IP ArchiTechs
MANAGED SERVICES
www.iparchitechs.com

WISP Design – Using eBGP and OSPF transit fabric for traffic engineering

PRESENTED BY:

KEVIN MYERS,
NETWORK ARCHITECT

Profile: **About Kevin Myers**

Background:

- 19+ years in Networking
- Designed/Built Networks on 6 continents
- MikroTik Certified Trainer
- MikroTik, Cisco and Microsoft Certified

Community Involvement:

Packet Pushers (Podcast Guest / Blogger)

Group Contributor (RouterOS / WISP Talk and others)

Delegate/Roundtable contributor (NFD14)

MT Forum (Forum Veteran – Member since 2012)

Network Collective (Podcast Guest)

Profile: About IP ArchiTechs

IP ArchiTechs
MANAGED SERVICES

Expert Networking

Whitebox | ISP | Data Center | Enterprise

- ✓ Global Consulting
- ✓ Managed Networks
- ✓ Monitoring
- ✓ Load Testing
- ✓ Development

Locations in: **US** | **Canada** | **South America**

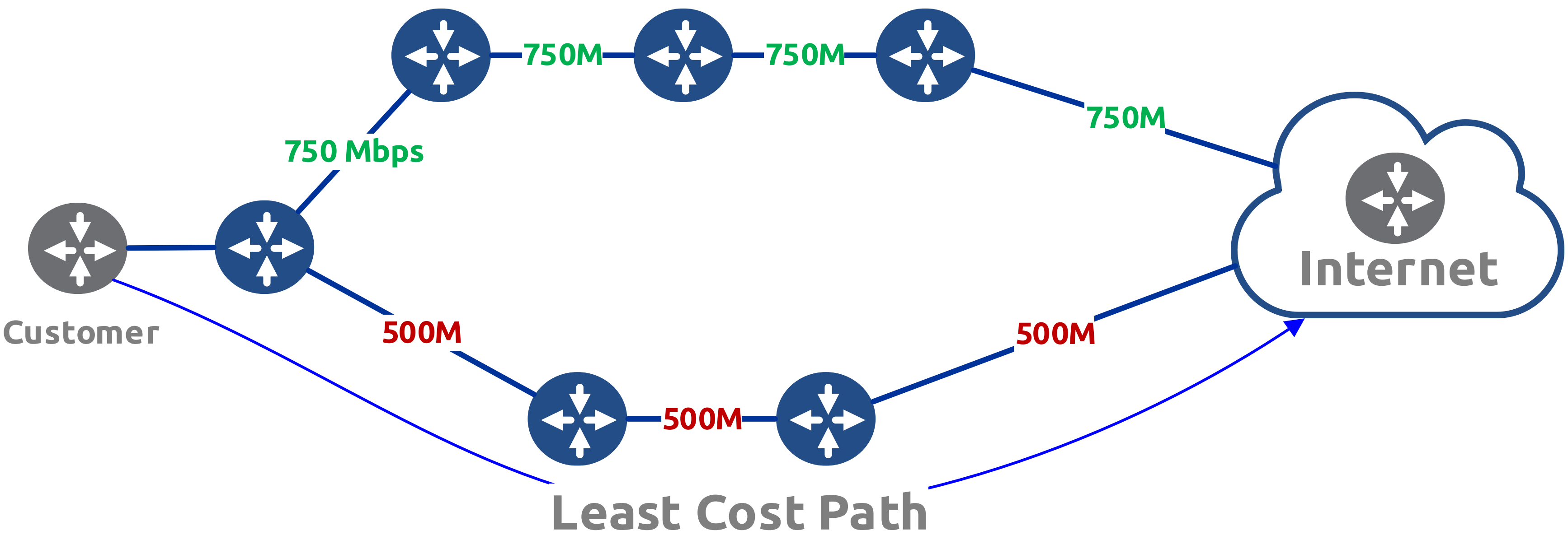
Call us at: +1 855-645-7684
E-mail: consulting@iparchitechs.com
Web: www.iparchitechs.com

Design: **Why eBGP/OSPF Transit Fabric?**

Goal of this presentation: When the presentation is finished, hopefully you will have walked away with a few key concepts:

- limitations in using OSPF or other non-BGP routing protocol when attempting to influence traffic paths in a WISP network
- How to leverage all bandwidth between two towers using an OPSF Transit Fabric
- The benefits of using eBGP and communities to build a scalable framework for identifying and managing traffic in tower networks.

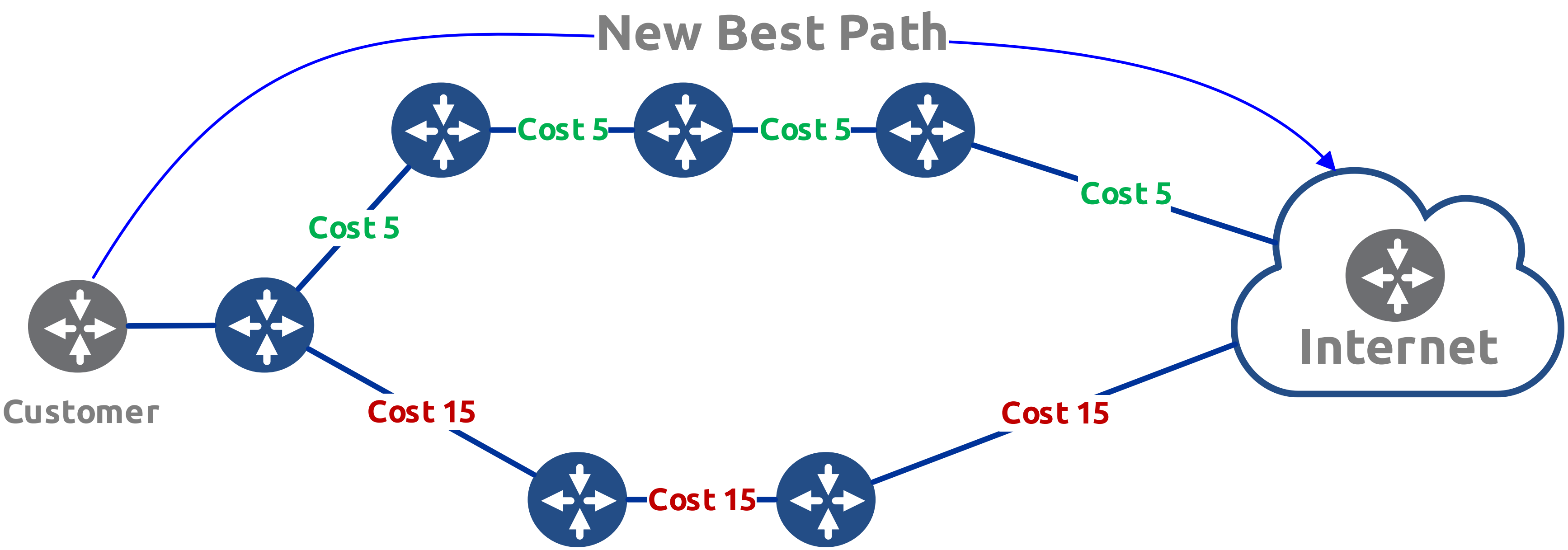
Problem #1: Many WISPs utilize OSPF (Static routing or Bridging also) for the forwarding of traffic. As the network grows, this creates suboptimal traffic flows.



Example: OSPF when using default Gigabit cost (10) will select the "shortest" path – but not the fastest.

