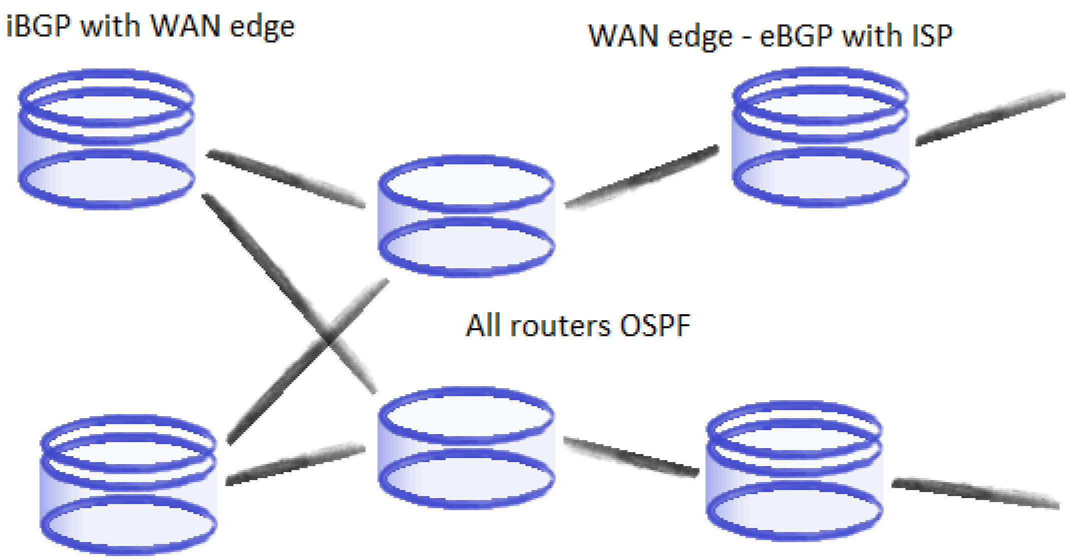
ADVANTAGES OF MPLS : AN EXAMPLE

While MPLS is already explained on this blog, I often still get questions regarding the advantages over normal routing. A clear example I've also already discussed, but besides VRF awareness and routing of overlapping IP ranges, there's also the advantage of reduced resources required (and thus scalability).



Given the above design: two routers connecting towards ISPs using eBGP sessions. These in turn connect to two enterprise routers, and those two enterprise routers connect towards two backend routers closer to (or in) the network core. All routers run a dynamic routing protocol (e.g. OSPF) and see each other and their loopbacks. However, the two middle routers in the design don't have the resources to run a full BGP table so the WAN edge routers have iBGP sessions with the backend routers near the network core.

If you configure this as described and don't add any additional configuration, this design will **not** work. The iBGP sessions will come up and exchange routes, but those routes will list the WAN edge router as the next hop. Since this next hop is not on a directly connected subnet to the backend routers, the received routes will not be installed in the routing table. The enterprise routers would not have any idea what to do with the packets anyway.

Update January 17th, 2014: the real reason a route will not be installed in the routing table is the iBGP synchronisation feature, which requires the IGP to have learned the BGP routes through redistribution before using the route. Still, synchronisation can be turned off and the two enterprise routers would drop the packets they receive.

There are a few workarounds to make this work:

- Just propagating a default route of course, but since the WAN edge routers are not directly connected to each other and do not have an iBGP session, this makes the eBGP sessions useless. Some flows will go through one router, some through the other. This is not related to the best AS path, but to the internal (OSPF) routing.
- Tunneling over the middle enterprise routers, e.g. GRE tunnels from the WAN edge routers towards the backend routers. Will work but requires multiple tunnels with little scalability and more complex troubleshooting.
- Replacing the middle enterprise routers by switches so it becomes layer 2 and the WAN edge and backend routers have a directly connected subnet. Again this will work but requires design changes and introduces an extra level of troubleshooting (spanning tree).

So what if MPLS is added to the mix? By adding MPLS to these 6 routers ('mpls ip' on the interfaces and you're set), LDP sessions will form… After which the backend routers will install all BGP routes in their routing tables!

The reason? LDP will advertise a label for each prefix in the internal network (check with 'show mpls ldp bindings') and a label will be learned for the interfaces (loopback) of the WAN edge routers… After which the backend routers know they just have to send the packets towards the enterprise routers with the corresponding MPLS label.

And the enterprise routers? They have MPLS enabled on all interfaces and no longer use the routing table or FIB (Forwarding Information Base) for forwarding, but the LFIB (Label Forwarding Information Base). Since all packets received from the backend routers have a label which corresponds to the loopback of one of the WAN edge routers, they will forward the packet based on the label.

Result: the middle enterprise routers do not need to learn any external routes. This design is scalable and flexible, just adding a new router and configuring OSPF and MPLS will make it work. Since a full BGP table these days is well over 450,000 routes, the enterprise routers do not need to check a huge routing table for each packet which decreases resource usage (memory, CPU) and decreases latency and jitter.

Source : http://reggle.wordpress.com/2014/01/08/advantages-of-mpls-an-example/