The Two-dimensional Nature of the Core Scaling Problem

Dimitri Papadimitriou - Alcatel-Lucent Bell NV

<dimitri.papadimitriou@alcatel-lucent.be>

Workshop on "Management of Network Virtualisation"

Organised by the AGAVE and EMANICS IST projects

with the support of the European Commission

6 November 2007

Outline

Internet traffic and routing system growth

Problem statement

Concept

Approach

- Transit AS
- Stub-AS

Architecture

- Forwarding plane
- Control plane

Benefits

Conclusion

Disclaimer: the problems, ideas and orientations discussed here are the thoughts of the author and may or may not represent future Alcatel product direction

Expansion of Internet between 2005 and 2006

Prefixes: 173,800 – 203,800 (+17%)

AS Numbers: 21,200 – 24,000 (+13%)

Average advertisement size is getting smaller (8,450 – 8,100)

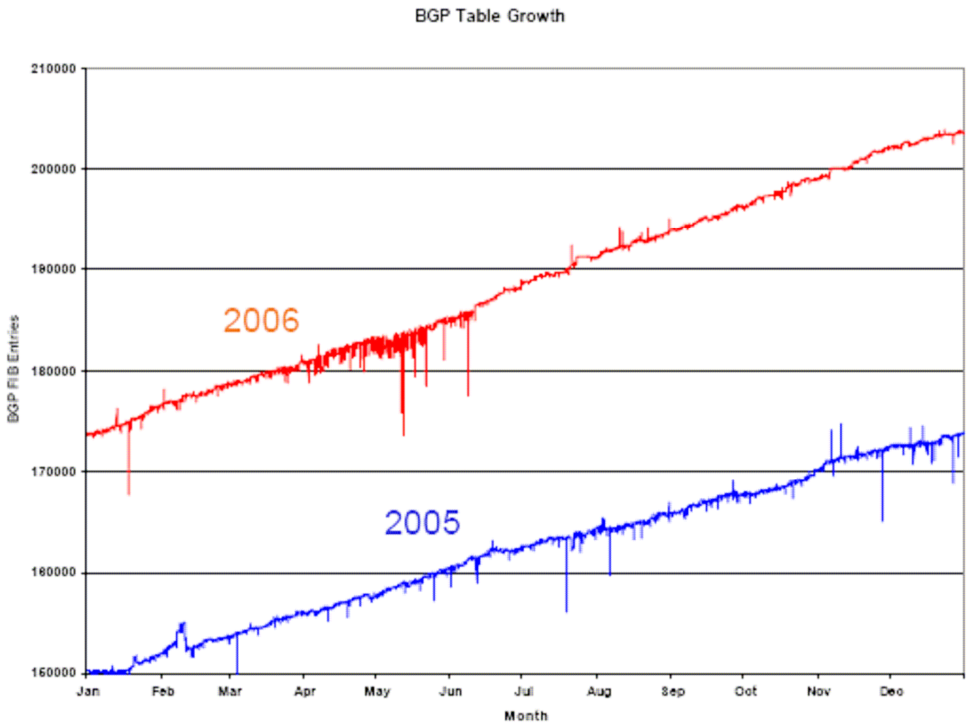
Average address origination per AS is getting smaller (69,600 – 69,150)

Average AS Path length steady (3.4)

AS transit interconnection degree rising (2.56 – 2.60)

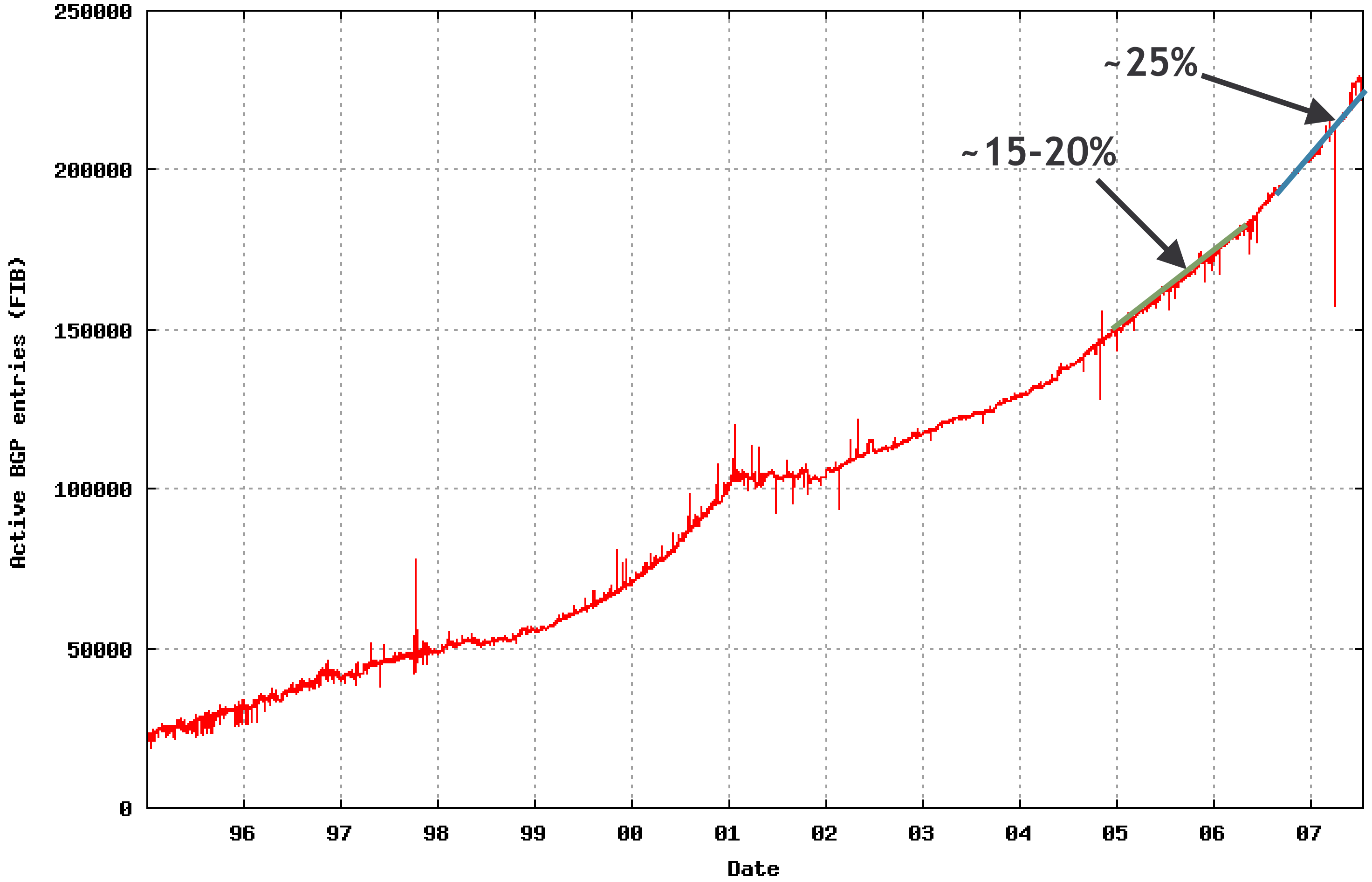
=> IPv4 network becomes denser (more interconnections), with finer levels of advertisement granularity (more specific advertisements)

