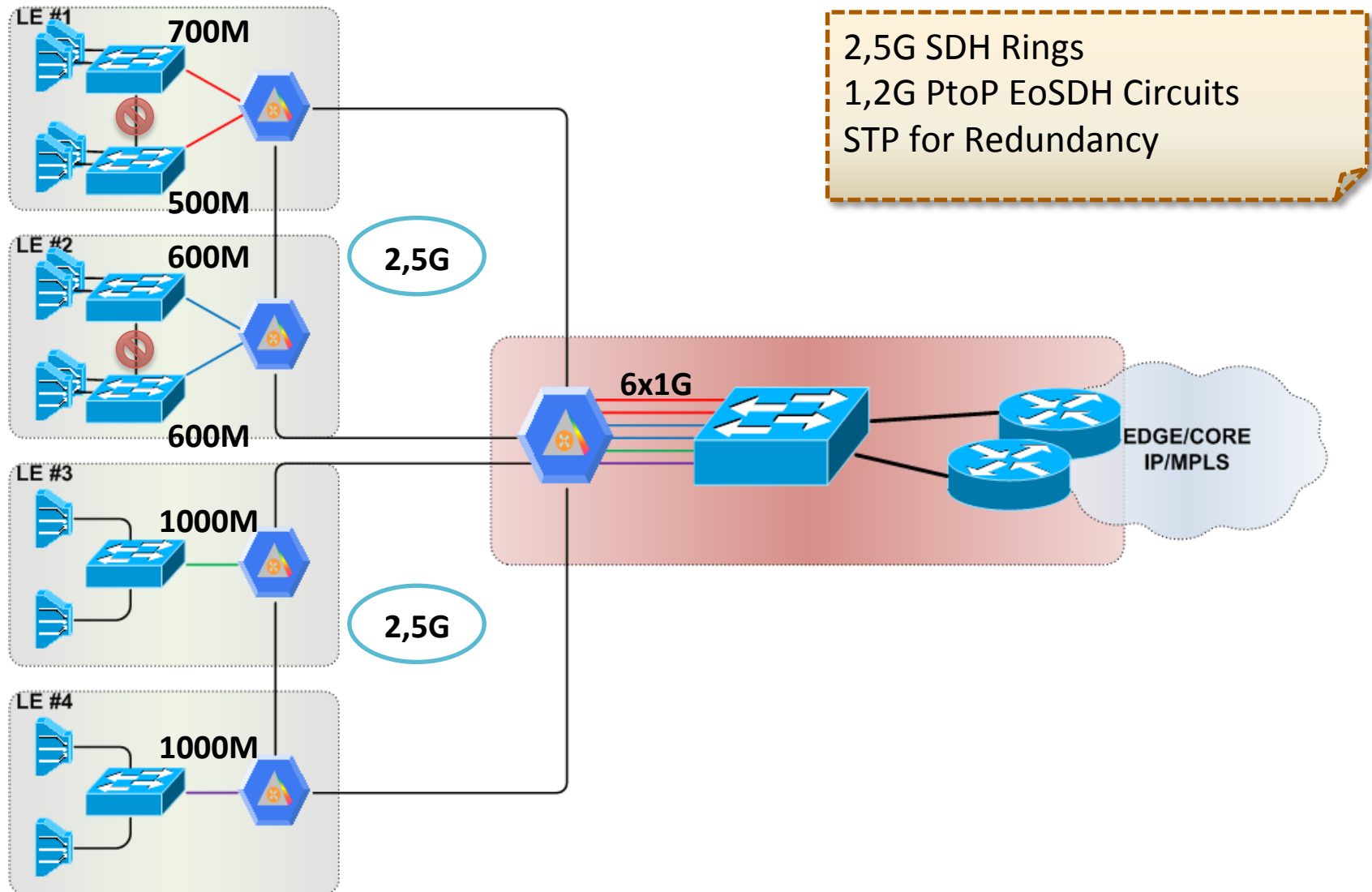
# Reinventing the Access Network

## Access Network in 2010



# Reinventing the Access Network

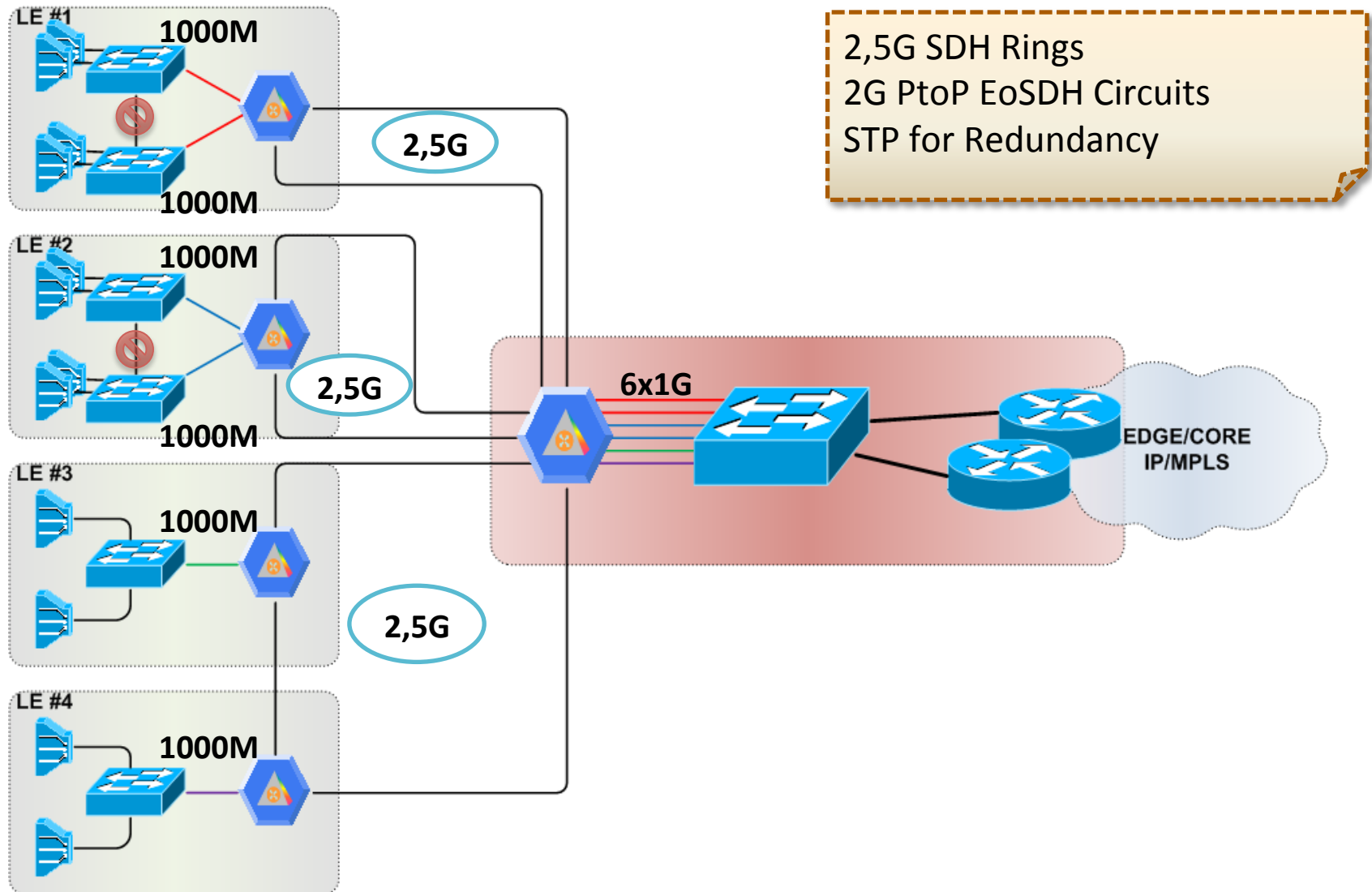
## Access Network in 2011





# Reinventing the Access Network

## Access Network in 2012



### Issues with old design

- ❑ Legacy SDH STM-16 transport
- ❑ Large L2 domains (macs/broadcasts/loops)
- ❑ Limited vlans (even with QinQ)
- ❑ Max 2x1G capacity
- ❑ Active/Standby Redundancy (based on STP)
- ❑ Limited mac-address space

# Reinventing the Access Network

## Requirements of new Design

### General Requirements of new Design

- ☐ n x 10G only (40G/100G in the future)
- ☐ L2CP transparency (especially for business services)
- ☐ QoS bits transparency
- ☐ Jumbo frames (> 9000 bytes)
- ☐ Active/Active Redundancy whenever possible
- ☐ No loss (< 50 ms) upon any Direct Link/Node failure
- ☐ Minimal loss (< 1 sec) upon any Remote Link/Node failure
- ☐ No need for very large scalability
- ☐ As much formulation/standardization as possible

### L2 HW Solutions

Multiply 1G uplink of Access Switch



Install 10G Access Switch



### L3 HW Solutions

Install 10G Access Router



### Transport Solutions

Upgrade SDH to STM-64



Replace SDH with WDM



Remove SDH



# Reinventing the Access Network

## Thinking about L2

### **G.8032v2**

- ☐ Too cumbersome
