

MPLS-based traffic shunt

Nicolas FISCHBACH [nico@colt.net]
Senior Manager - IP Engineering/Security
RIPE46 - Sept. 2003



Contributors

- COLT Telecom
 - Andreas Friedrich
 - Marc Binderberger

- Riverhead Networks
 - Yehuda Afek
 - Anat Bremler-Barr
 - Boaz Elgar
 - Roi Hermoni

- Cisco Systems
 - Roy Brooks
 - Paul Quinn

Agenda

- DDoS Protection
- Deployed mitigation methods
- MPLS-based traffic shunt
- Conclusion
- Securing the infrastructure ?
 - To be discussed at the nsp-sec BoF Tuesday evening !

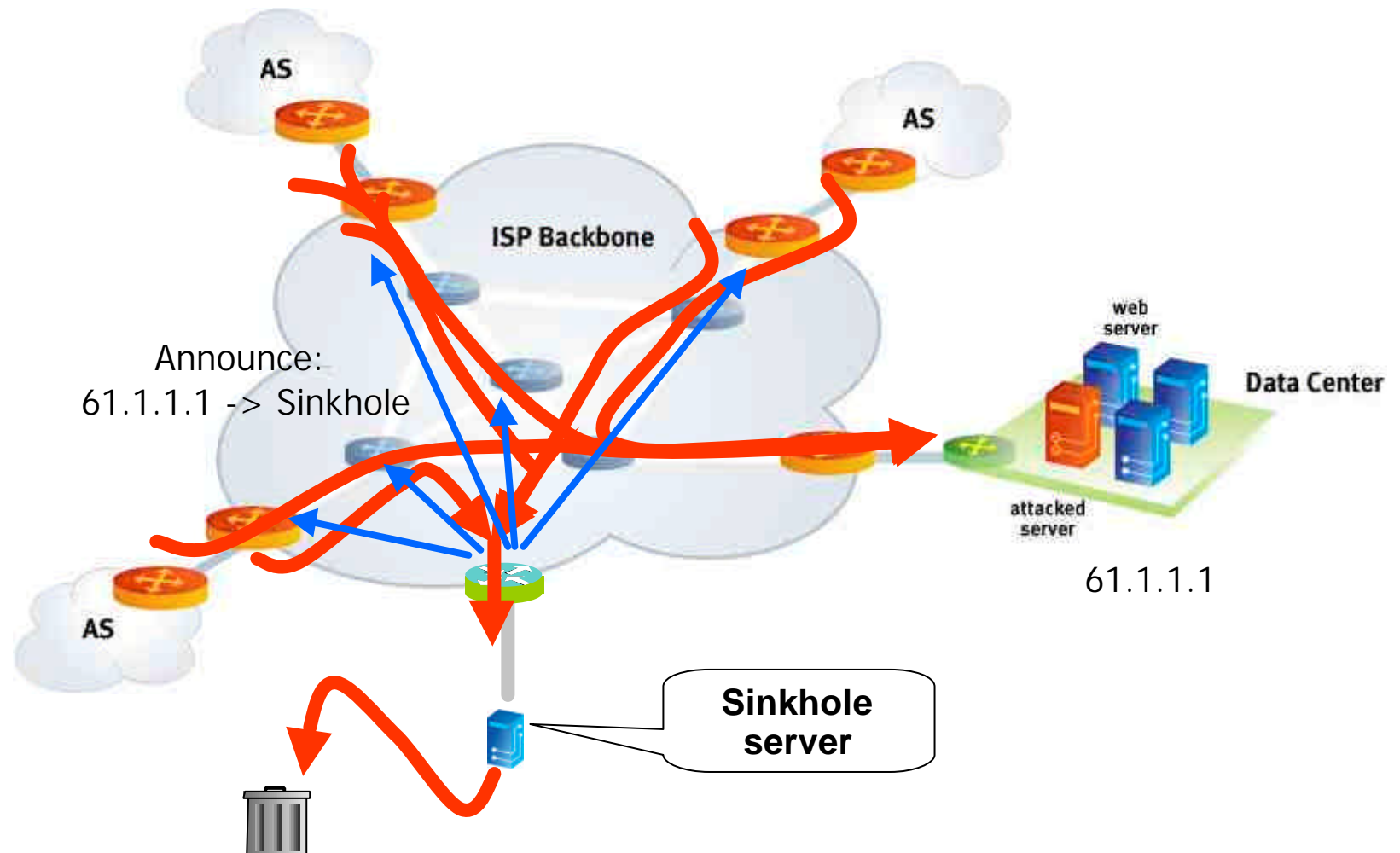
Distributed Denial of Service Protection