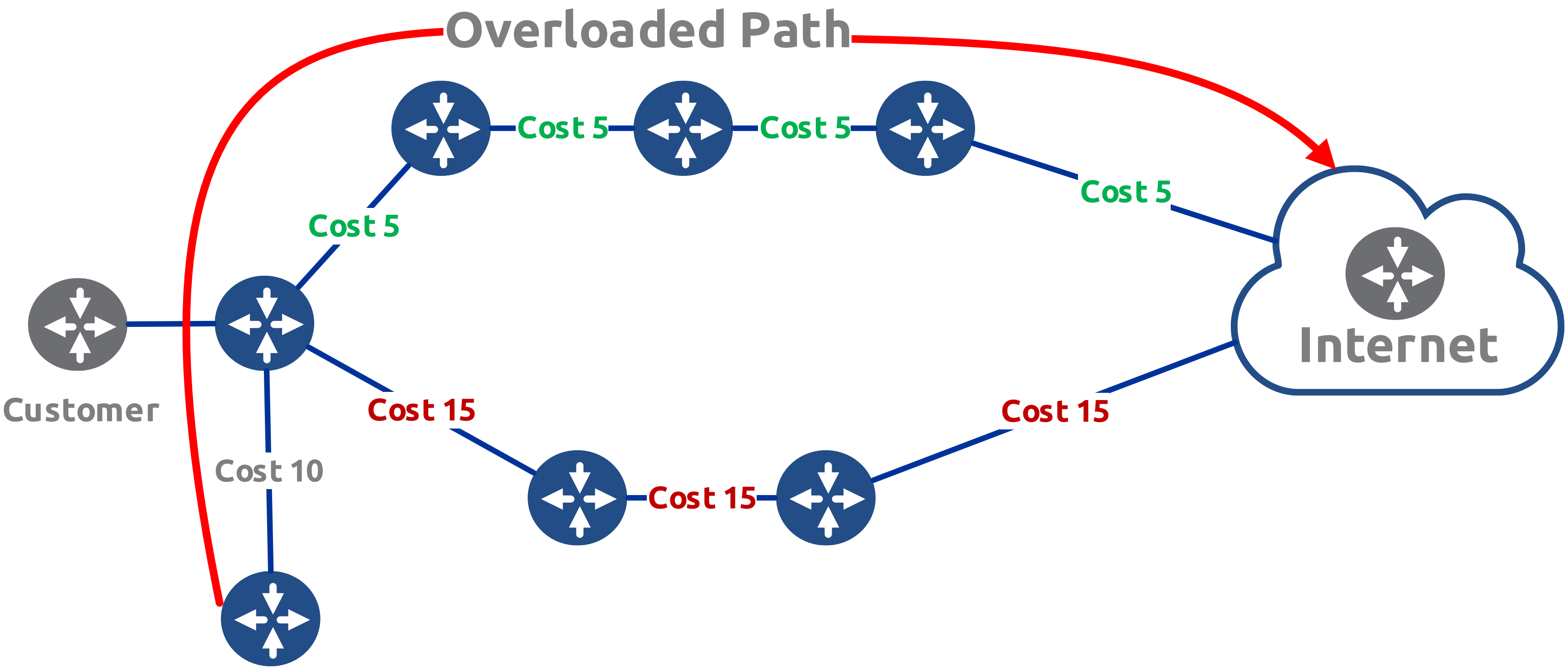
Design: **Why eBGP/OSPF Transit Fabric?**

Better or Worse?: We've modified the OSPF cost so that traffic takes the lowest cost path. How well does this scale?



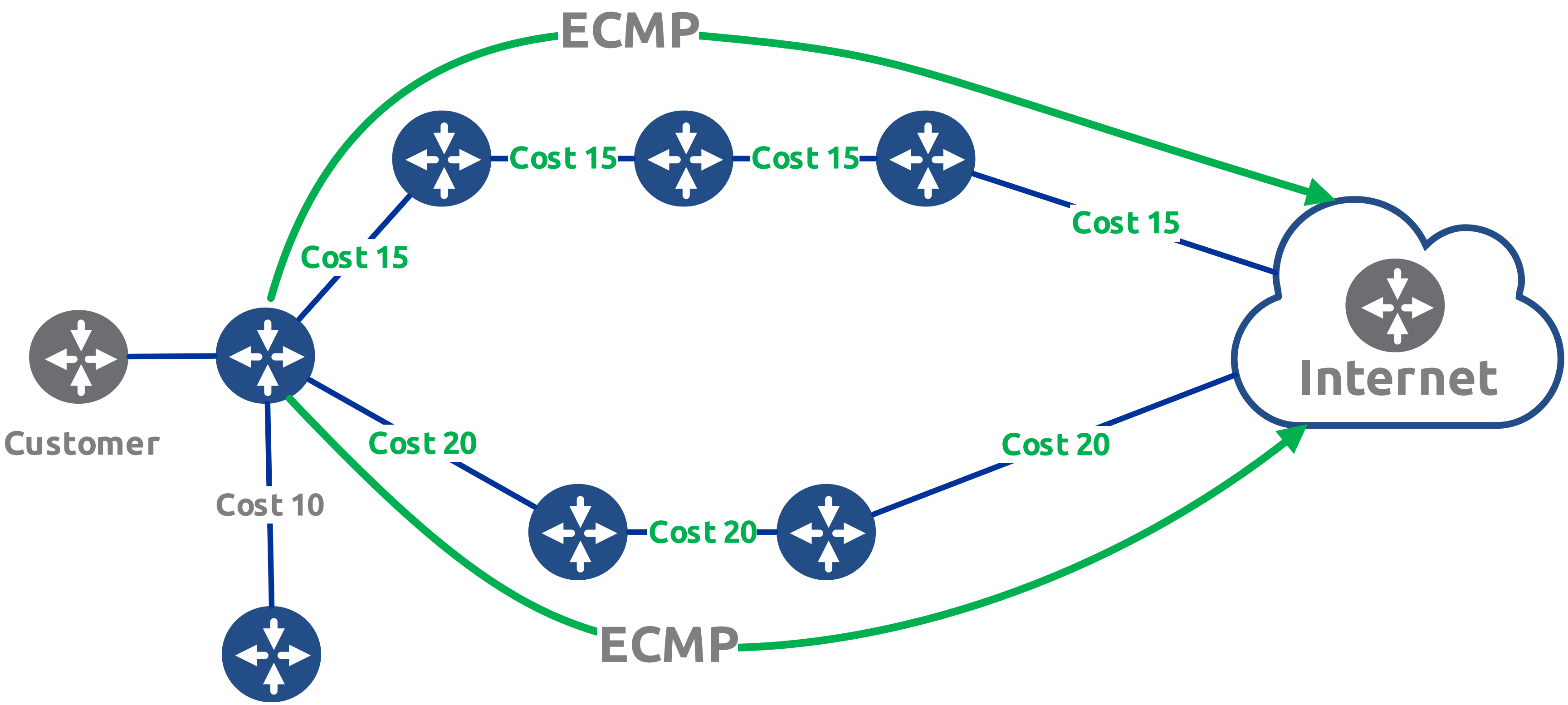
Example: Cost has now been modified to 5 on each link for the faster path and 15 on each link for the slower path.

It's Worse!: Now we have a new tower to add. The "fast" path is getting overloaded and we need to use some of the bandwidth on the other path.



Example: With the addition of the new tower, we need to utilize bandwidth down both paths.

