

# BGP Free Edge

NTT Multimedia Communications Laboratories, Inc.

Robert Raszuk

2012.10.25



*SEAMLESS CLOUD FOR THE WORLD*

Expand your business. Globally.



**Global ICT Partner**  
Innovative. Reliable. Seamless.

# Introduction

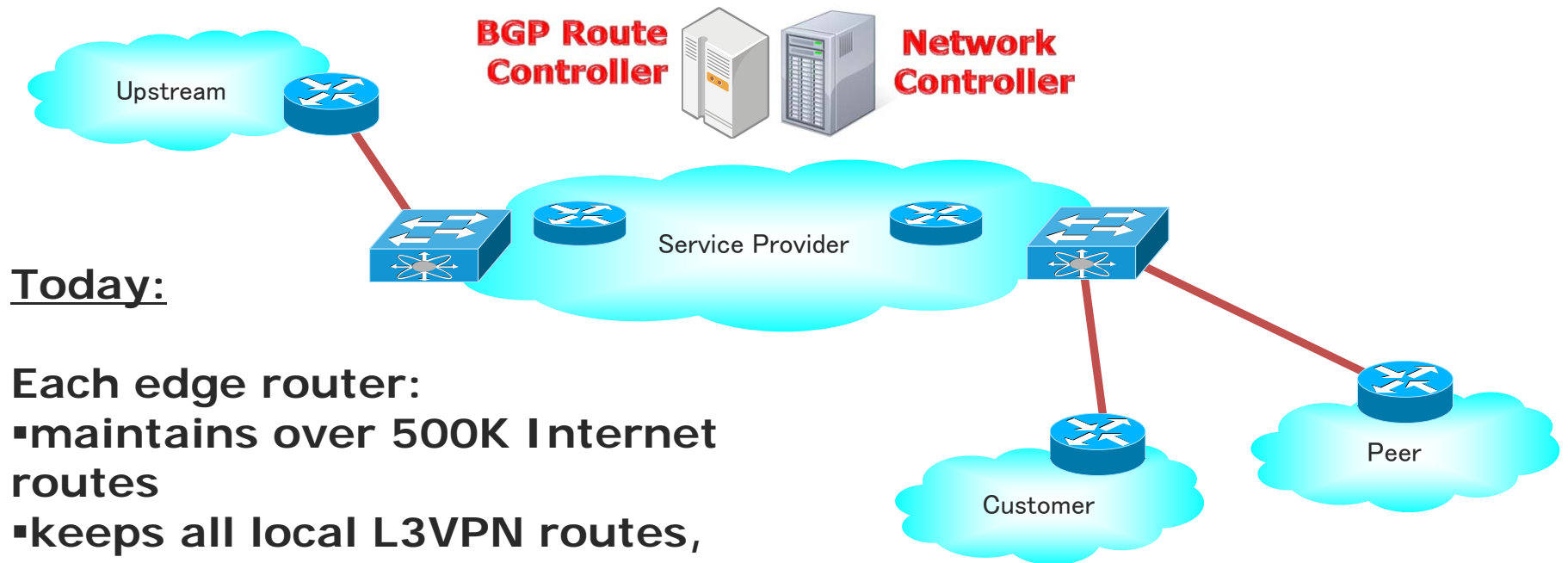
---

- **BGP Free Core is becoming popular among network operators who run some form of encapsulation in the core. Motivations:**
  - **Simplified core architecture**
  - **Lower cost of core infrastructure**
  - **Increase in core speed**
  - **Simplified core management**
  - **Better control on traffic patterns in the core**
  - **Direct preparation for optical switching**

# Introduction

- BGP Free Edge can revolutionize the edge network architectures.
  