- Data-center vs infrastructure approach
- Why strict filtering isn't (always) the answer
 - usually means the attacker “won”
 - some traffic can't be filtered at the router level
 - layer 4+
 - traffic requiring *real* state information (not only “bit is set)
 - after “everything on top of IP” the trend is “everything on top of HTTP”... wanna filter 80/tcp ? ;-)
 - is your network's physical and logical structure enabling you to filter at the Edge and not in the Core ?
 - you are tired of arguing with your network architecture team (“we are here to transport packets” vs “the Internet firewall” ;-)

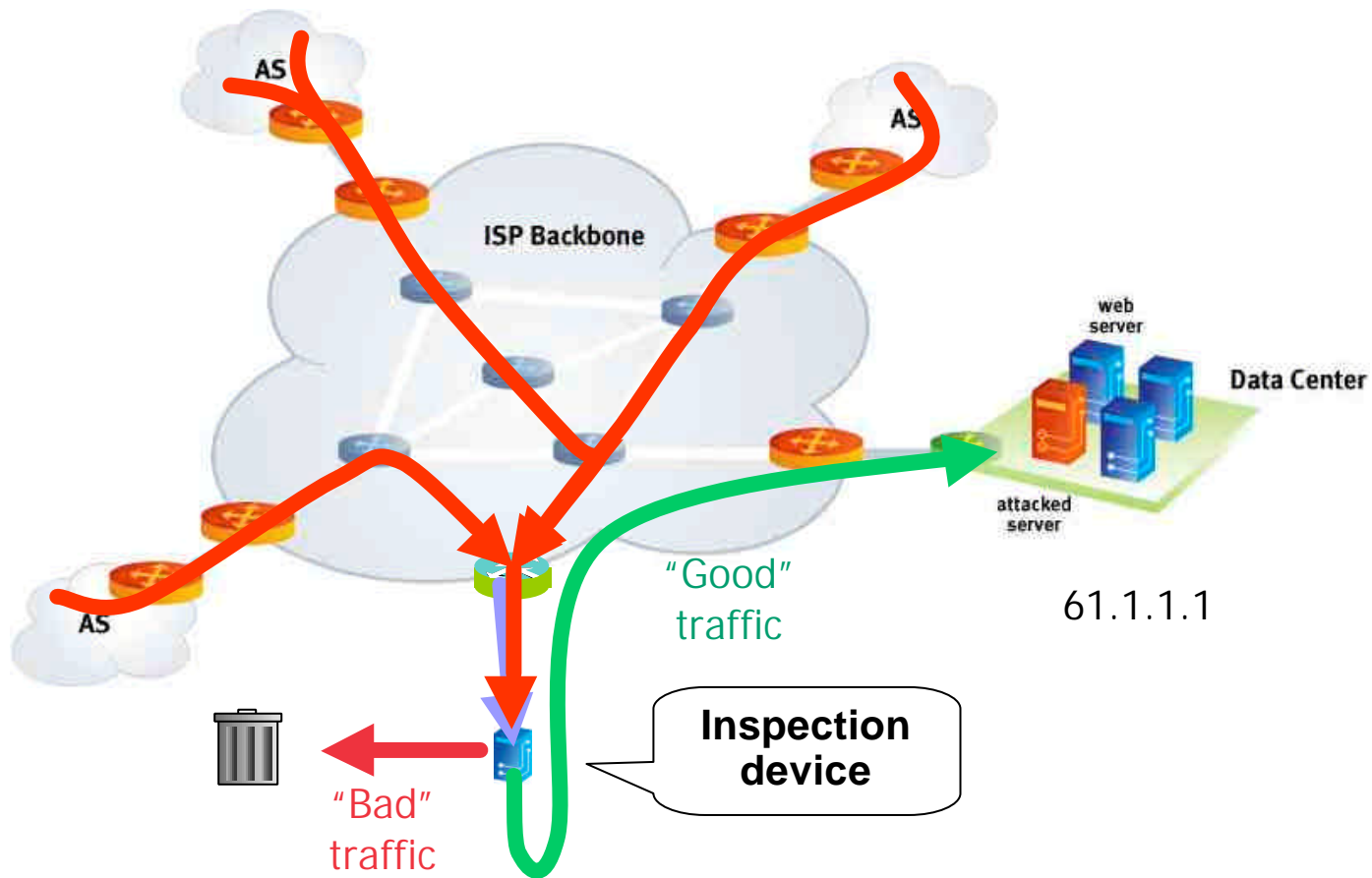
Deployed mitigation methods

- What do/should SPs support/do ?
 - (propagated) blackholing
 - (de-aggregate and) stop to announce - bad practice ?
[dampening, BGP table size, filters, etc.]
 - sinkholes
 - rate-limiting
 - ACLs
 - iACLs (infrastructure)
 - tACLs (transit)
 - re-coloring

Sinkhole



Traffic Shunt



Sinkhole vs Shunt

■ Sinkhole

- Uni-directional
 - Data in, no data out
- IP based
- Blackholing traffic, forensics
- [CenterTrack, NANOG17]

■ Shunt

- Bi-directional
 - Data in, processed and data out
- Tunnels: GRE, MPLS, L2TPv3, etc.
- DDoS cleaning, reserve proxy, traffic analysis
- [Bellwether, NANOG19]

IP-based Traffic Shunt

- Tunnels examples
 - From the peering/upstream routers to the inspection device
 - From the inspection device to the CPE/end-system
 - A mix/combination of both
- Limitations
 - Careful setup required to avoid loops
 - Returned traffic must not pass through a peering router
 - Cisco GSRs and Juniper require a dedicated interface card to act as a tunnel server (GRE/IPIP)
 - Processing overhead

MPLS-based Traffic Shunt

- Advantages

- Doesn't require a special/dedicated interface card
- No extra HW load or SW (IOS 12.0(17)ST+ and JunOS 5.4+)
- If your network is MPLS-enabled, operations knowledge should be there: no need for the network to be MPLS-only! "Normal" routed IPv4 traffic can be carried in parallel
- Minimal (initial) static configuration with dynamic LSPs (iBGP triggered)
- Low (zero ?) overhead [did someone just say "why not use Policy Based Routing" ? ;-]
- A MPLS-speaking inspection device isn't required (option)