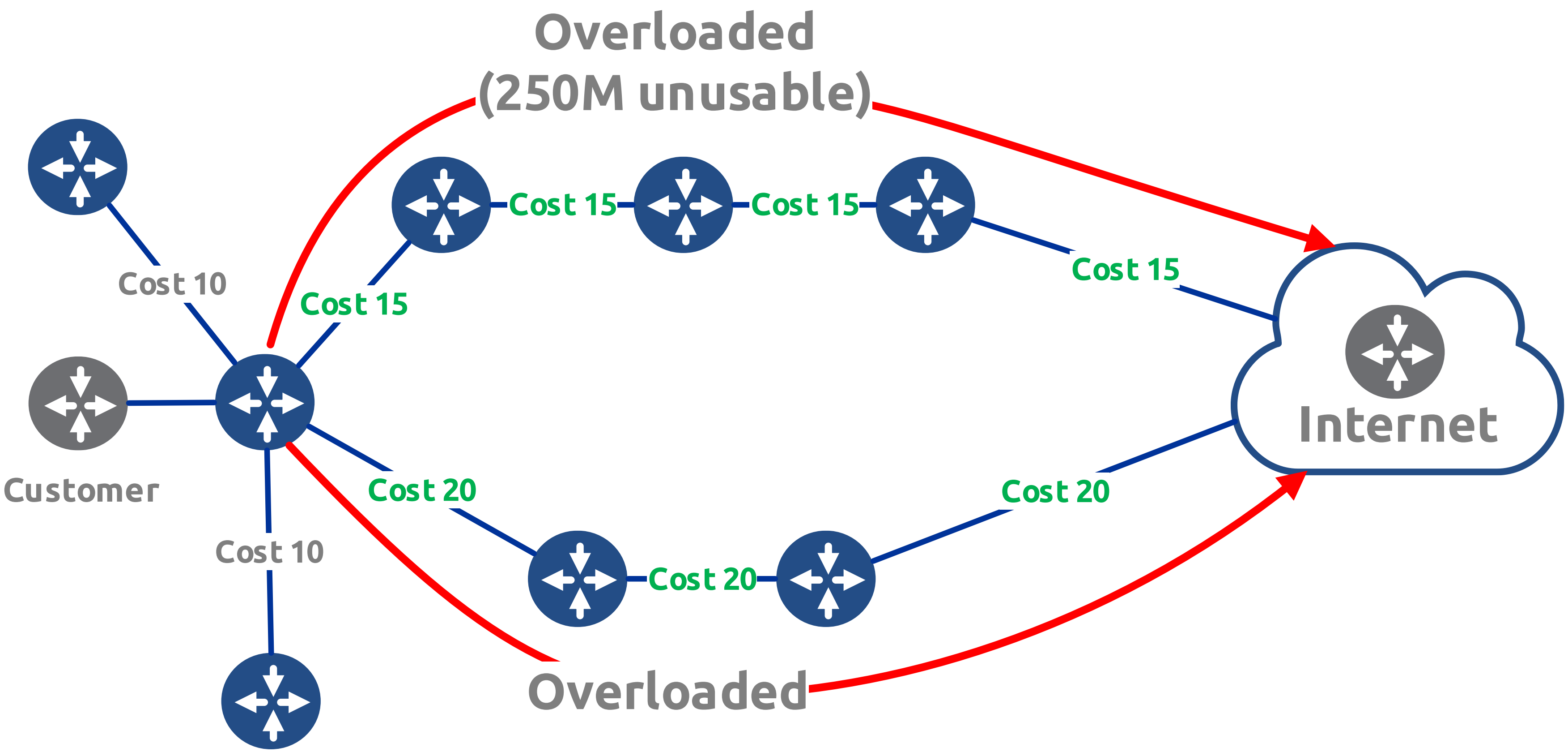
Sacrificing bandwidth: Now we have utilized both paths and alleviated the traffic problem yet again, but we had to sacrifice 250Mbps on the faster path to accomplish it.



Example: The OSPF costs have been set to enable ECMP so that both paths will carry traffic.

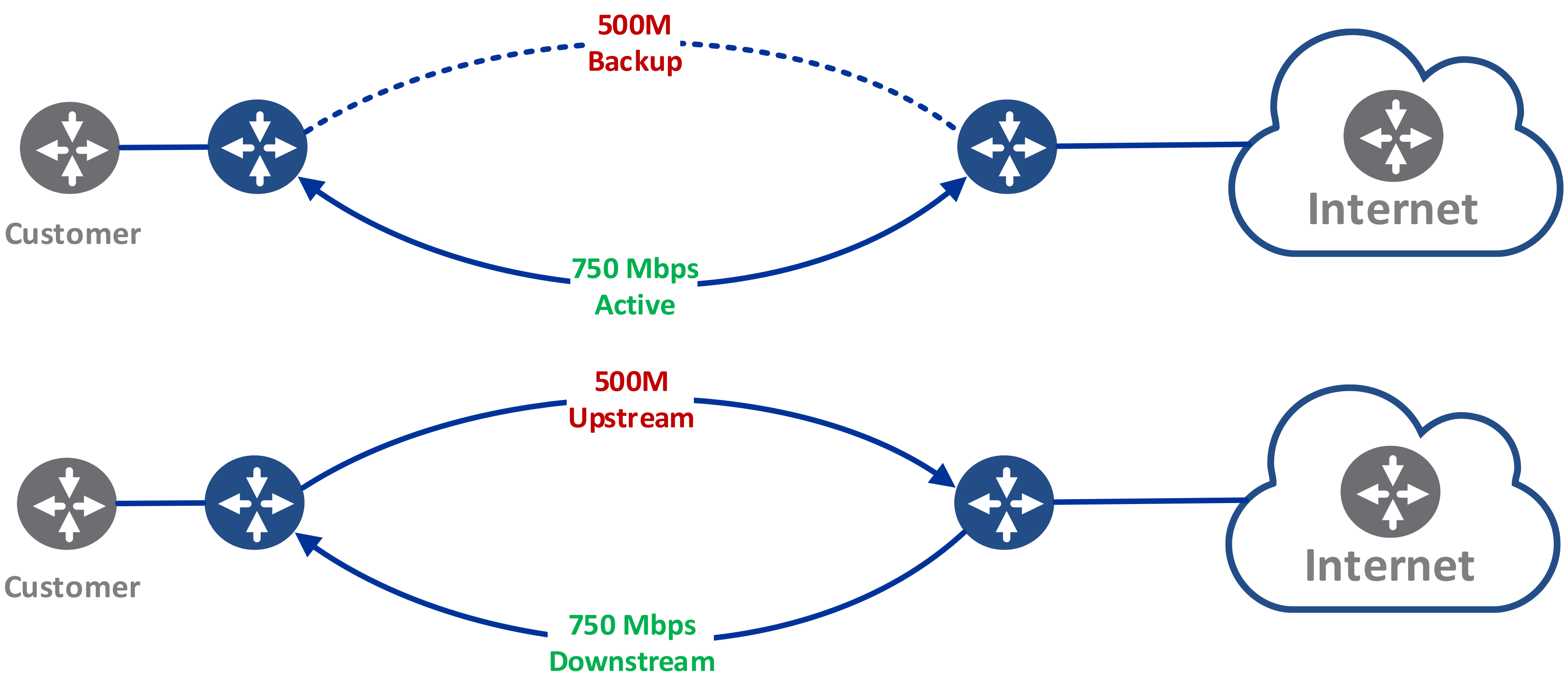
Design: Why eBGP/OSPF Transit Fabric?

Overloaded and low on options: Now we have added another tower. At this point, we have utilized both paths and have no more room on either due to ECMP. Even though there is 250Mbps of potential bandwidth, we can't use it without static or policy routing.