- Motivations:
  - Simplified edge architecture
  - Lower cost of edge infrastructure (commodity switches)
  - Increased in edge speed, **increase in new services and application adoption**
  - Enabled by **control plane and data plane decoupling** (OpenFlow/SDN)
  - Flexible traffic classification match into core pipes (L2..L7)
  - Direct preparation for optical switching in the core

# 10000 meters view ...



## Today:

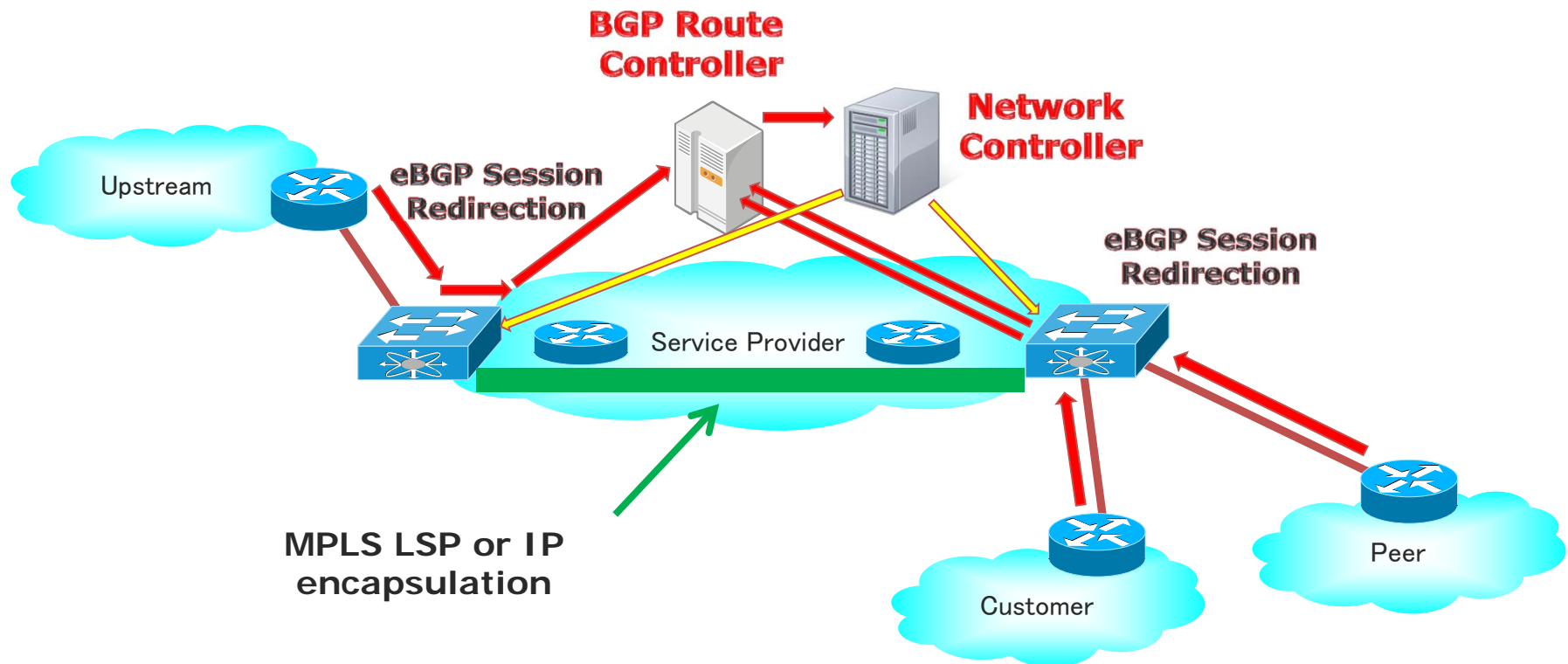
Each edge router:

- maintains over 500K Internet routes
- keeps all local L3VPN routes,
- keeps more than best path,
- connects to dual Route Reflectors
- needs to be configured at each service change
- is difficult to innovate with customized service introduction