Total Advertised BGP Prefixes



Source: IEPG, <http://www.potaroo.net>

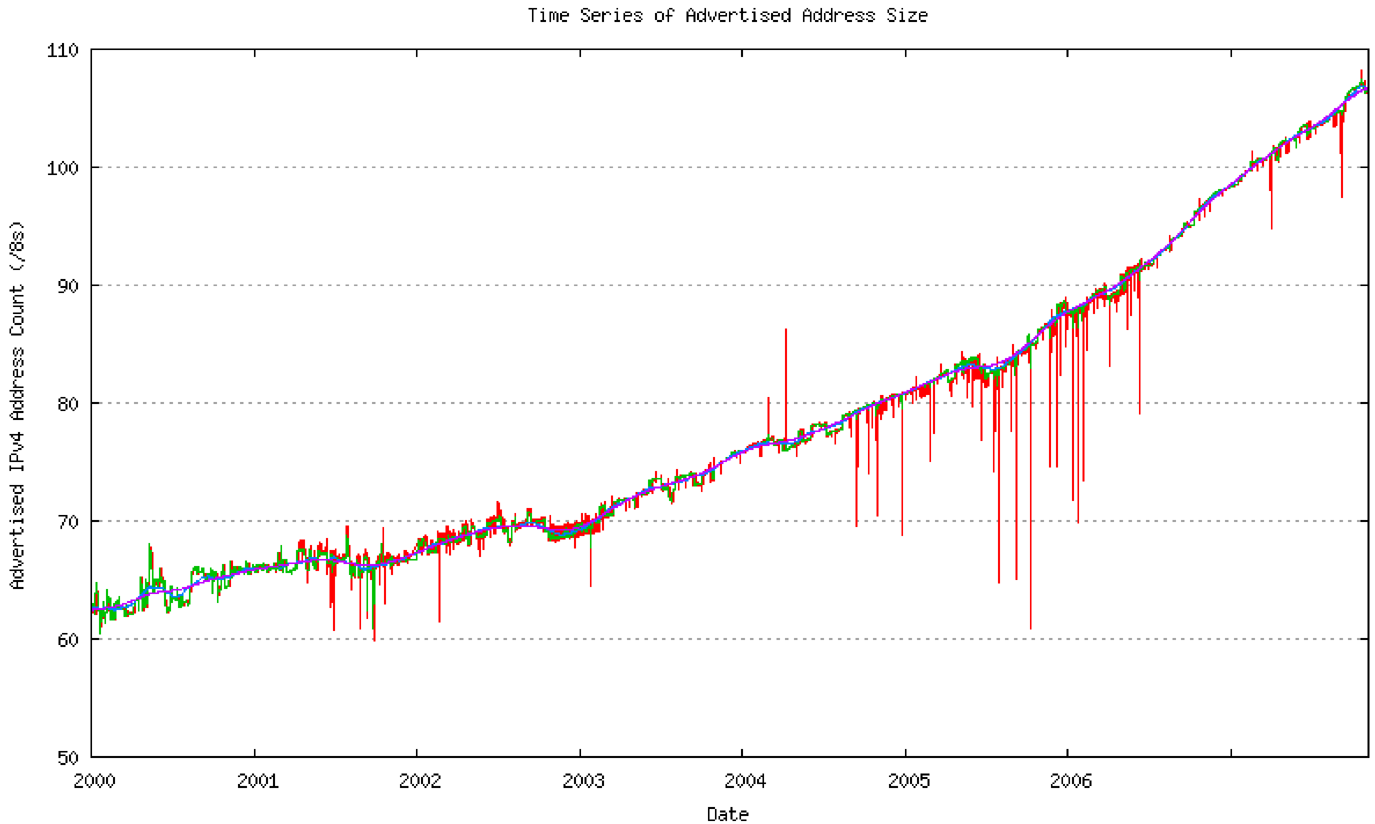
Growth of the BGP Table (from 95 to mid'07)



Source: BGP Routing Table Analysis Reports - http://bgp.potaroo.net

Advertised IPv4 Addresses

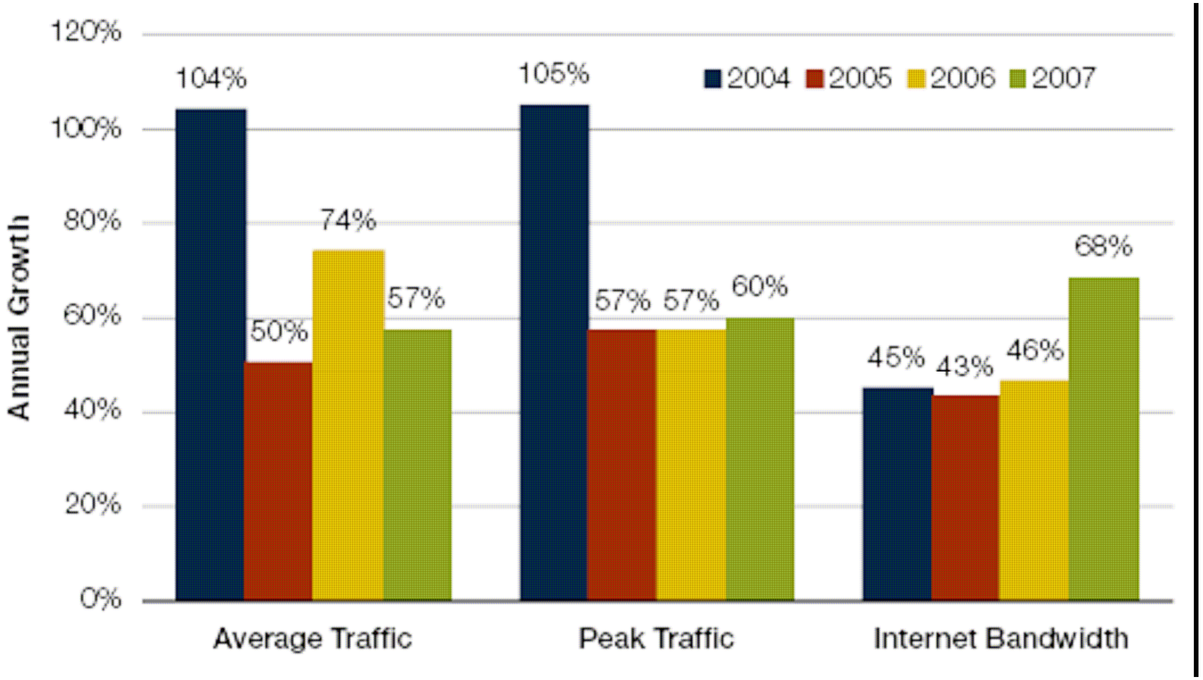
- Total span of address space advertised in BGP routing table since 1999
- Day-by-day sequence (all sample values recorded over a day are averaged into a single daily value)
- Daily average sequence smoothed by applying a sliding window of 93 days average across the sequence in two passes