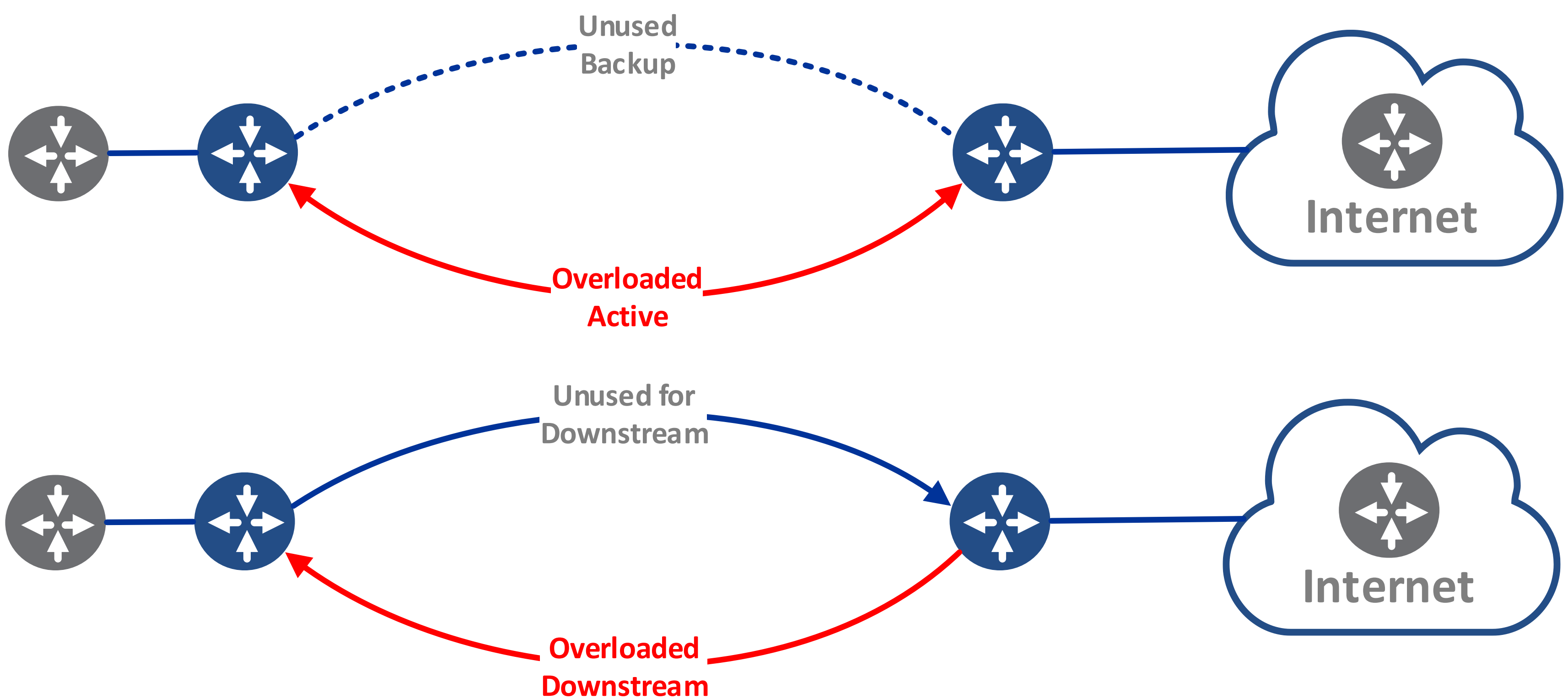
Design: Why eBGP/OSPF Transit Fabric?

Problem #2: In WISP deployments, there is often more than one set of PtP radios connecting the same two towers. Traditionally, these are configured into an active/backup setup or one link is configured to handle upstream, and the other is configured for downstream traffic.



Design: Why eBGP/OSPF Transit Fabric?

Unused capacity: The biggest challenge with using PtP links in active/backup or upstream/downstream configurations is the unused capacity in the lower speed links.



Design: **OSPF Transit Fabric and eBGP overview**

OSPF alone doesn’t work: Looking at the previous slides, we saw a number of issues and workarounds to manage traffic using only OSPF as the routing protocol. With every new workaround, a new issue or limitation with scaling the design came up. Why is that?

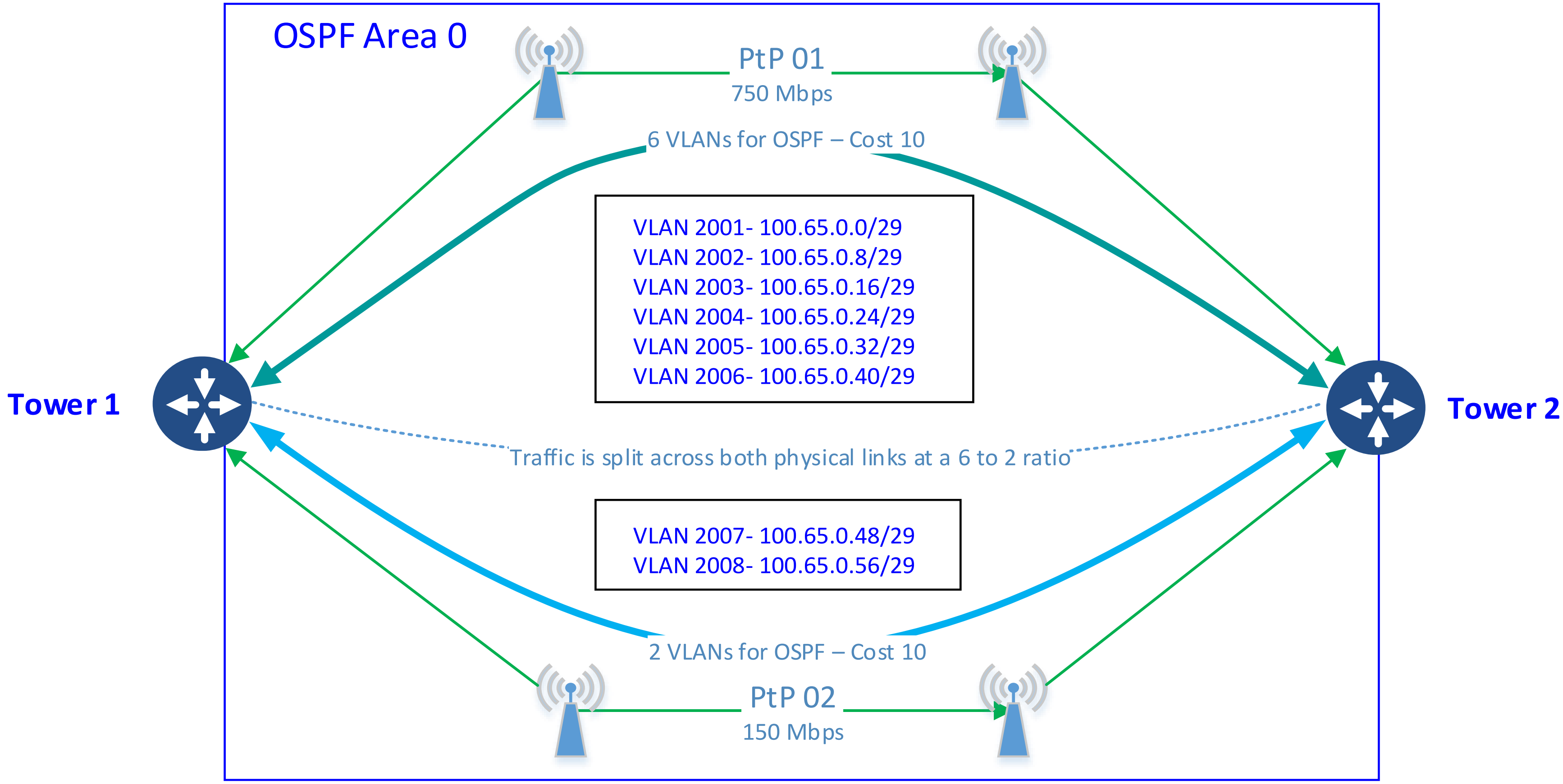
- Because OSPF is designed to be a reachability protocol – NOT a policy protocol. The function of a link state routing protocol is to map the available paths and speeds to prefixes.
- OSPF is not well suited to implementing policy as we saw in the previous slides

Design: **OSPF Transit Fabric and eBGP overview**

Potential Solution: Now that we've identified the problems and limitations associated with turning an IGP like OSPF into a policy protocol, let's look at a potential solution.