## Tomorrow:

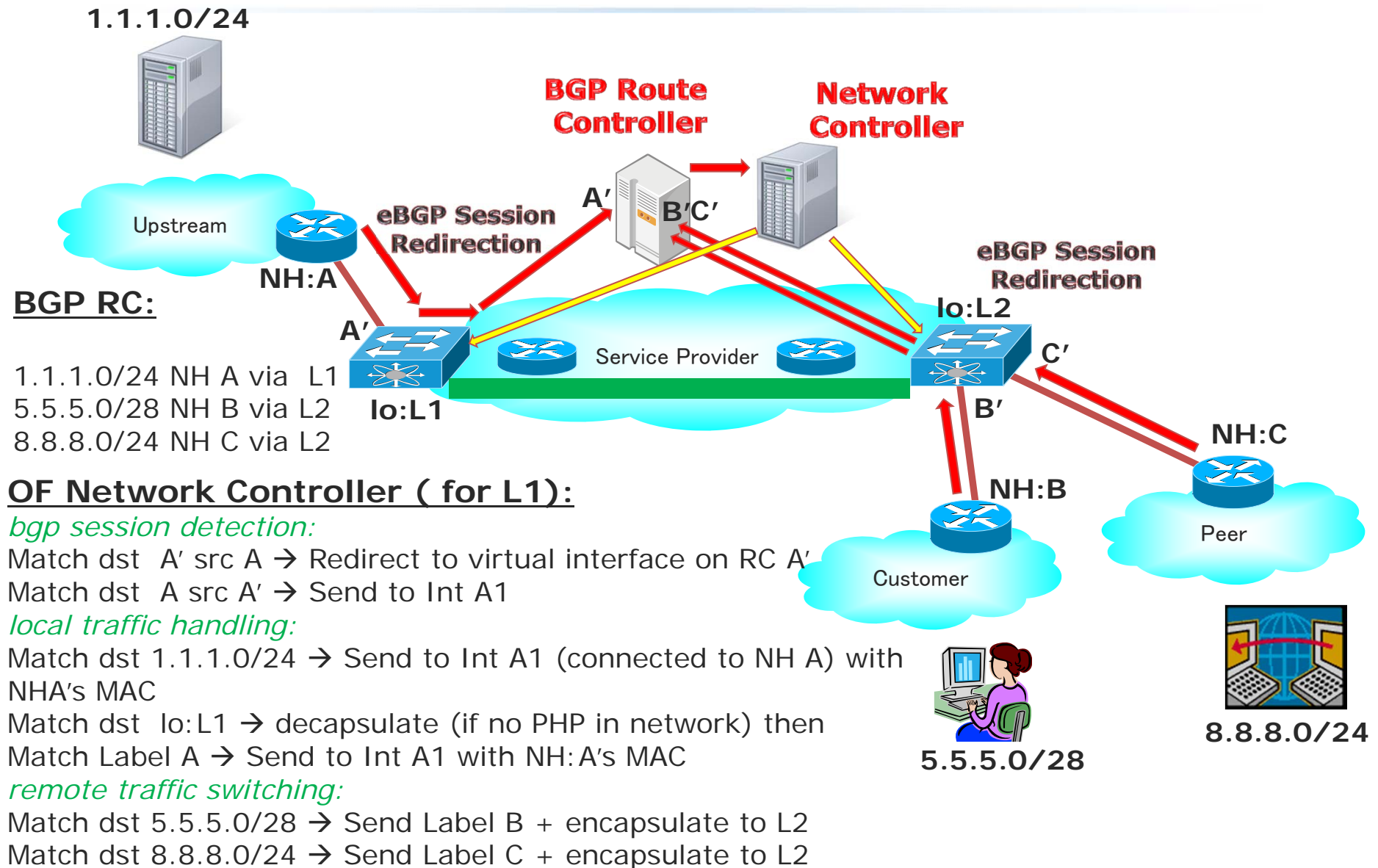
Same service level for customers, no difference from today's operation. Separated control plane from data plane – stored and processed on **clusters of computing devices**