Source: IPv4 Address Report, http://www.potaroo.net/tools/ipv4/index.html

Internet Traffic Growth

International Internet Traffic and Bandwidth Growth, 2004-2007



Source: Global Internet Geography, TeleGeography research, 2007

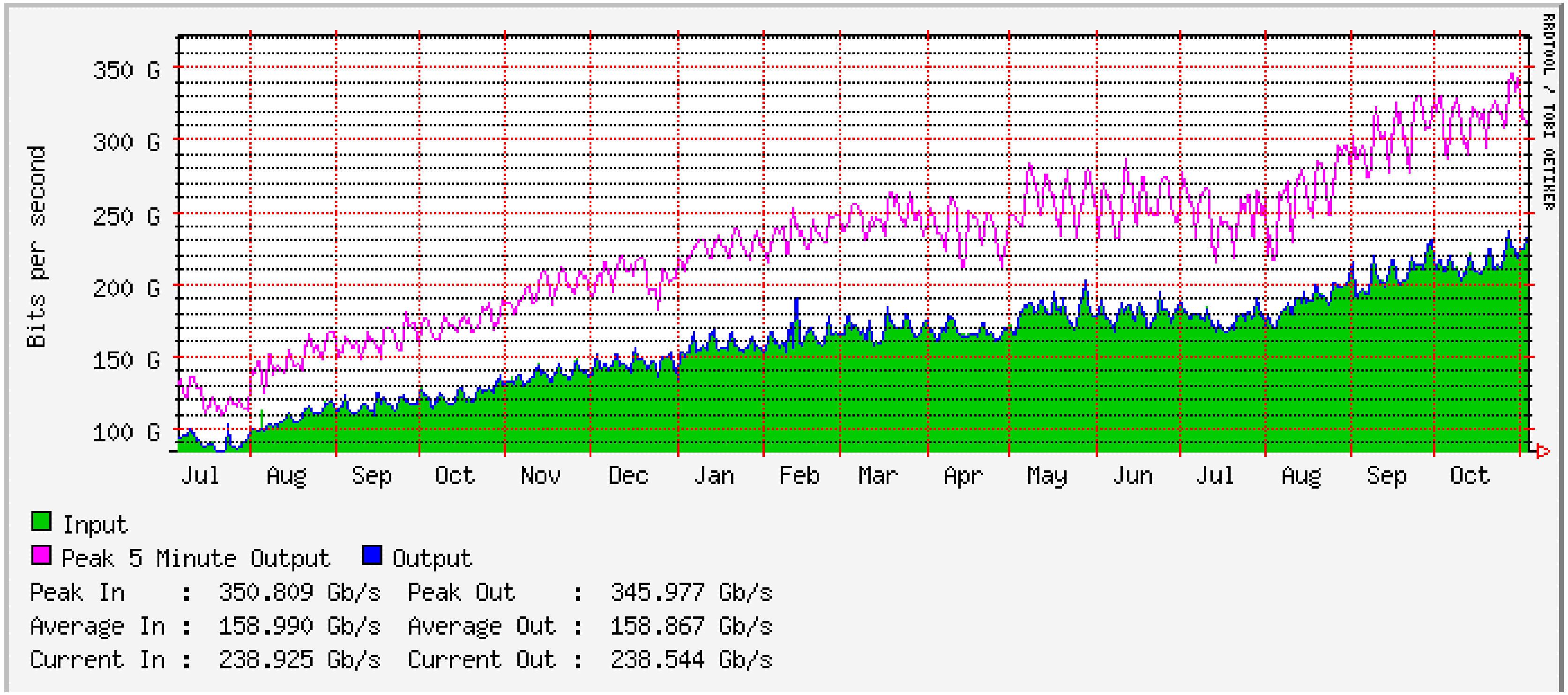
Notes: (mid-year) data reflect traffic over Internet bandwidth connected across international borders

Intra-European Internet traffic grew 85 % in 2006 and 71 % in 2007

Current (mid-2007) annual Internet traffic growth rates : between 50-**100%**

Example: European IXP

AMS (Amsterdam) - Internet eXchange (AMS-IX)



Source: http://www.ams-ix.net

Problem statement

In core / large-scale architectures for packet networks

- if control plane / traffic is aggregated, then it is aggregated on the same platform that aggregates data plane / traffic
- imposes set of two–dimensional requirements on that platform

⇒ platform must scale in terms of bandwidth and throughput

+ protocol messaging and processing

Consequences

- Routings platform (in part. core routers) must include state–of–the–art capabilities for both dimensions
- Cost and complexity of platform

Problem: how to address/reduce the impact of the two–dimensional nature of core scaling (traffic growth ~5 times more important than routing table growth)