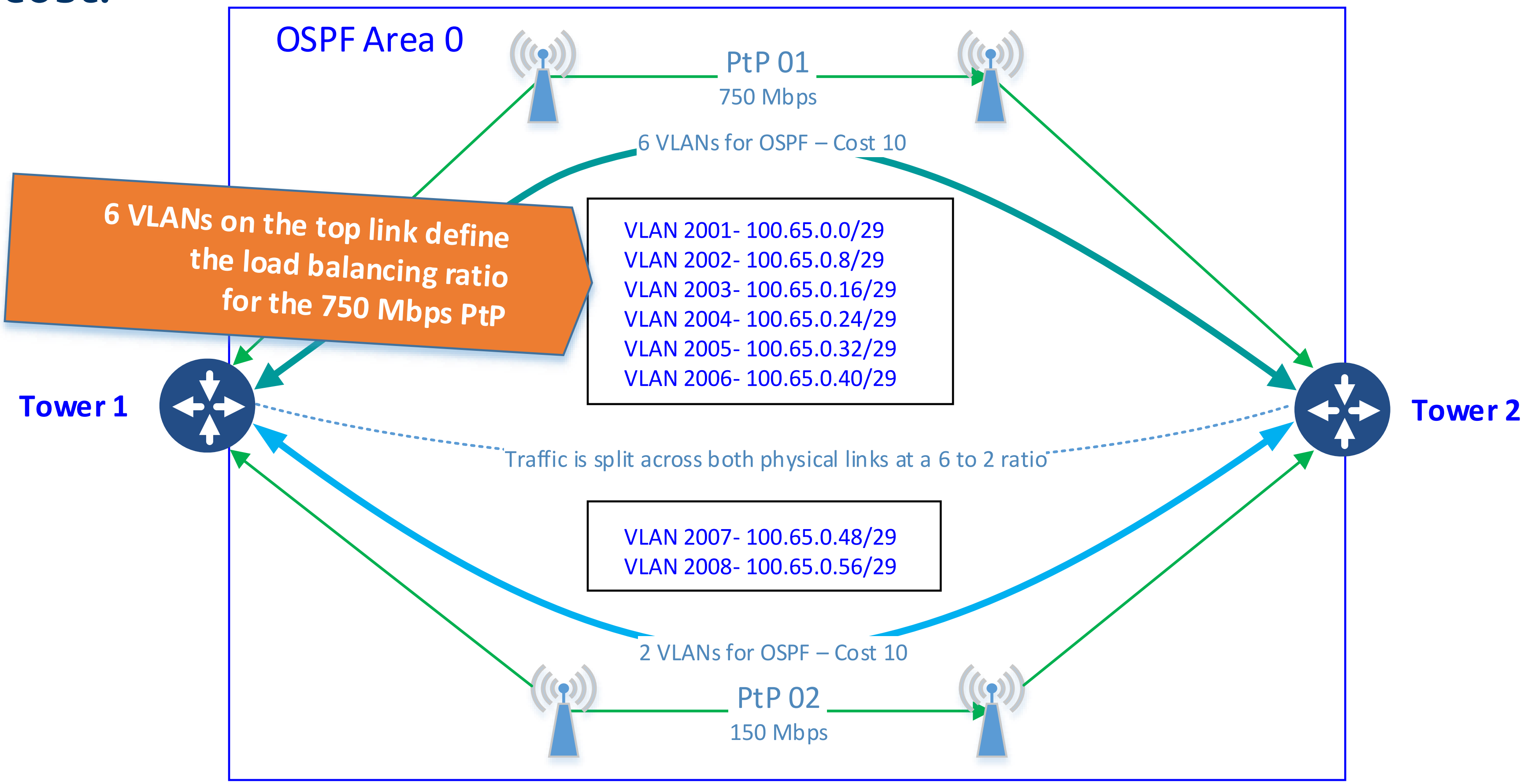
- Current design approach is to use OSPF to advertise transit subnets and loopbacks to form an iBGP AS. This works very well but can run into limitations in policy options due to the dependence on an IGP to advertise the next hop.
- Combining eBGP and the OSPF Transit Fabric allows for total control of the tower path and allows for the use of all bandwidth between any two towers.
- It's not all or nothing, iBGP can still be utilized where it makes sense and eBGP/Transit Fabric deployments can be utilized at key aggregation points to make complex traffic decisions.

Defining the OSPF Transit Fabric: The OSPF Transit Fabric (TF) is a design that came out a real world deployment for an IP ArchiTechs client. We needed a way to use all of the bandwidth between two towers on links of unequal speed.



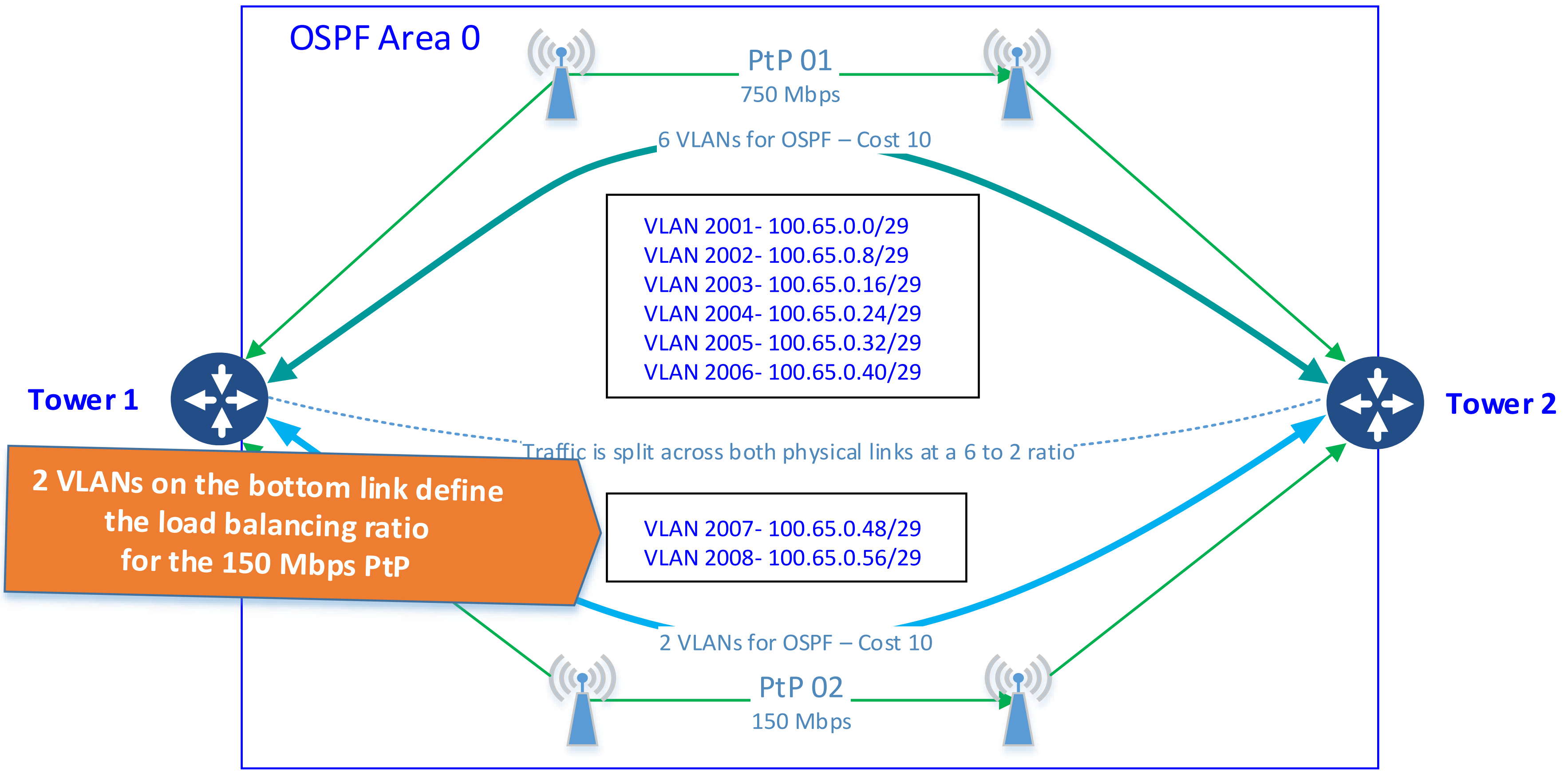
Design: **What is an OSPF Transit Fabric?**