- ☐ Extra vlans per LE
- ☐ Limited public exposure

### **TRILL & SPB**

- ☐ Mostly focused on DC (not applicable for Carrier Ethernet)
- ☐ Limited OAM functionality (under development)
- ☐ Non-existent support by CE products

### **Vendor Proprietary Solutions**

- ☐ Might do the job quite well
- ☐ Possible vendor lock-in
- ☐ Prefer vendor agnostic solutions, unless no such solution exists

### **Final Decision**

- ☐ Move toward unified network architecture based on IP/MPLS
- ☐ Expand L3 deployment from Core/Edge to Aggregation/Access

# Reinventing the Access Network

## Thinking about L3

### IGP + optimizations

- ☐ Simple and works in every case
- ☐ Slow convergence (hello/timer tuning might fix it)

### MPLS TE/FRR

- ☐ Too complex (although used in other parts of the network for TE)
- ☐ No easy way for automation (affinity/manual)
- ☐ Explicit paths for inter-area tunnels

### (r)LFA

- ☐ Plug & Play (unless IETF blows it up)
- ☐ Some topologies not covered 100%
- ☐ Micro-loops are possible

### BFD

- ☐ Use to detect losses due to virtual ifs (otherwise it would be detected at PHY, due to FO)
- ☐ SW-based in some platforms

### EoMPLS/VPLS

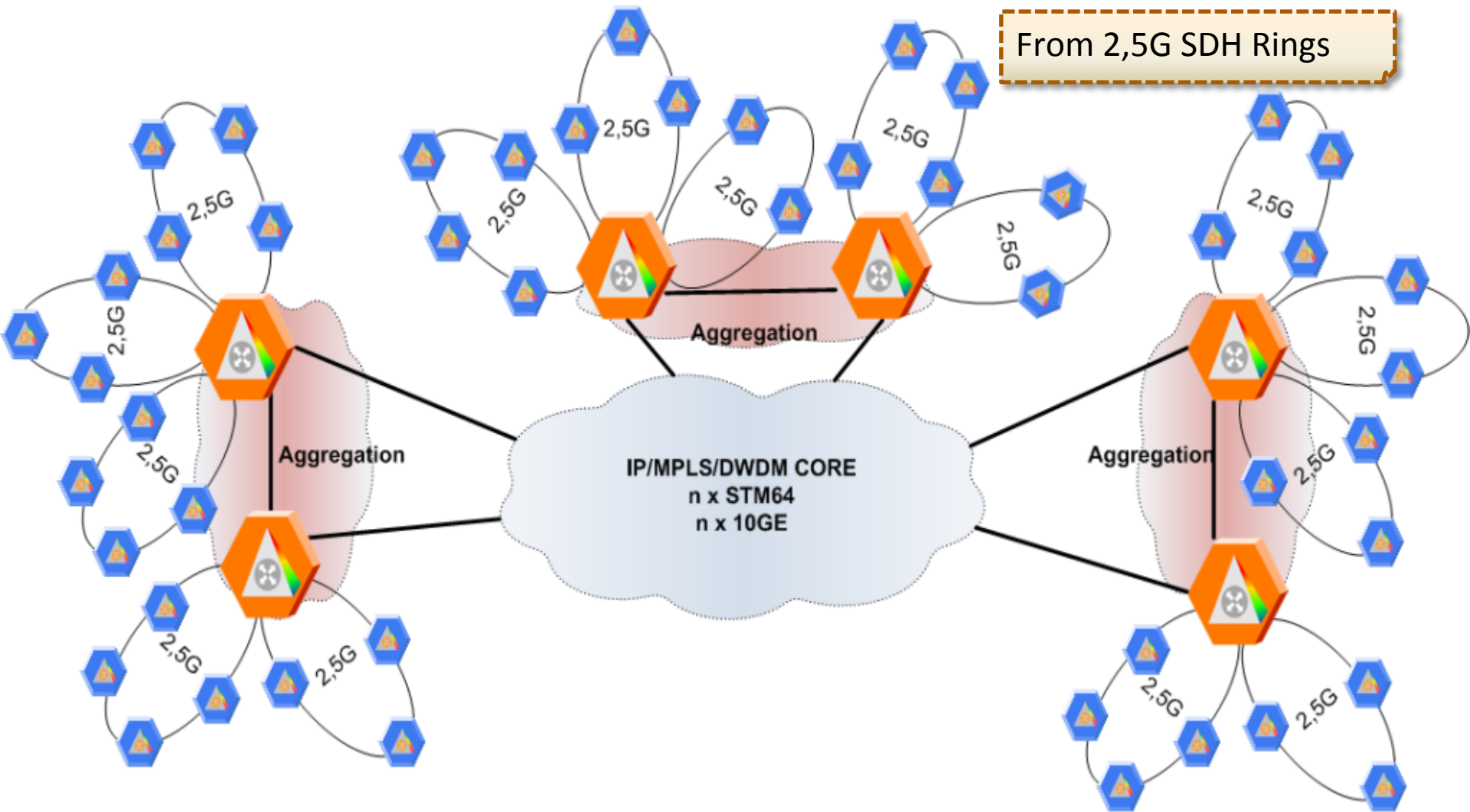
- ☐ A/A and A/S Pseudowires to transfer L2 services from Access to Aggregation