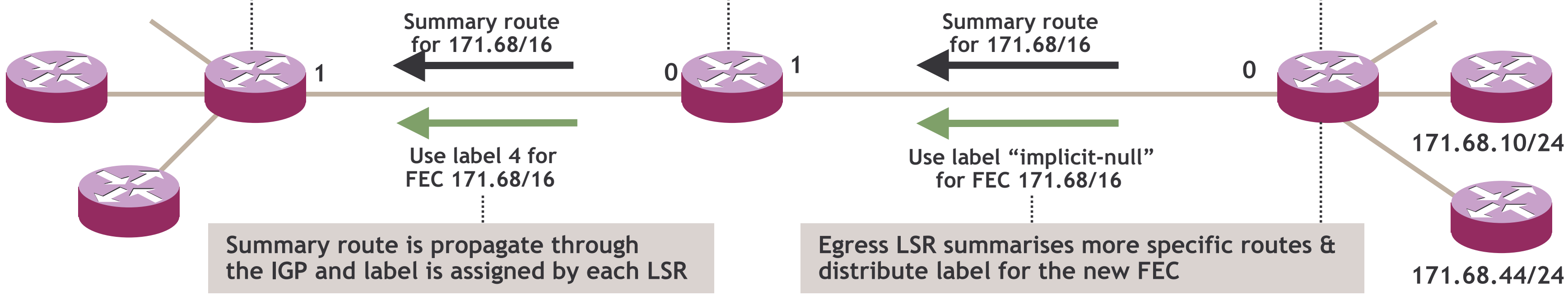
MPLS and the promises …

1. Forwarding plane scalability - "multipoint-to-point aggregation trunks"

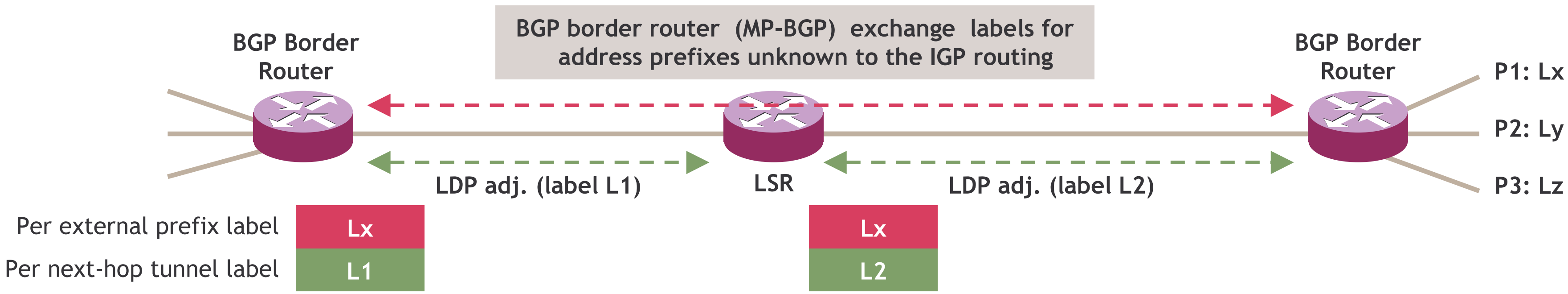
In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	-	171.68/16	1	4
…	…	…	…	…

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	4	171.68/16	2	pop
…	…	…	…	…

Address Prefix / mask	Next-Hop	Interface
171.68.10/24	10.68.9.1	Serial1
171.68.44/24	10.68.12.1	Serial2
171.68/16	…	Null



2. Control plane scalability - "transit tunnels" between border routers



… main issue: traffic nesting into a non-decomposable label sequence

Concept

Decouple control/routing from forwarding plane aggregation functions

- As traffic increase vs routing entries
- As number of AS increases (periphery)
- Path remaining sensibly identical (length)

Transit AS needs to accommodate more traffic with less increasing number edges/routes

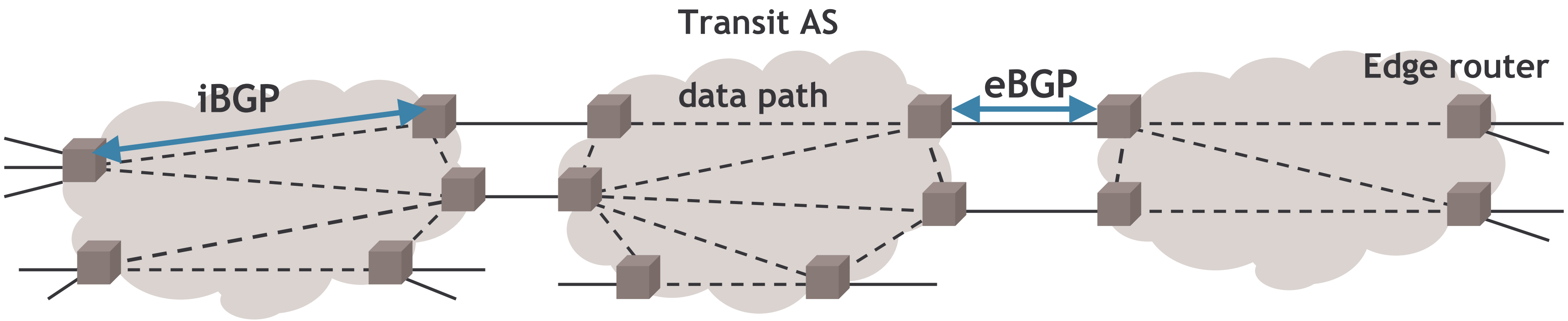
Motivations / Drivers

- Technical complexity associated to each aggregation problem can be addressed separately
- Each aggregation problem can be addressed with a specific, rather than generalized platform (potential cost reduction)
- Differences in expansion rates in logical and physical space are no longer dependent
  - Internet traffic growth: ~ 50-100% per year
  - Routing table growth: ~ 20% per year

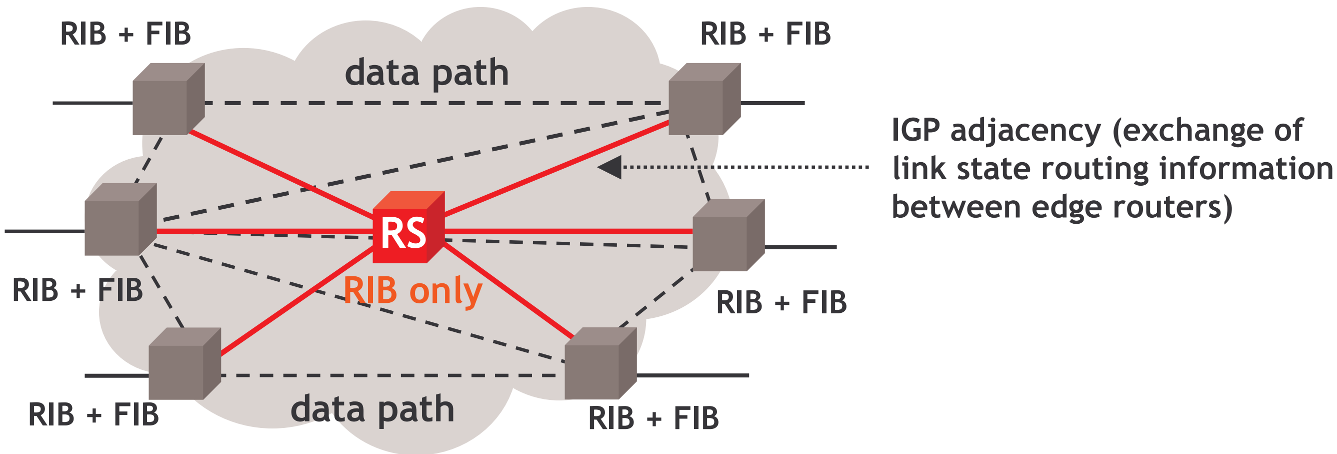
=> Does not require upgrading both the physical and logical scaling platforms at the same time, as they are no longer linked