Using VLANs to create unequal load balancing: The idea behind the TF is to use VLANs at a default cost to build a ratio for balancing traffic across unequal links. ECMP balances traffic by flow when there is a next hop reachable via interfaces of equal cost.



Design: **What is an OSPF Transit Fabric?**

Using VLANs to create unequal load balancing: Because the VLAN interface on each PtP has the same cost, RouterOS sees 8 equal-cost next hops to load balance traffic onto. This forms a 6 to 2 ratio of traffic in favor of the higher speed link.



The end result: What the RouterOS routing table looks like with a TF.

```
[admin@IPA-LAB-TWR-1] > ip route print detail
Flags: X - disabled, A - active, D - dynamic, C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 ADb  dst-address=0.0.0.0/0 gateway=100.127.0.1
        gateway-status=100.127.0.1 recursive via 100.65.0.89,100.65.0.81,100.65.0.49,100.65.0.73,100.65.0.97,
               100.65.0.57,100.65.0.41 VLAN3016,VLAN3015,VLAN3011,VLAN3014,VLAN3017,VLAN3012,VLAN3010
        distance=200 scope=40 target-scope=30 bgp-local-pref=100 bgp-origin=incomplete
        received-from=TO-CORE1
```

Live traffic: Here is a look at the physical and VLAN interfaces under load in a TF - 6.2 Mbps on BH-06/ 22.8 Mbps on BH-07

;;; Transit Through BH-06							
R	VLAN3010	VLAN		3.3 Mbps	140.4 kbps	282	290
R	VLAN3011	VLAN		2.9 Mbps	188.2 kbps	245	388
;;; Transit Through BH-07							
R	VLAN3012	VLAN		4.0 Mbps	512 bps	337	1
R	VLAN3013	VLAN		0 bps	512 bps	0	1
R	VLAN3014	VLAN		6.7 Mbps	167.7 kbps	561	343
R	VLAN3015	VLAN		8.1 Mbps	163.9 kbps	690	340
R	VLAN3016	VLAN		3.5 Mbps	330.7 kbps	296	683
R	VLAN3017	VLAN		535.0 kbps	182.9 kbps	49	377

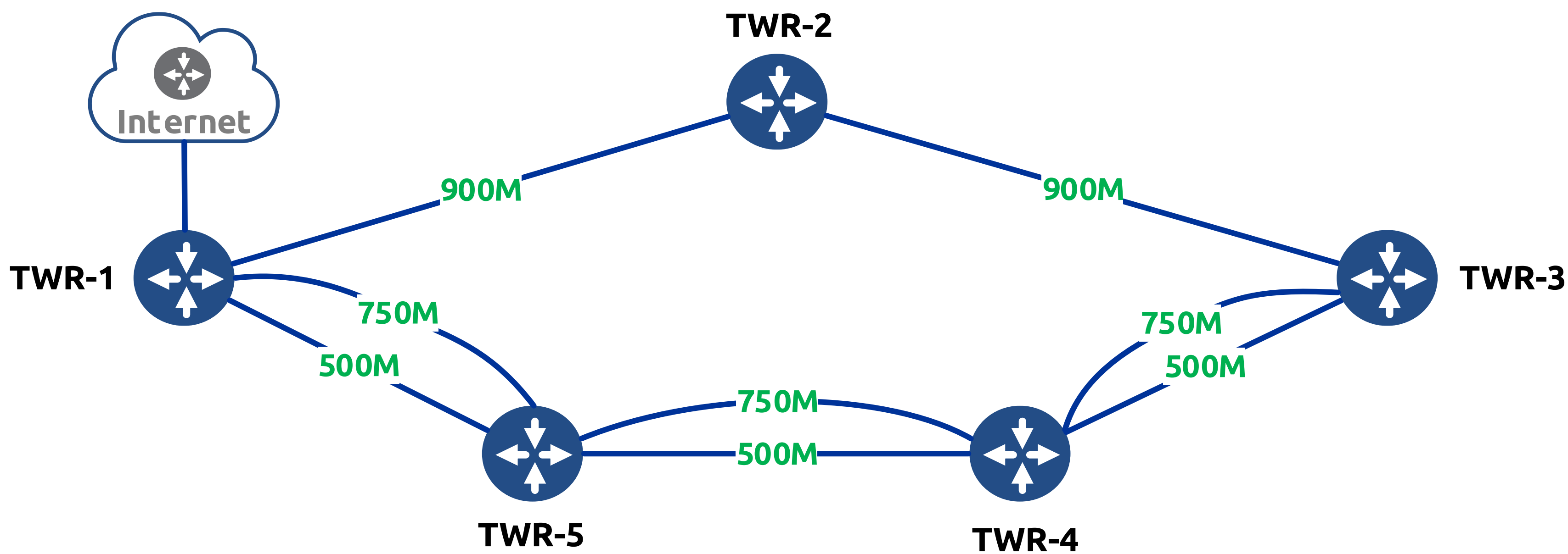
Design: What is an OSPF Transit Fabric?

Now I want an OSPF TF, but how does it scale?

- One of the challenges of implementing the OSPF TF is scaling it to multiple towers
- WISPs that are under 50 towers can utilize the OSPF TF without using another routing protocol on top of OSPF such as BGP
- WISPs that are over 50 towers or that have complex traffic management requirements may want to consider coupling the OSPF TF with eBGP
- eBGP provides a way to limit OSPF strictly between the towers by using eBGP multihop peering and loopbacks.