### PW-HE

- ☐ Remove completely Vlans/L2 from Aggregation
- ☐ More Pseudowires from Access to Edge

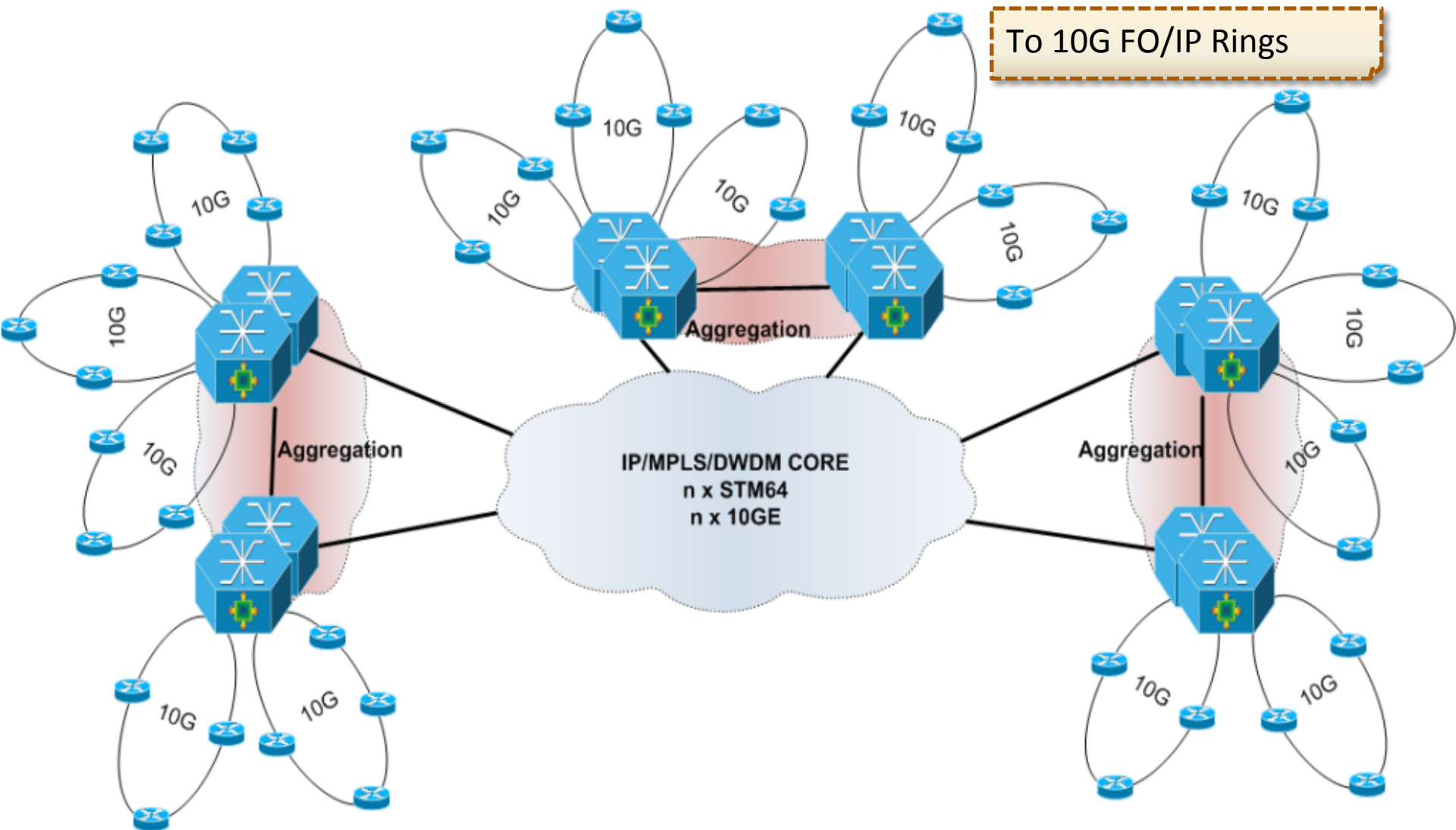
# Reinventing the Access Network

## Old design



# Reinventing the Access Network

## New design





### Old Design

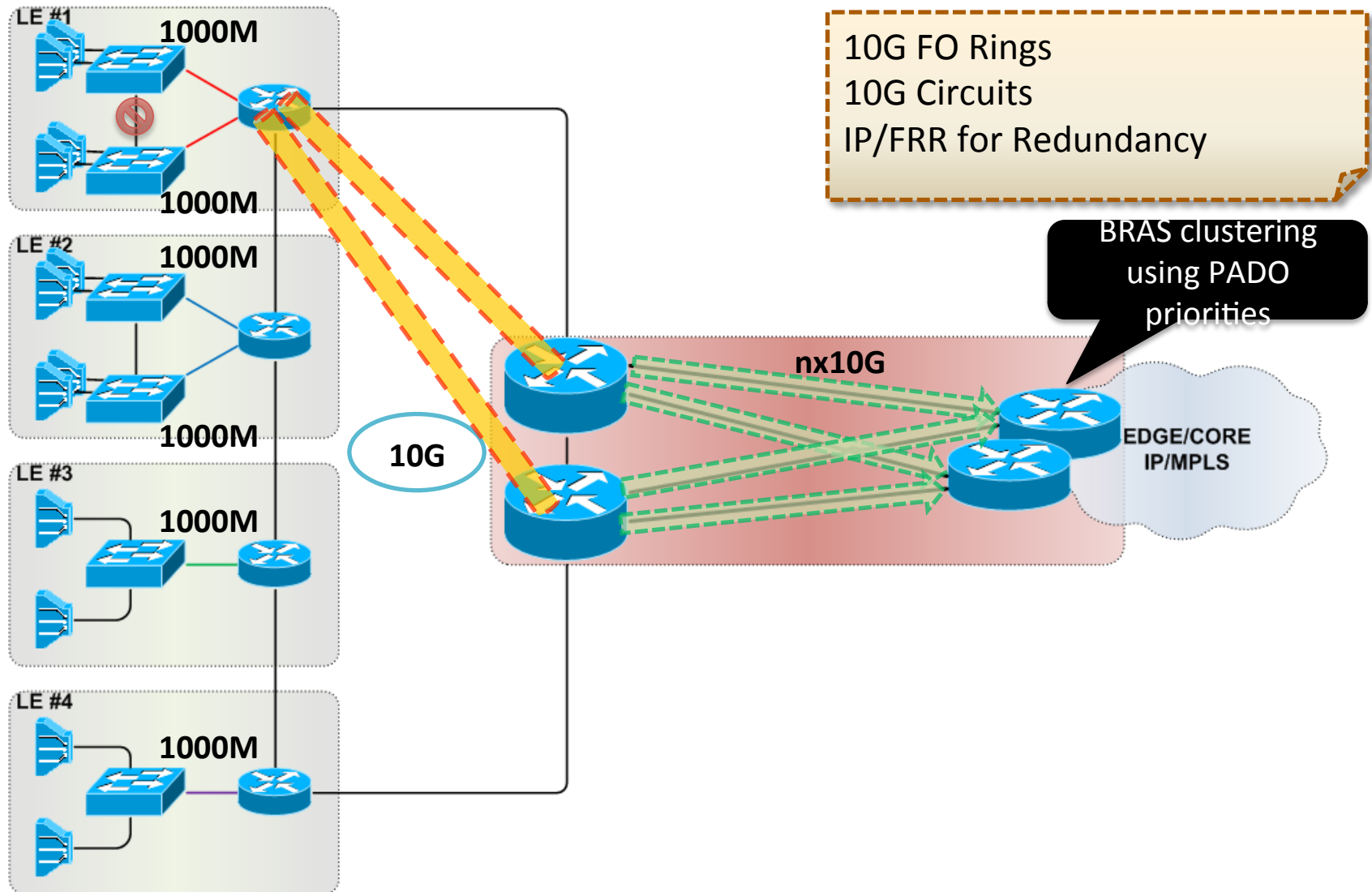
- ❑ Legacy SDH STM-16 transport
- ❑ Large L2 domains
- ❑ Limited vlans (even with QinQ)
- ❑ Max 2x1G capacity
- ❑ Redundancy based on SDH/STP
- ❑ Limited mac-address space

### New Design

- ❑ Direct FO, no SDH
- ❑ Limited L2 domains
- ❑ Many PWs
- ❑ Max 2x10G capacity
- ❑ Redundancy based on IP
- ❑ x10 mac-address space