Approach (1) - Transit AS

Inter-domain routing level



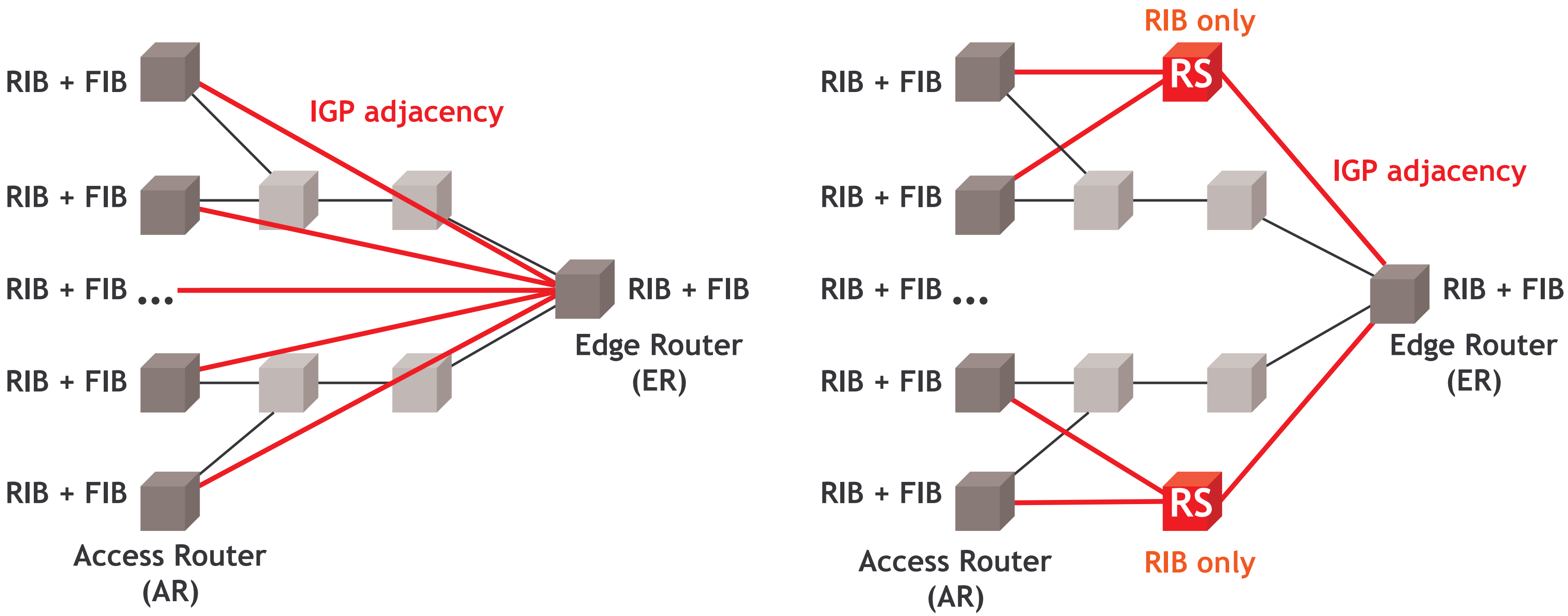
Intra-domain routing level



RS IGP routing information distribution server (stand-alone routing engine i.e. RIB + IGP link-state routing protocol)

Approach (2) - Stub-AS

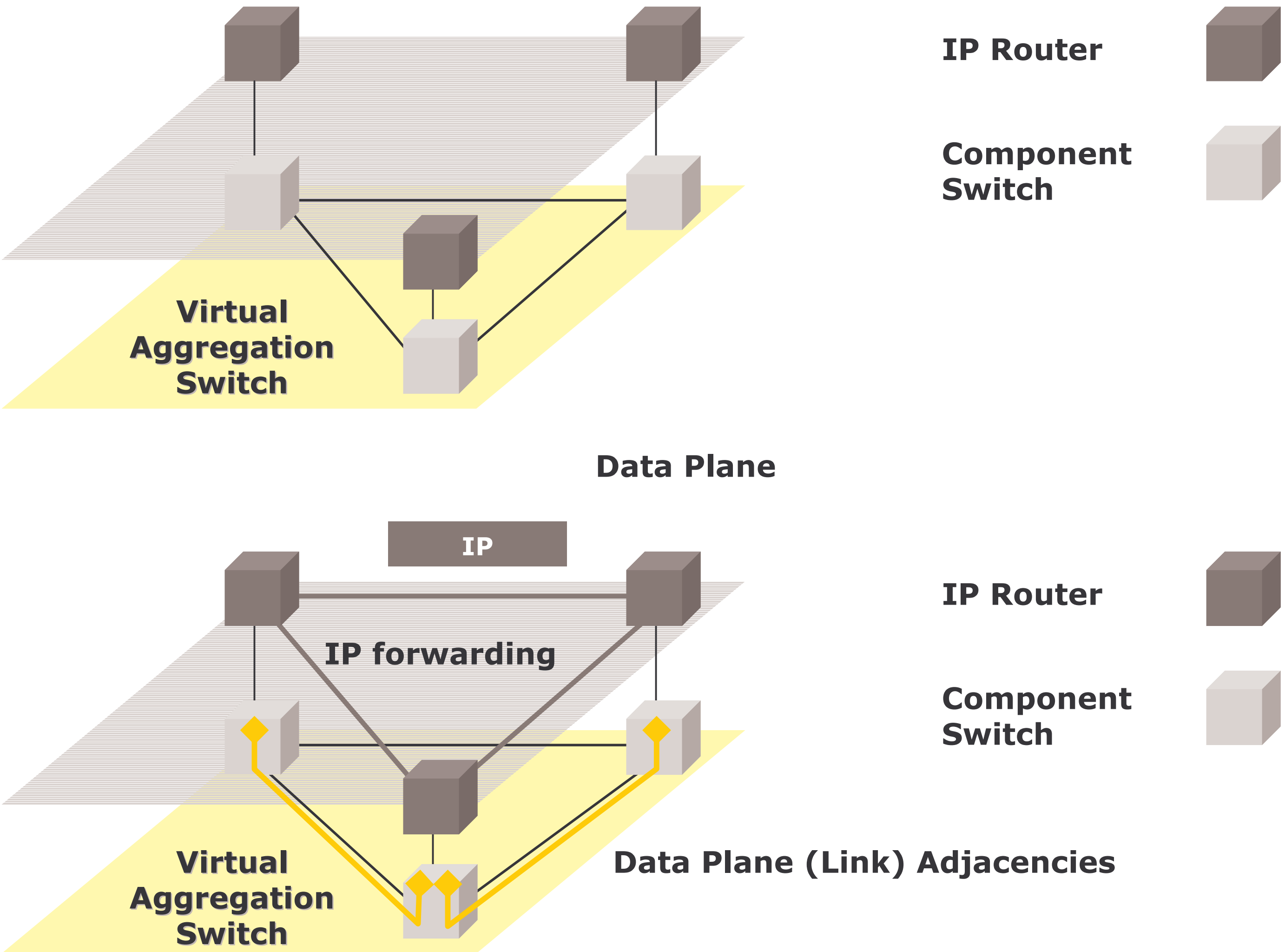
The same approach can be beneficial for stub-AS