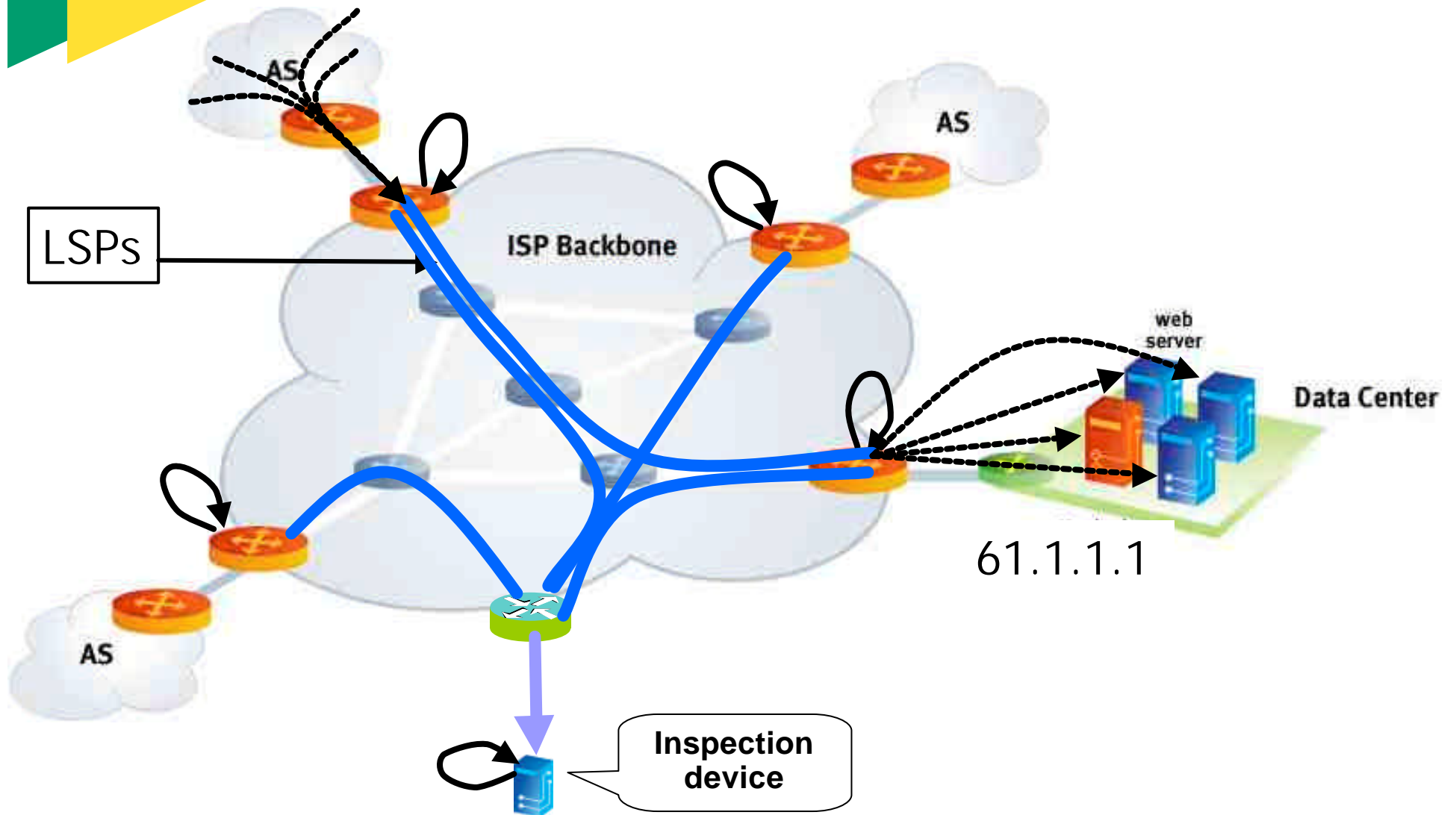
MPLS-based Traffic Shunt

- Advantages (cont.)
 - Enables you to overcome the “this device is in-line only” and “you need one inspection device per peering/upstream)” limitations: profile traffic and (potential) victims, select key POPs/IXes and deploy there
 - Not on the critical path and quite scalable
 - LDP only carries the loopback address of the inspection device
- Caveats
 - You may carry the traffic through the backbone (depending on how distributed your deployment is)
 - Latency: a few more ms (extra hops/distance)
 - Peering Router that also acts as an Access Router (unless you (can) use more specific routes)