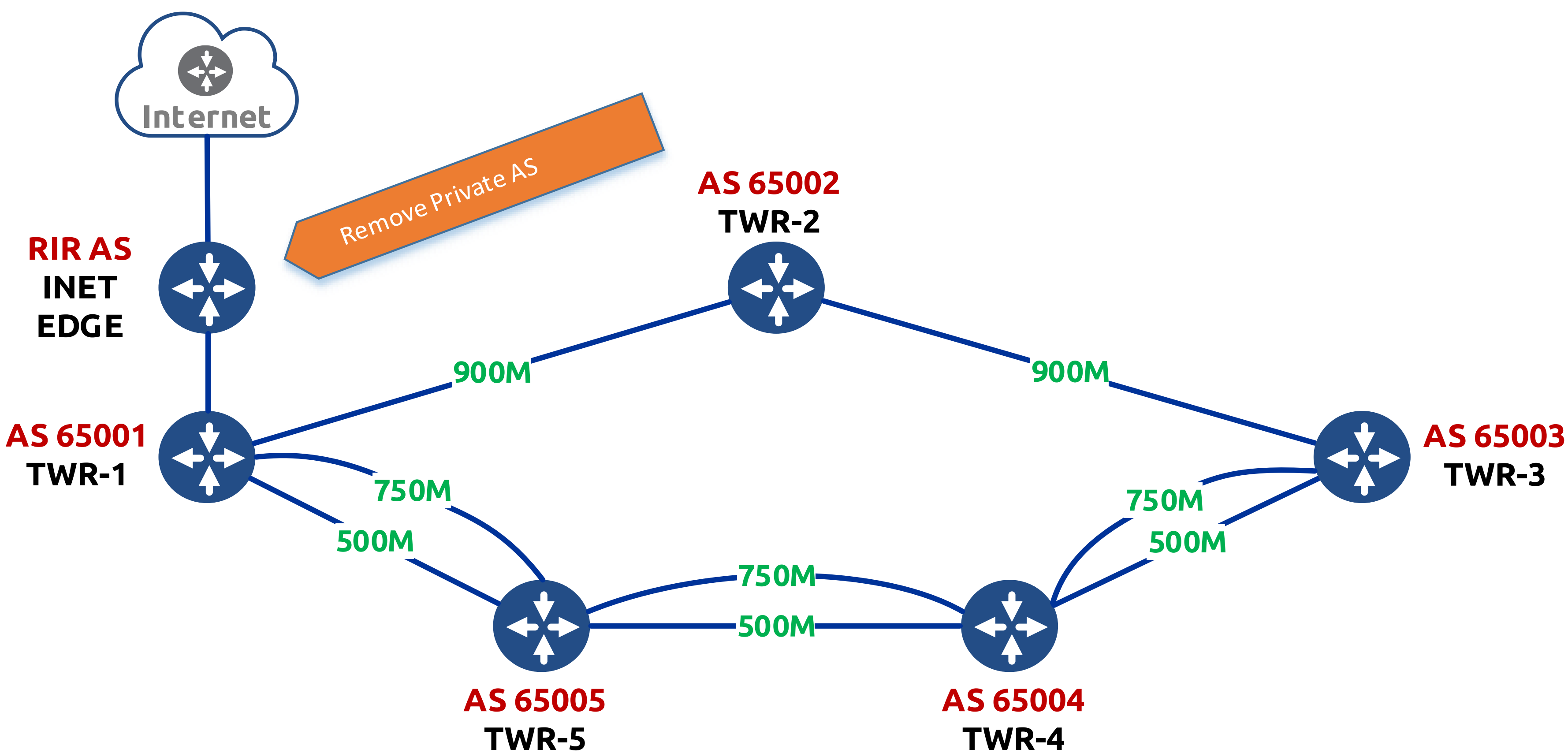
Design: **Utilizing eBGP at the tower to set traffic paths**

eBGP as a traffic engineering mechanism: One of the major benefits of using eBGP is the ability to definitively set the next hop by matching on a number of attributes. When coupled with BGP communities, a vast array of options for traffic engineering are available.



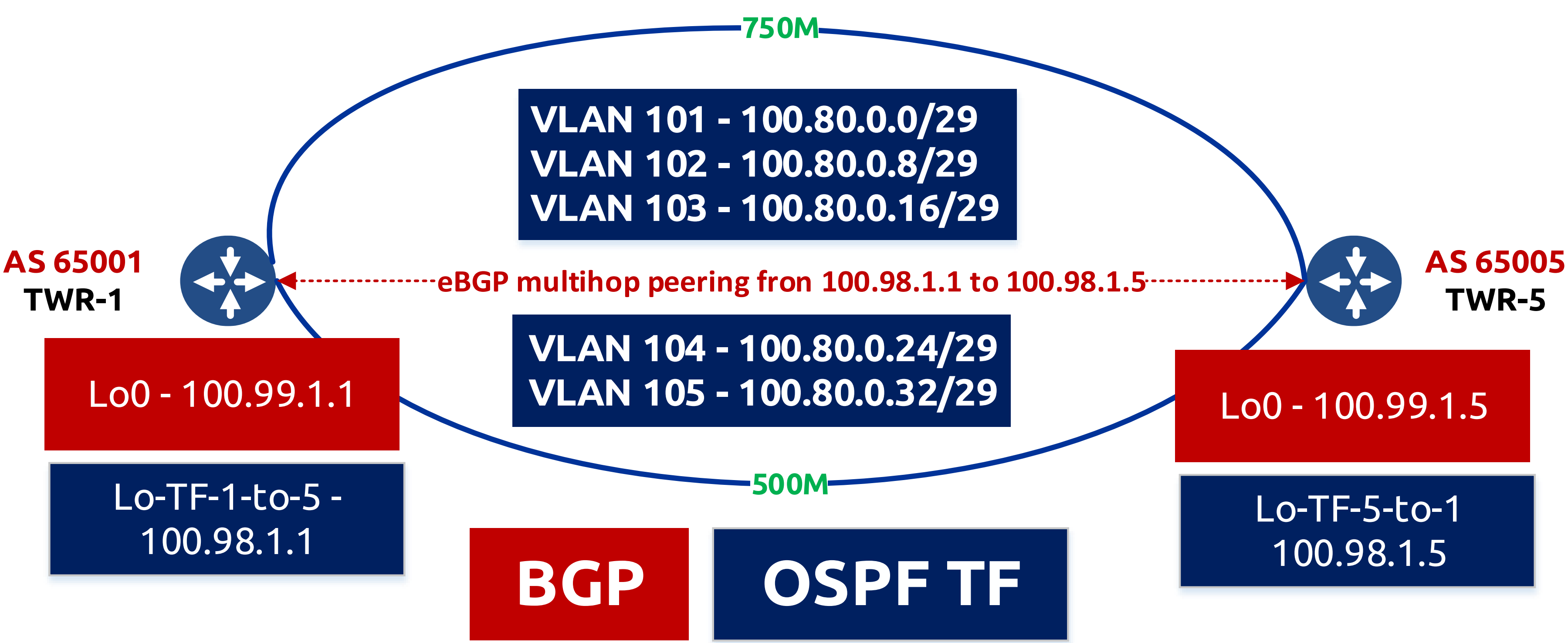
Design: **Utilizing eBGP at the tower to set traffic paths**

How does eBGP differ from iBGP?: eBGP or External BGP means the AS you are peering to is different than your AS. We are going to look at a network that combines eBGP and the OSPF Transit fabric to solve traffic engineering and unequal load balancing issues.



Design: Utilizing eBGP at the tower to set traffic paths

Combining OSPF TF and eBGP: In order to build an eBGP design but still leverage the OSPF TF for capacity, we have to build a multihop eBGP peering using the OSPF TF loopbacks in blue. Then the BGP loopbacks in red can be advertised as the reachable loopback outside of that router.