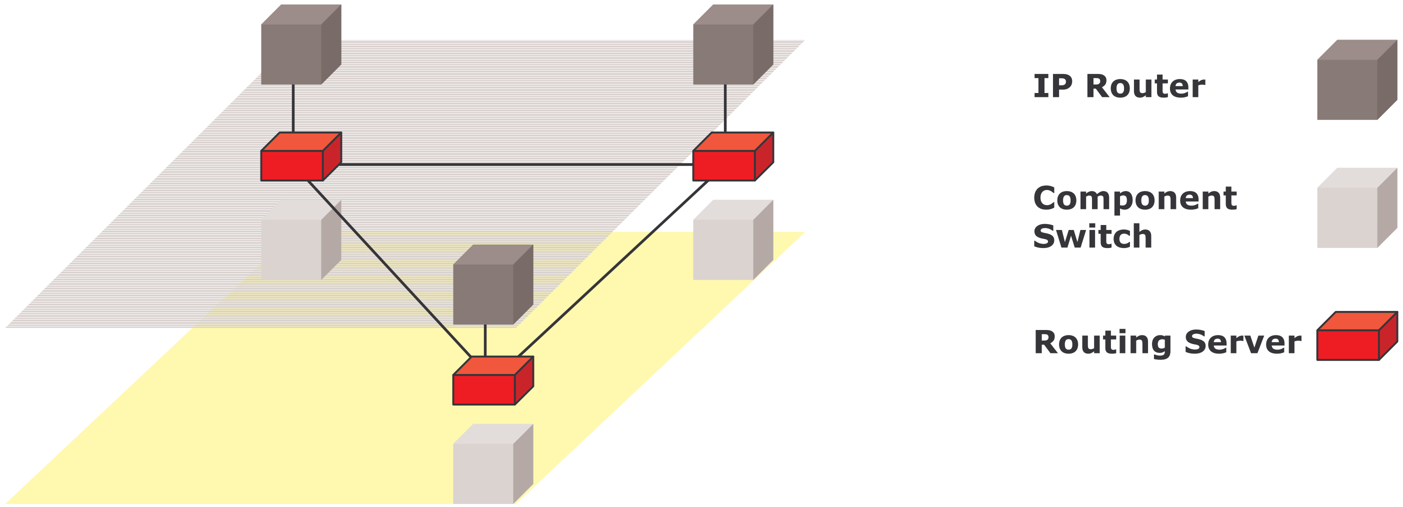
O(1k) AR => O(1k) routing adjacencies at ER

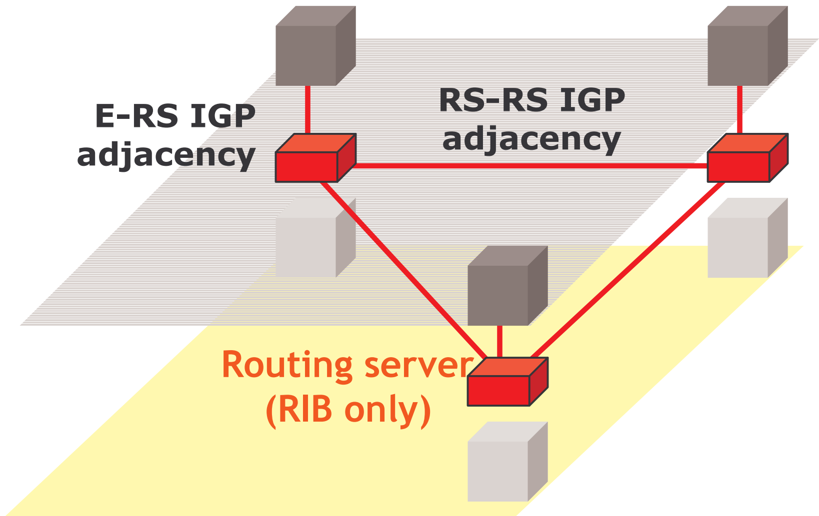
O(1k) AR and O(#RS) routing adjacencies at ER

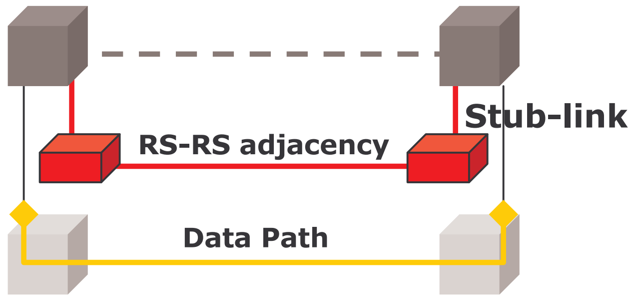
Architecture: Forwarding plane and adjacencies



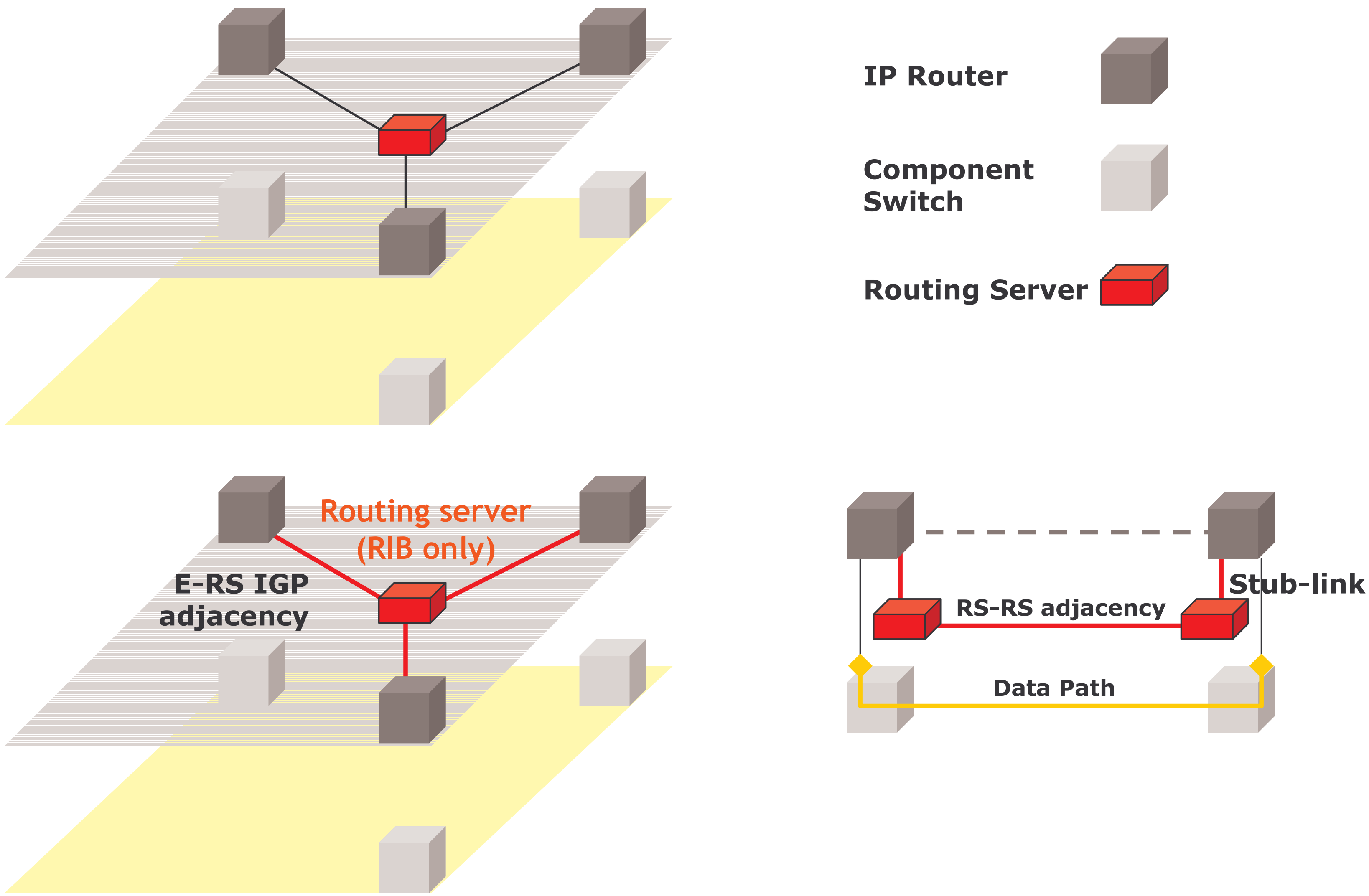
Architecture: Routing plane and adjacencies (distributed)