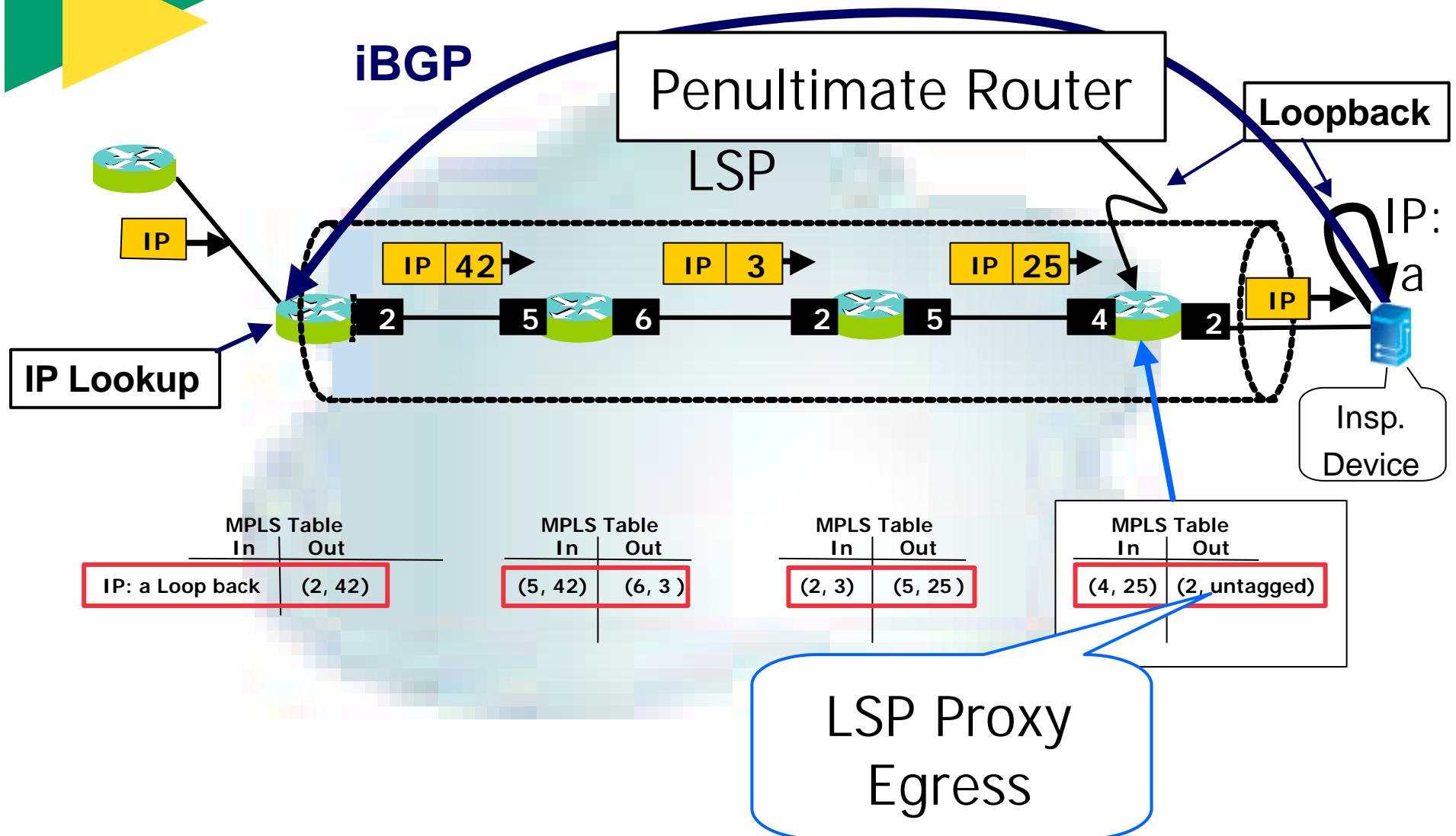
MPLS-based Traffic Shunt

- Two methods
 - Pure MPLS using Proxy Egress LSP (*)
 - Penultimate hop popping
 - RFC 3031
 - MPLS VPNs using VRFs
 - see: <http://www.nanog.org/mtg-0306/afek.html>
[NANOG28]