# Zoom in - BGP sessions – No shim

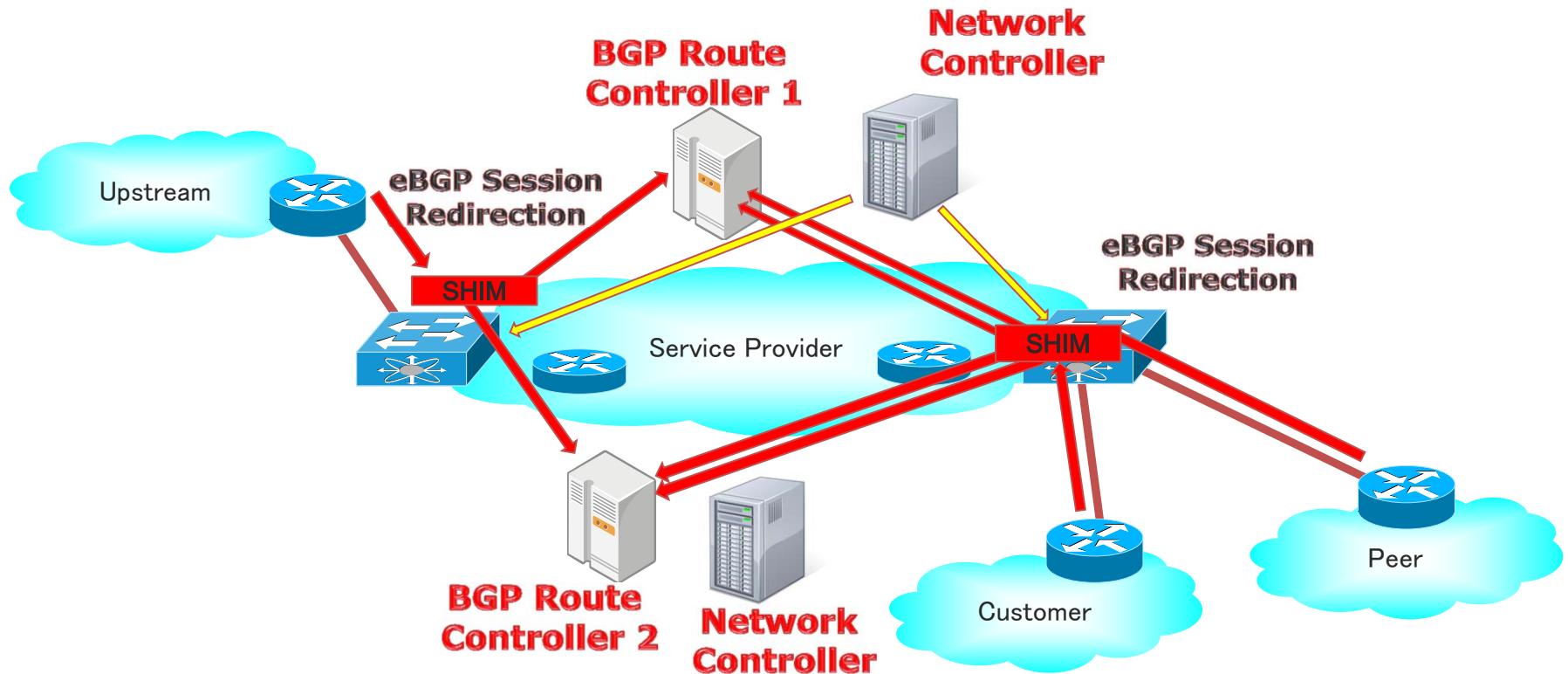


- Incoming BGP sessions is recognized (using proactive match rule) by OF controlled Edge switch and forwarded to BGP Route Controller in the network
- Functional equivalent of L3VPNs per CE label allocation mode with greatly simplified OF based edge configuration and customer's route handling by Route Controller (with multi instance BGP, VRF-lite or hybrid)