Architecture: Routing plane and adjacencies (centralized)

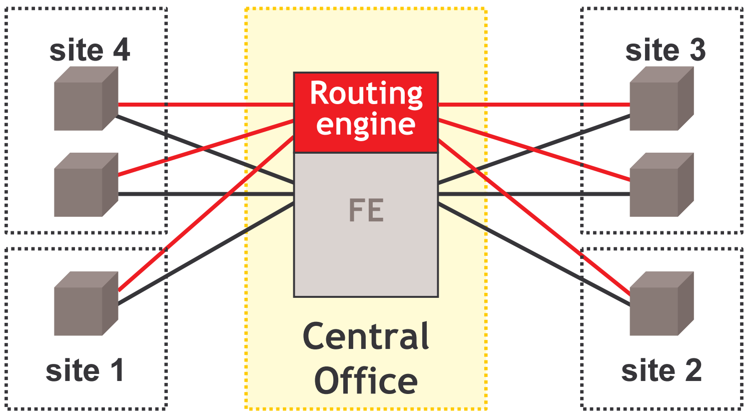


Benefits (1)

RS acting as IGP routing information “re-director” : IGP routing information exchanged via established adjacencies with peering routers (routing plane level)

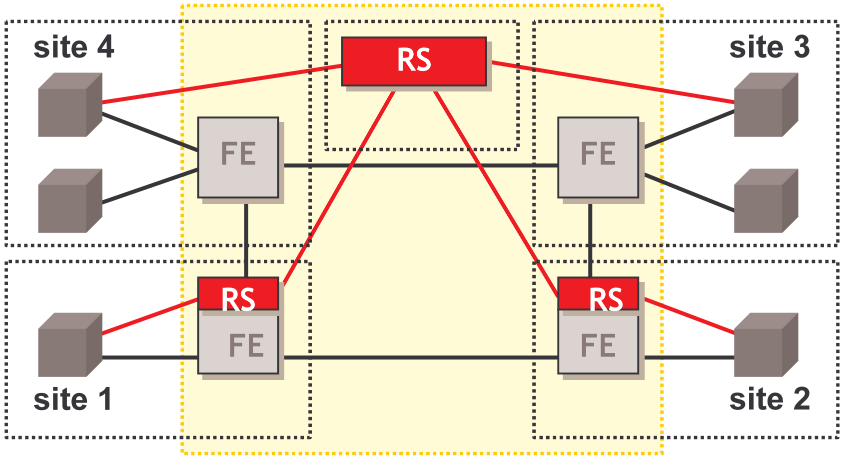
=> Forwarding capacity vs routing capacity differences in expansion rates in both logical and physical spaces are no longer dependent

With classical core router



Core switch capacity N

With distributed core router



Distributed switch
Total cap. N

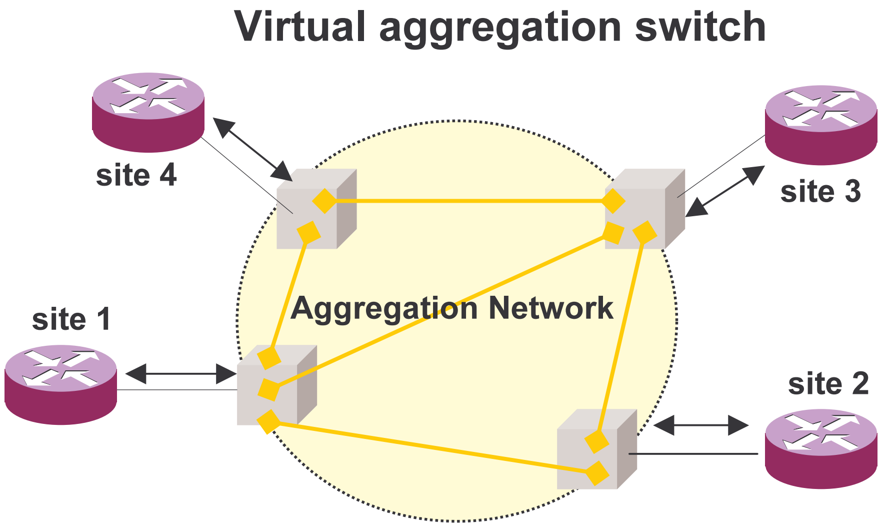
FE: Forwarding Engine

Benefits (2)