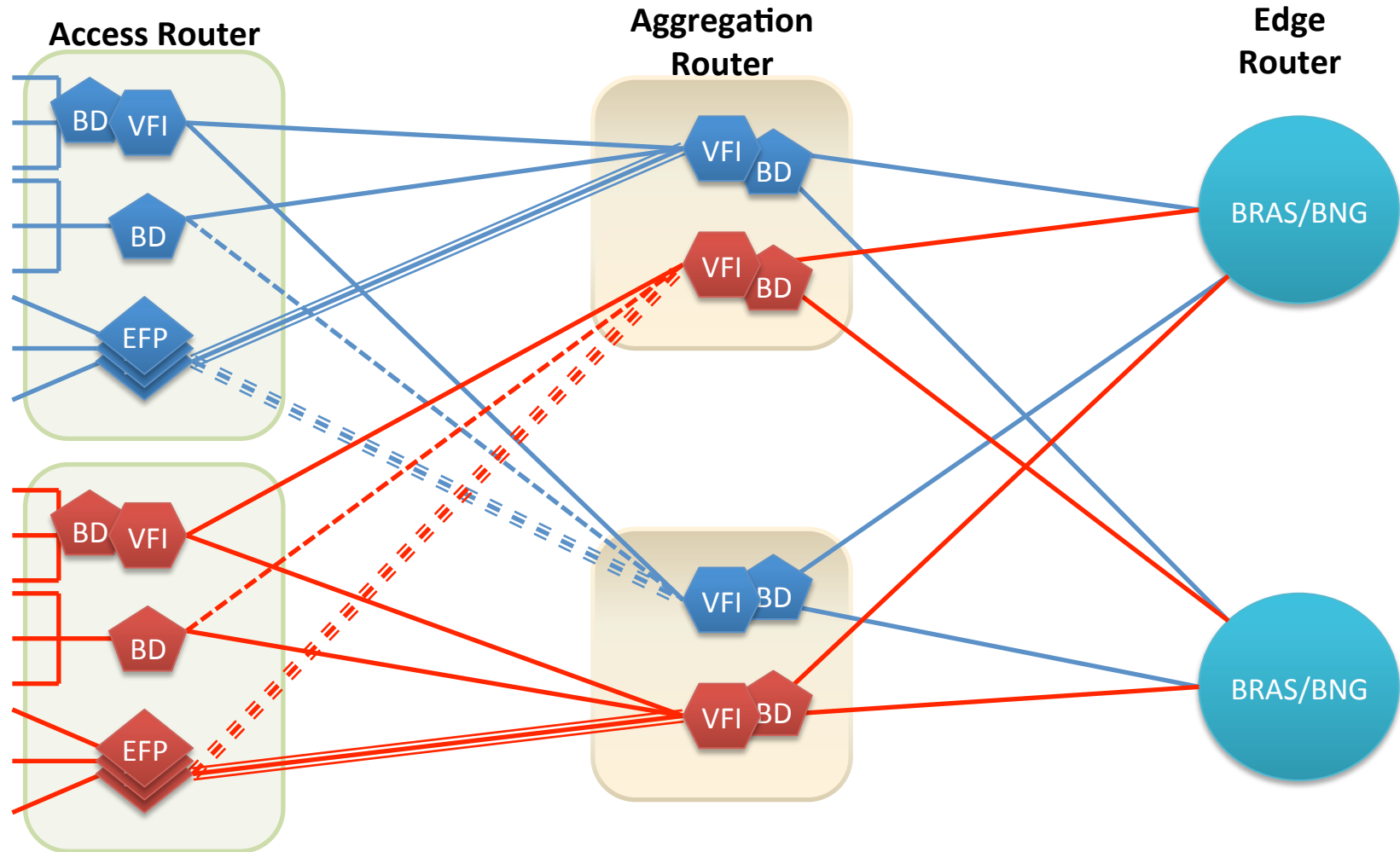
# Reinventing the Access Network

## Access Network in 2013



# Reinventing the Access Network

## Low Level Design



### IGP Details

- ☐ IPv4 10.X.Y.Z addressing (no IPv6 for LDP yet)
- ☐ OSPF (already used in the network)
- ☐ Area 0 for aggregation
- ☐ Area 0.X.Y.Z for all access rings between POPs X & Y
- ☐ Multi-area adjacency (if needed for intra-area vs inter-area)
- ☐ No external prefixes, no ASBRs
- ☐ Every adjacency configured as point-to-point
- ☐ BFD tx/rx 100 multiplier 3 (**buggy**)
- ☐ LFA support only for Loopbacks (**buggy**)

### PW Details

- ☐ Numbering based on **VLAN-NUMBER-SERVICE** (i.e. 3456011100)
- ☐ MTU > 9000
- ☐ Control Word enabled (avoid 4/6 mac issue in LB)
- ☐ Active/Standby if attached to EFP/BD
- ☐ Active/Active if attached to VFI
- ☐ Split-Horizon disabled if > 2 IPoE Access PWs from same LE and no Access BD

### Access Router Management

- ❑ Different Loopback for Management
- ❑ Recursive Static Default Route pointing to an Aggregation Router
- ❑ Double Default Route through IGP (with different metrics)
- ❑ Global => VRF in Aggregation Router towards Management Network
- ❑ OOB over old EoSDH when possible

*Note: Mgmt VRF in Access Routers also under consideration*

# Reinventing the Access Network

## Load-Balancing

### Issue

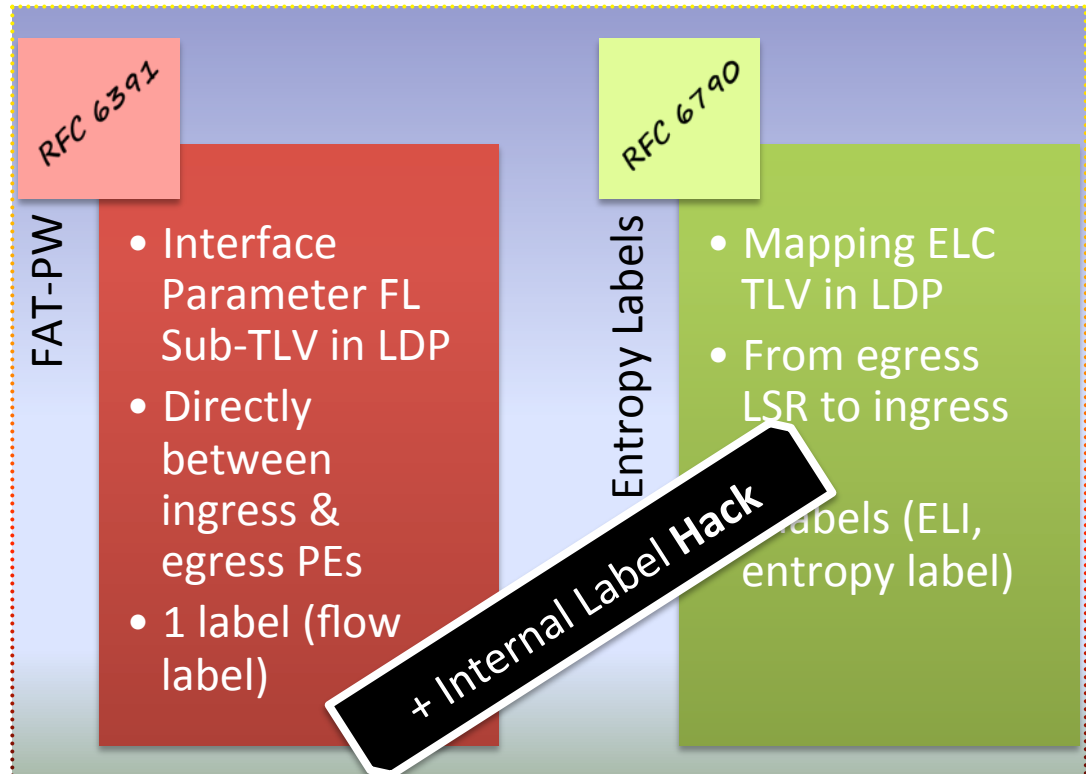
PWs between Access and Aggregation transport large volumes of PPPoE traffic, especially in the downstream direction. Need to make sure that they are split over multiple links.

### Solution

Improve granularity on the hashing of traffic running over PWs by introducing one or more additional labels  
Intermediate nodes need only to make an ECMP choice based on a hash of the MPLS label stack