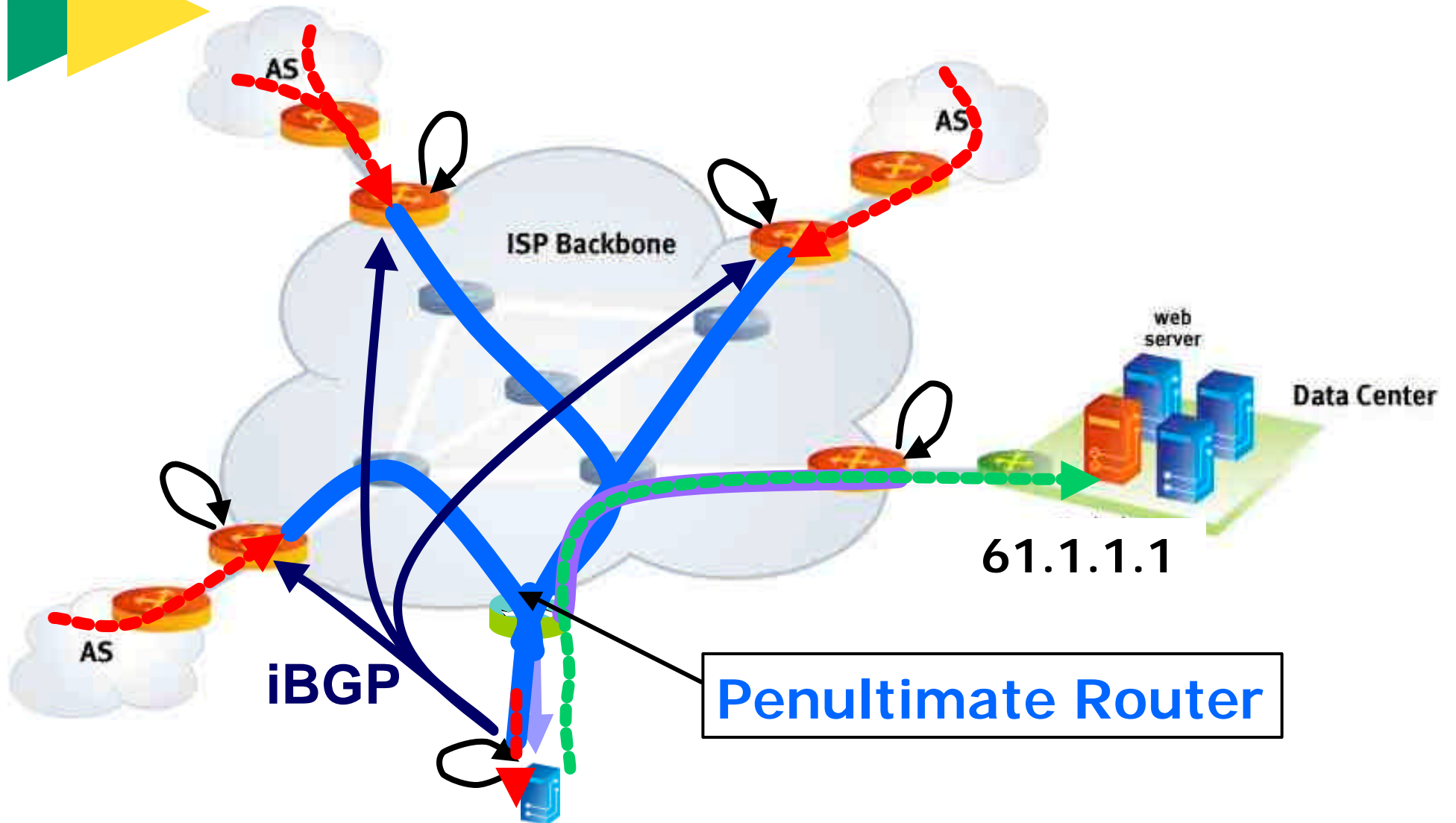
MPLS LSPs based on loopbacks



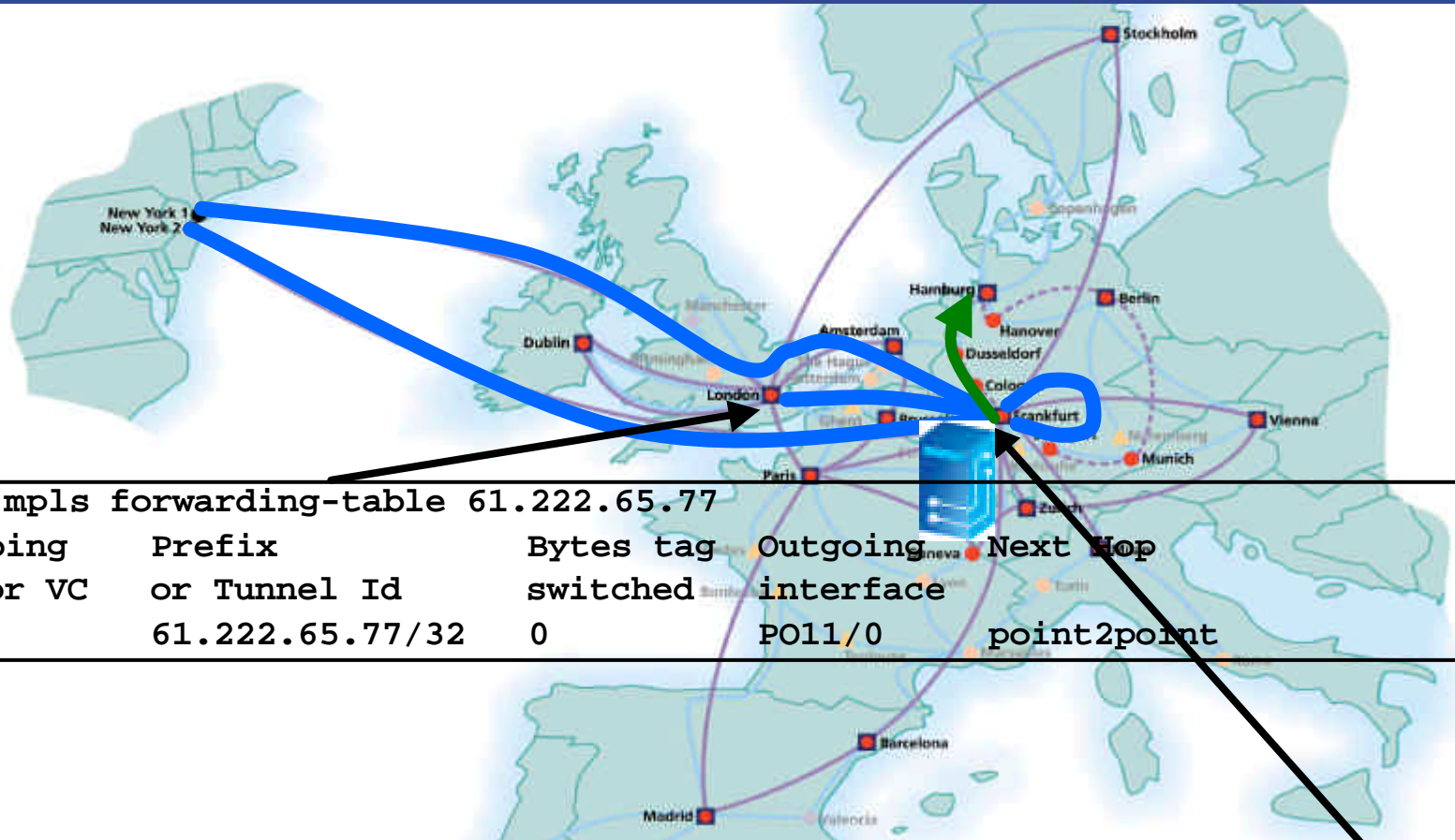
MPLS LSP Proxy Egress



MPLS LSP Proxy Egress



Deployment example

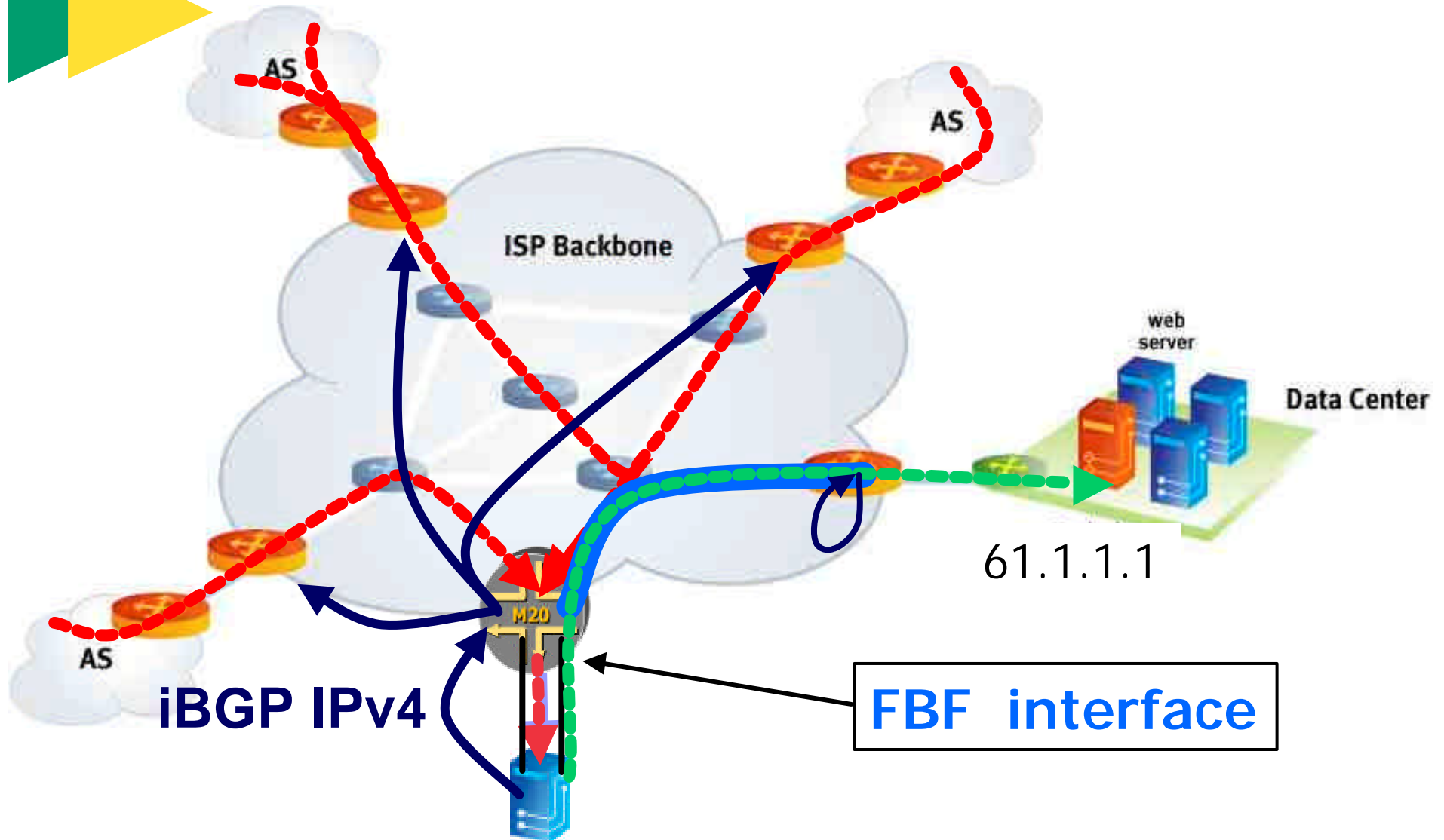


```
LONDON#show mpls forwarding-table 61.222.65.77
Local   Outgoing   Prefix          Bytes tag  Outgoing     Next Hop
tag     tag or VC   or Tunnel Id   switched  interface
503     560        61.222.65.77/32  0         PO11/0       point2point
```

```
FRANKFURT#show mpls forwarding-table labels 16
Local   Outgoing   Prefix          Bytes tag  Outgoing     Next Hop
tag     tag or VC   or Tunnel Id   switched  interface
16      Untagged   61.222.65.77/32  24831266  Gi6/0        61.44.88.111
```



The Juniper way (courtesy of Riverhead)



Conclusion

- Actually deployed, not only in the lab
- Proved easy to deploy, maintain and use
- Improved DDoS detection, mitigation and analysis/post-mortem in conjunction with Netflow-based detection solution and customer profiling (filtering templates)
- Any question ?
- Technical Notes & configurations examples: boaz@riverhead.com

Thank you