# Zoom in - BGP sessions – No shim

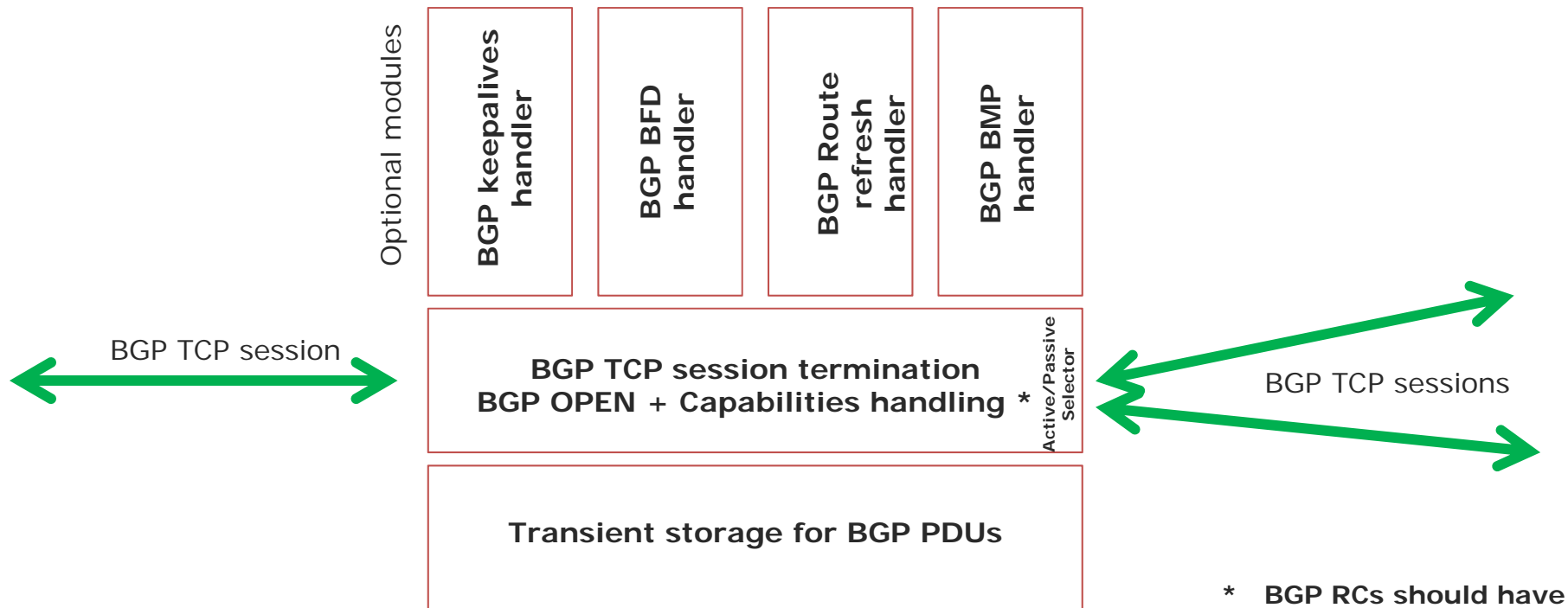
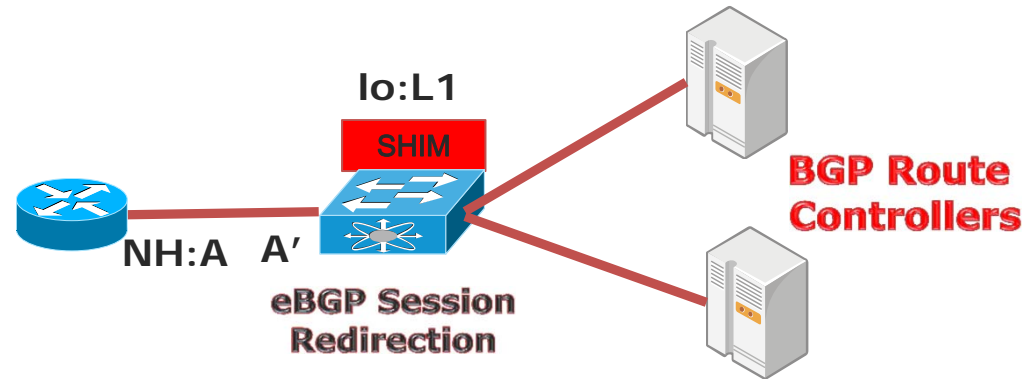