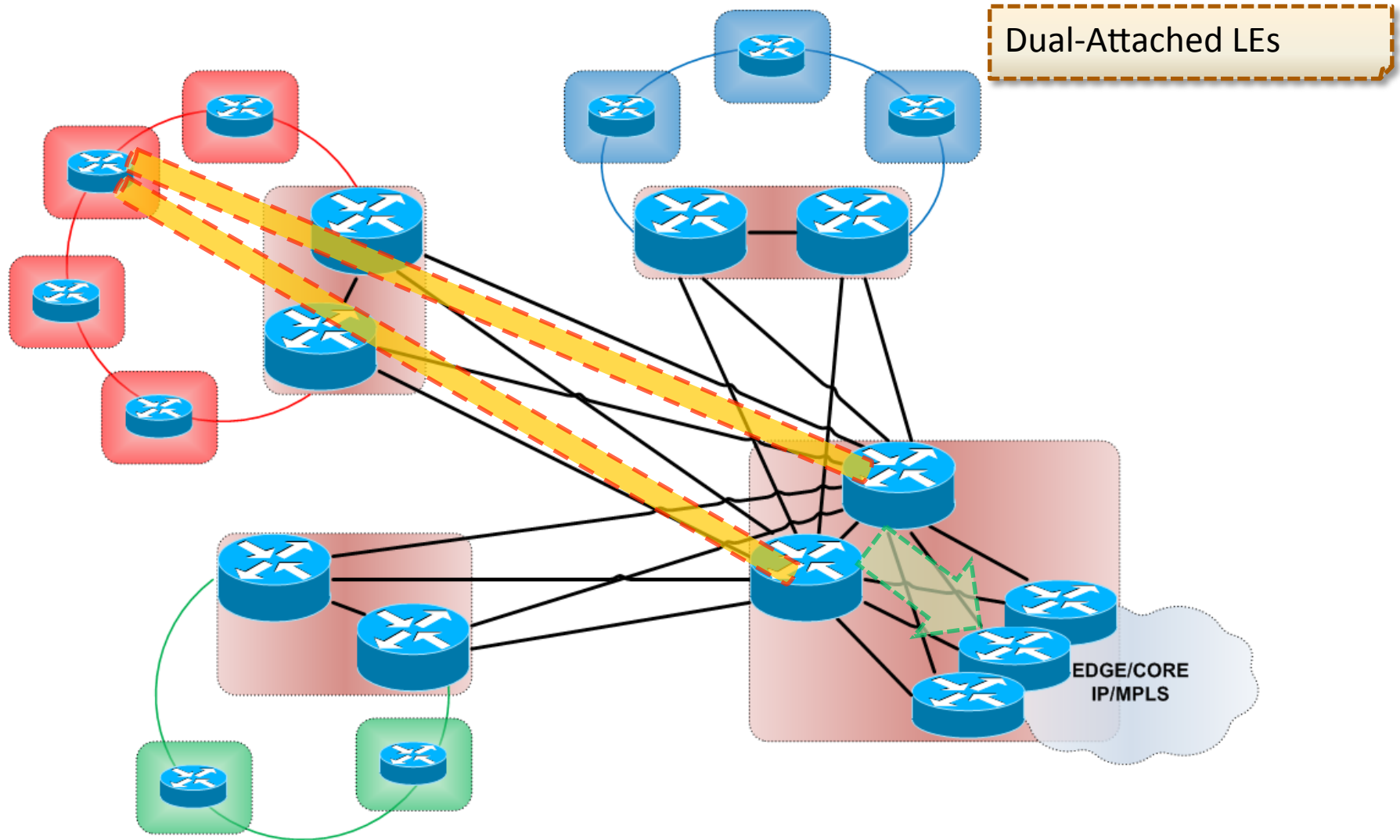
Packet ordering must be preserved only within the context of each individual transported IP/Eth flow

The requirement to load-balance over multiple PSN paths occurs when the ratio between the access PW utilisation and the PSN's core link capacity is large (e.g.,  $\geq 1:10$ )



# Reinventing the Access Network

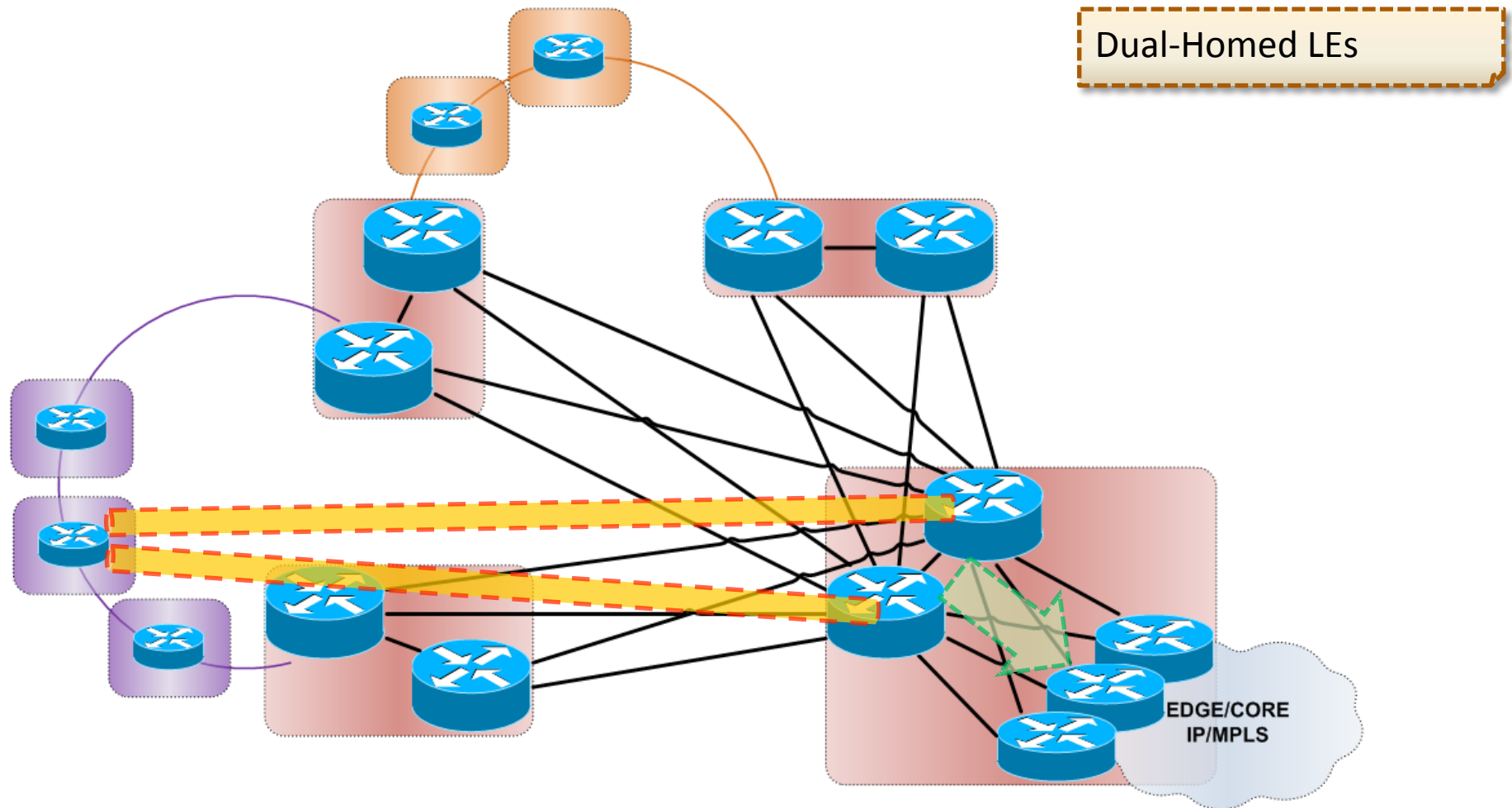
## Access Network in 2013 – Dual-Attached LEs