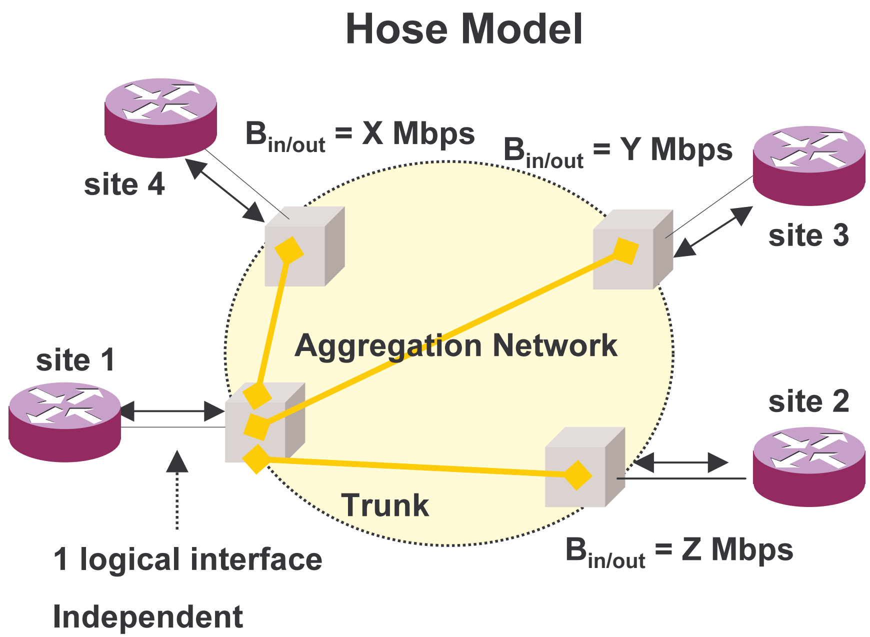
Advantages

- Reduced routing system scaling requirements
- Increased robustness/stability and resiliency
- Ability to instantiate multiple IP networks (using MT routing) relying on the same aggregation network, *without fate-sharing their control planes*
- Preserve traffic engineering of router-to-router flow in the network (using aggregates) whilst providing
  - Advantages of an IP network
  - Original control plane separation between IP and transit/aggregation network
- Within aggregation network
  - IP routing plane protocols (new paradigm possible)
  - Commoditized interfaces to the IP routers connected to it

Virtual aggregation switch - Hose model



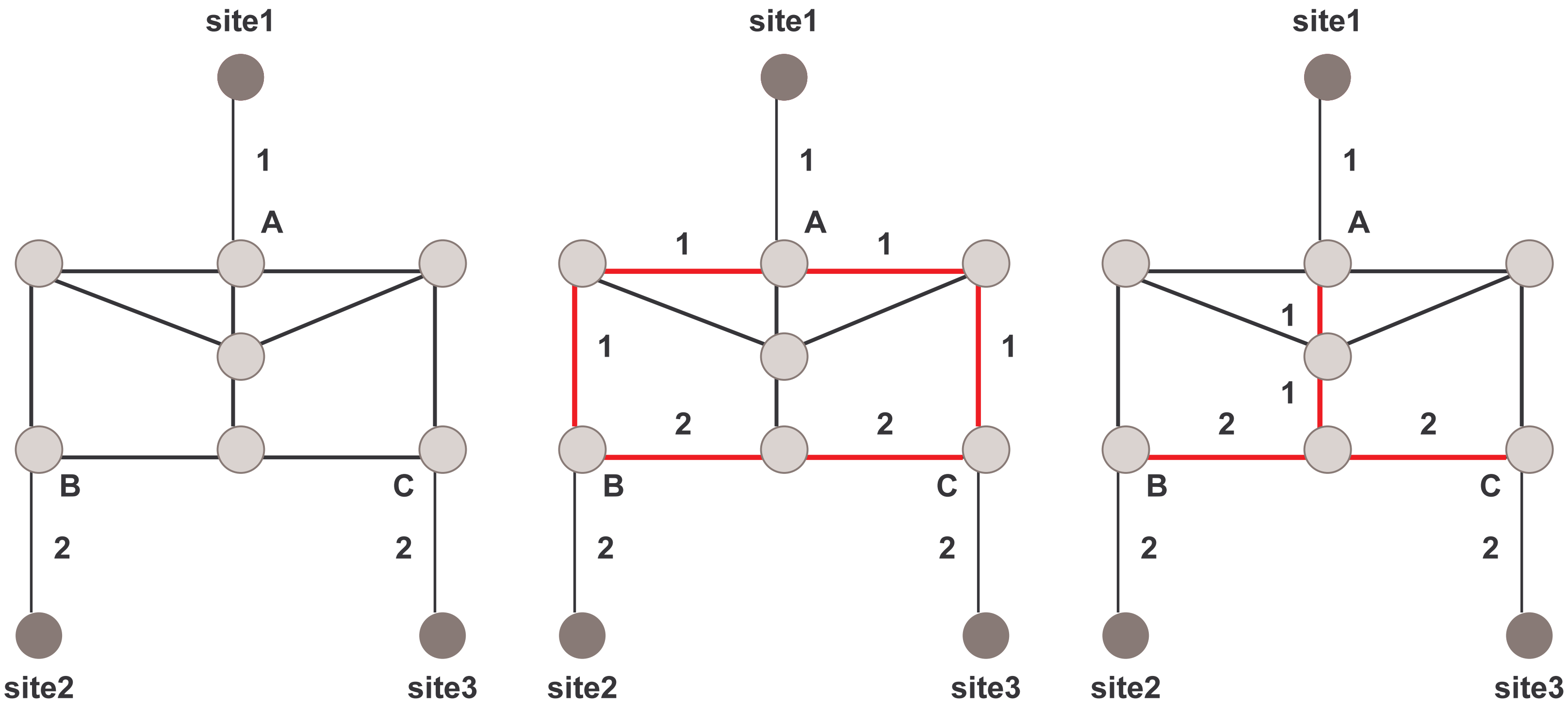
Paradigm: in order to conserve bandwidth and realize the multiplexing benefits of the hose model, paths entering into and originating from each hose endpoint need to share as many links as possible



Hose model advantages over virtual pipe model (= MPLS)

1. **Simplicity:** only one ingress and egress bandwidth per endpoint to be specified, compared to bandwidth for each pipe between pairs of endpoints
2. **Flexibility:** traffic to and from endpoints can be distributed arbitrarily over other endpoints as long as the ingress and egress bandwidths of each endpoint are not violated
3. **Multiplexing Gain:** due to statistical multiplexing gain, hose ingress and egress bandwidths can be less than the aggregate bandwidth required for a set of point to point pipes
4. **Characterization:** requirements easier to characterize because the statistical variability in the individual source-destination traffic is smoothed by aggregation into hoses

Example



Independent Shortest
Paths
Total = 8 units

Link sharing
Total = 6 units

Conclusion

To reduce the two-dimensional nature of the core scaling problem => decouple routing from forwarding plane aggregation

RS acting as IP routing information (distributed) server => forwarding vs routing capacity differences in expansion rates in logical and physical space are no longer dependent

Retain benefits of both IP traffic engineering and original control plane “separation” of overlay networks

Core routing without core router:

- Approach applicable to larger scale IP networks with e.g. Ethernet as (intra-domain) aggregation technology
- Maintains distributed traffic aggregation (no hyper-node aggregation)

  => robustness and resiliency against both node and link failure

Acknowledgements

This work was carried out within the framework of the IWT TIGER project sponsored by the Flemish government institute for Innovation through Science and Technology in Flanders (IWT)

Thanks !