# Zoom in - BGP sessions – With shim



Incoming eBGP TCP sessions are recognized (proactive match rule) by OF controlled Edge switch, terminated by local shim and relayed to redundant BGP Route Controllers in the network (analogy to Route Reflectors)

# Zoom in - SHIM



\* BGP RCs should have identical capabilities



# Zoom in - SHIM

## Active/Passive switchover BGP state synchronization:

### Option 1: Use of BGP Enhanced Route Refresh Mechanism

### Option 2: Use of GR:

- Active SHIM client goes down
- SHIM triggers GR procedure on eBGP peers by artificial performing session bounce
- SHIM elects new active shim client and switches over to it
- SHIM forces complete route refresh to be received from the newly elected active SHIM client
- Upon reception of EOR all eBGP speakers purge all paths which the newly elected active shim client would differ with in comparison with previous active SHIM client

### Option 3: No GR

- Case the same as today where no GR/NSR is configured.
- SHIM can run NSR where redundant CPUs/RPs are available

**Note: Best if peers support either BGP enhanced-route-refresh or GR.**

# Challenges

- **Same or better network robustness**
  - Control plane centralization carries the risk of higher probability of network failures
  - Architecture must provide not worse than today network robustness
- **Solutions:**
  - Distributed control plane with self correlation
  - Multi-vendor control plane + in-house development
  - Add-on TCP session level SHIM to replicate incoming EBGP sessions to multiple control plane units (optional)
  - Move to concept of Control Plane in the Cloud model
  - Applicable both to BGP as well as to OpenFlow controllers
- **Open Flow session redundancy on the switches mandatory – full proactive mode provisioning only**

# Challenges

- **Dynamics of FIB download rate from Network Controller to OpenFlow enabled switches ... not a challenge to BGP Free Edge but to OpenFlow technology**
  - **Can be addressed by elimination of unnecessary programming of forwarding entries by using techniques of simple-va**
  
- **Requirement for BGP enhanced route refresh or GR support by peering routers as well as BGP Route Controller**
  - **Required for data plane non-disruptive switchover between Active and Passive BGP Route Controllers**
  - **Simpler alternative for otherwise much more complex SHIM design which would need to parse and keep BGP prefixes and associated attributes**
  
- **Partitioned network – Switch loses communication with all controllers**
  - **A problem for any solution which is centrally managed even RRs today – not specific to BGP Free Edge. Solution could be OpenFlow GR with relatively long timer.**

# Benefits

- **Very easy new feature introductions. Examples of real benefit:**
  - **Very easy and unified introduction and management for new network services – much faster revenue opportunities**
  - **BGP Origin Validation processing moved from each edge router to BGP central control plane**
  - **BGP Sec signature validation and crypto processing moved from each edge BGP router to a central computing server in the cloud**
  - **Flexibility to calculate customized BGP best paths for not only each ingress point, but also per customer basis**
  - **Scalable Internet as a VPN**
  - **Services separation over the same infrastructure**
  - **Significant reduction in BGP Wave Effect – Helps Internet stability, by elimination of leaking intra-AS path transitions**

# Benefits continued ...

- **Elimination of redundant state in the RIB and FIBs of network elements with overlapping prefix suppression (Simple-VA) executed on RC**
- **No more opex and capex to upgrade CPU and memory network wide for control plane grow – sometimes it is not possible and vendors ask to fork lift upgrade entire network infrastructure. Example new platforms new control plane operating systems ....**
- **Very easy BGP monitoring and reporting tools interface – no need to collect all BGP “raw” feeds (aka original intention of BMP) from all border routers – API from BGP RC will enable view of entire BGP in an Autonomous System**
- **Centralized BGP policy management**

---

# Questions ?

[rr@nttmcl.com](mailto:rr@nttmcl.com)