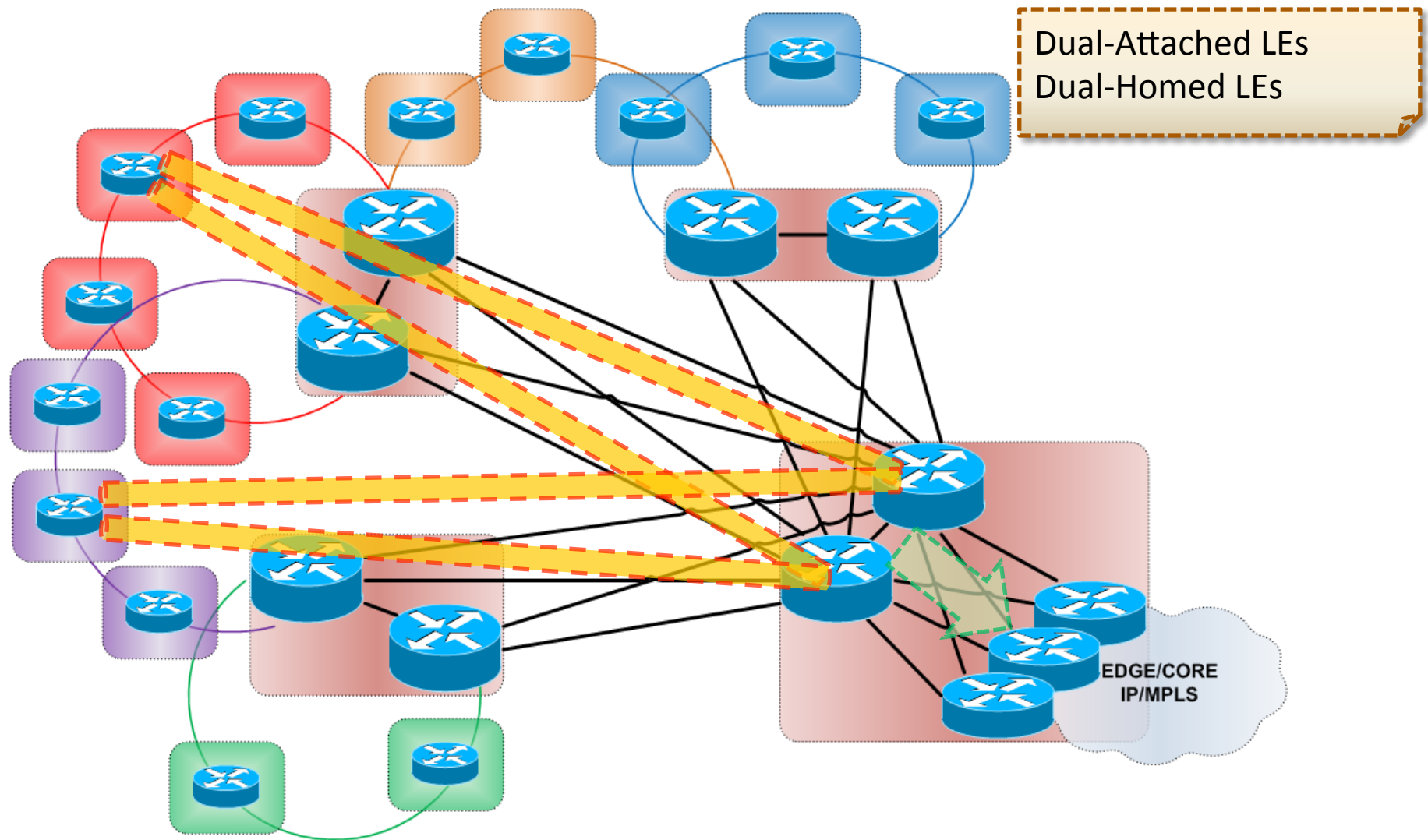
# Reinventing the Access Network

## Access Network in 2013 – Dual-Homed LEs



# Reinventing the Access Network

## Access Network in 2013 – Dual-Attached/Homed LEs

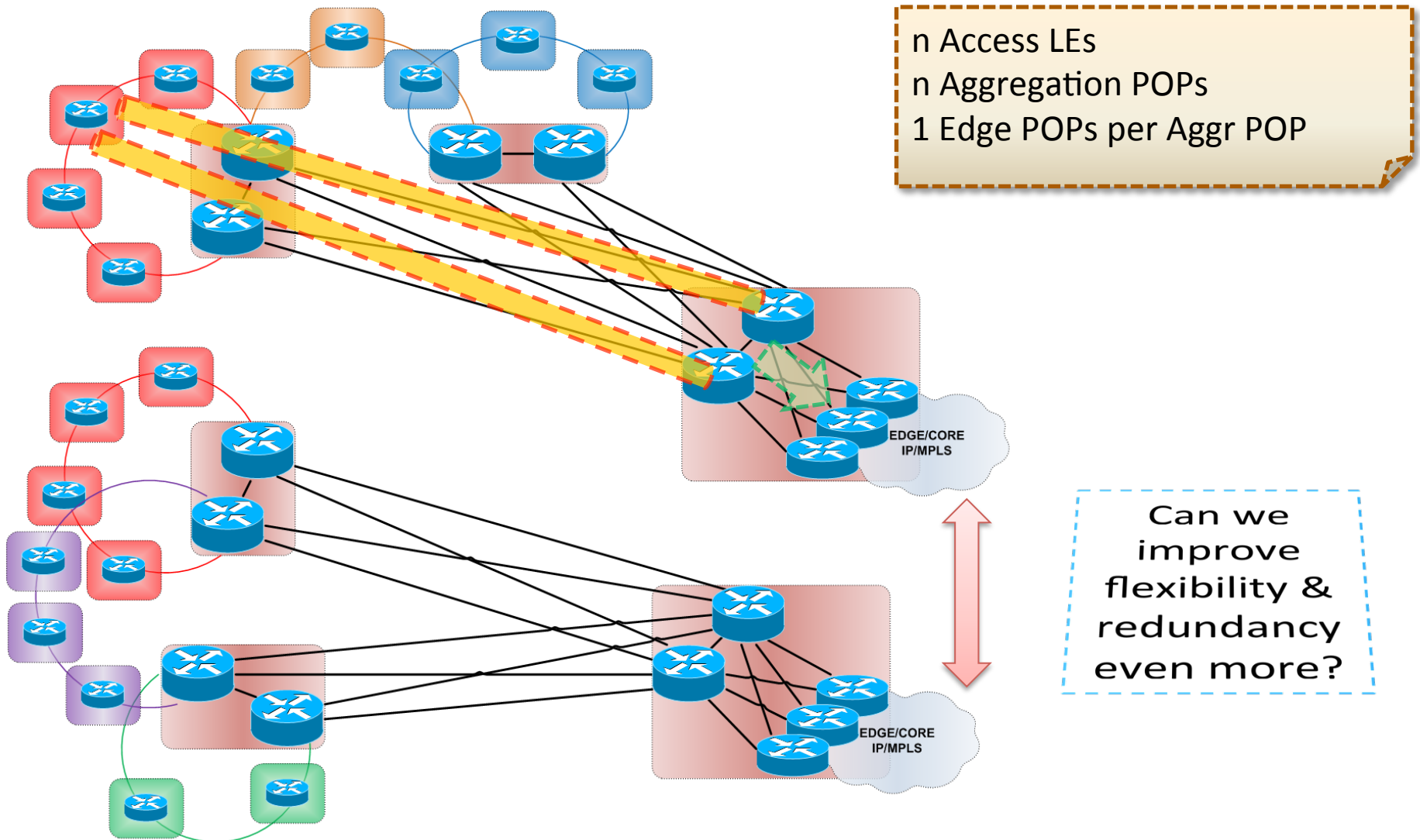


### Migration Steps

- ☐ Move from aggregation towards access
- ☐ Edge doesn't need to be changed
- ☐ New aggregation routers between old aggregation switches and edge routers
- ☐ Extra PWs between old and new aggregation devices
- ☐ Use parallel circuits to change from L2 to L3
- ☐ Many maintenance windows, many bugs

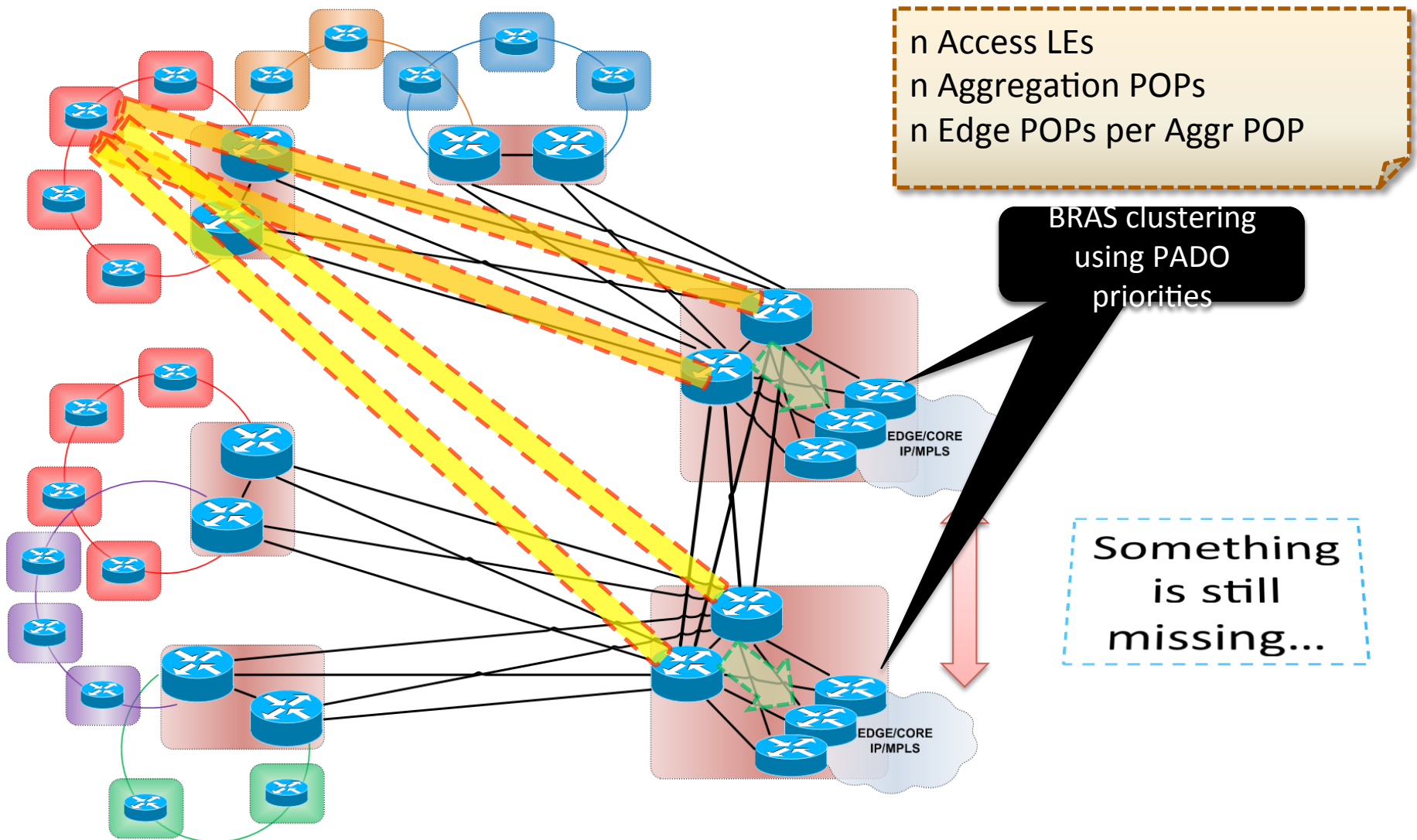
# Reinventing the Access Network

## Access Network in 2013 – Zoom Out



# Reinventing the Access Network

## Access Network in 2013 – Edge Cloud



### Poor Man's SDN

#### The Idea

Create a **framework** that...

# Poor Man's SDN

...**takes** actions...

...**provisions** services...

...with the **minimum** possible user interaction

### Future things to evaluate

- ☐ Multicast efficiency (draft-ietf-l2vpn-vpls-mcast)
- ☐ EVPN (draft-ietf-l2vpn-evpn)
- ☐ Labels in BGP (RFC 3107)
- ☐ VPLS Auto-discovery/Signaling with BGP (RFC 4761)
- ☐ TDM services (with or w/o MPLS-TP)
- ☐ Segment Routing (draft-filsfils-rtgwg-segment-routing)
- ☐ Enhanced ECMP and Large FAT (draft-yong-pwe3-enhance-ecmp-lfat)
- ☐ MRT-FRR (draft-ietf-rtgwg-mrt-frr-architecture)

**Thank you!**

**Q & A**