Design: Utilizing eBGP at the tower to set traffic paths

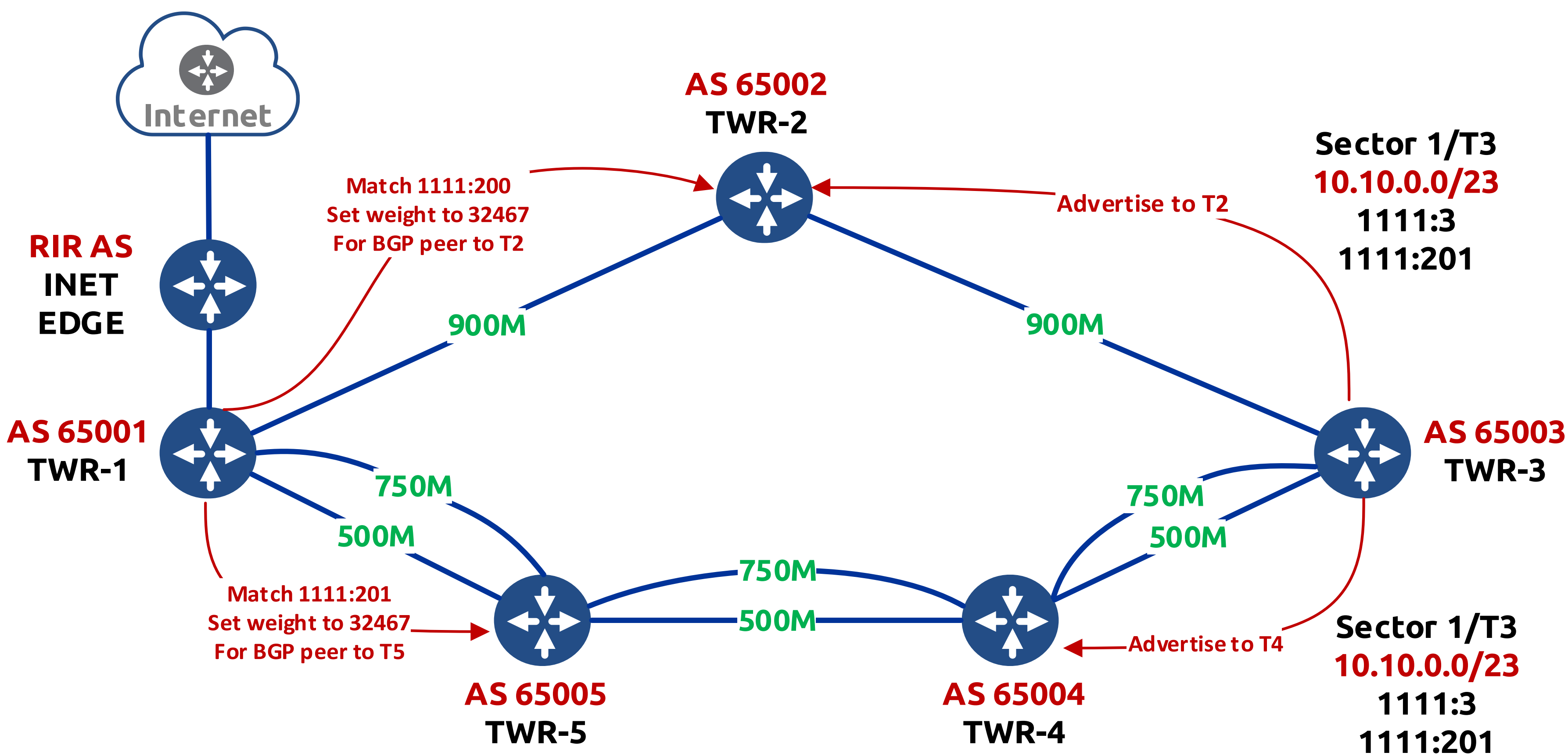
BGP Communities: BGP Communities are a field in a BGP route advertisement that can be used to take action on a prefix tagged with that community number. They can also be used for route identification. The format is 32 bits with the first 16 bits typically set to the ASN of the network and the last 16 bits are defined by the operator.

BGP Communities

1111:1	Tower 1
1111:2	Tower 2
1111:3	Tower 3
1111:4	Tower 4
1111:5	Tower 5
1111:200	Northern Tower Path (T3,T2,T1)
1111:201	Southern Tower Path (T3,T4,T5,T1)

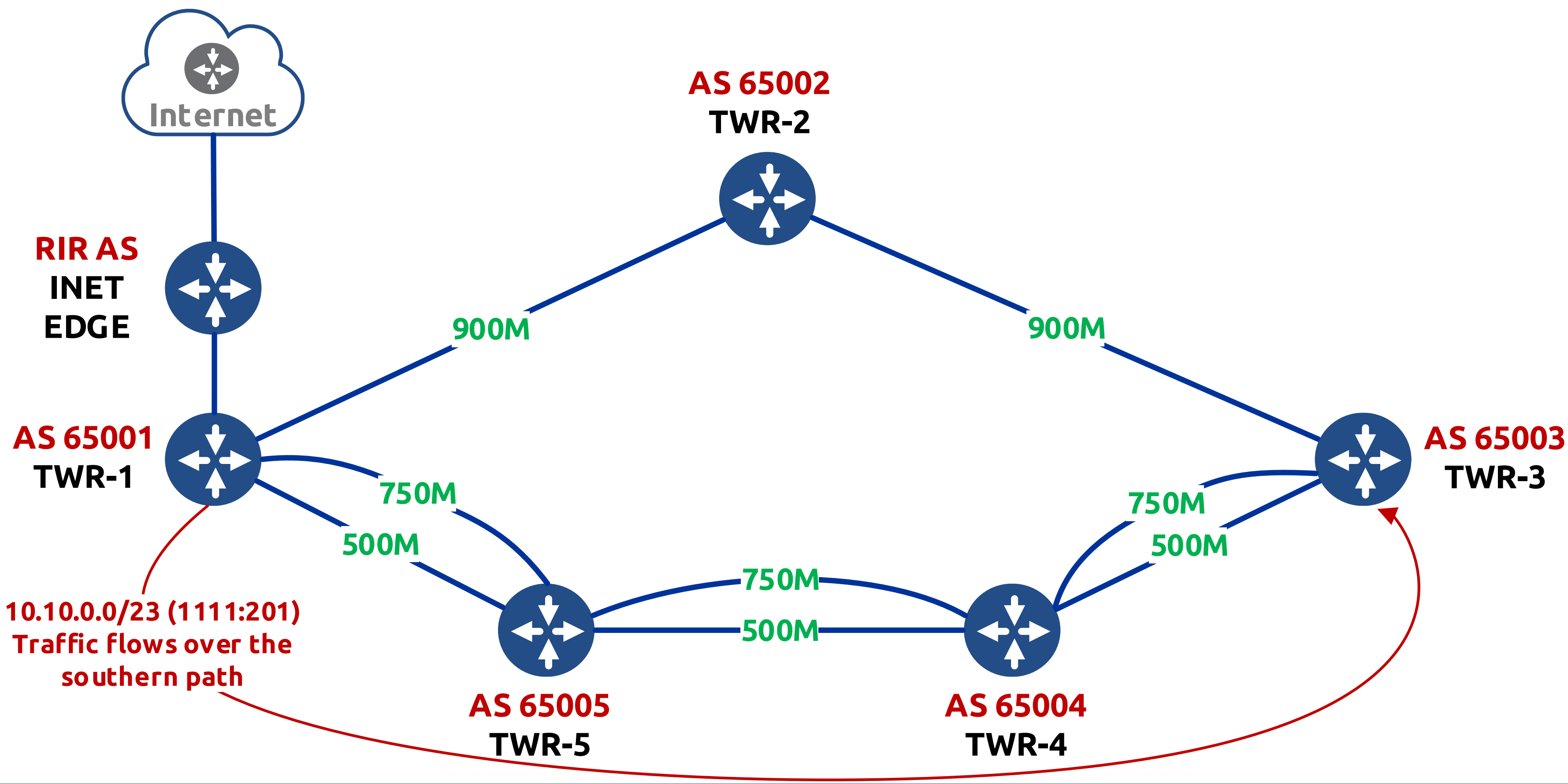
Design: **Utilizing eBGP at the tower to set traffic paths**

Changing the next hop: How a community is matched and weight set.

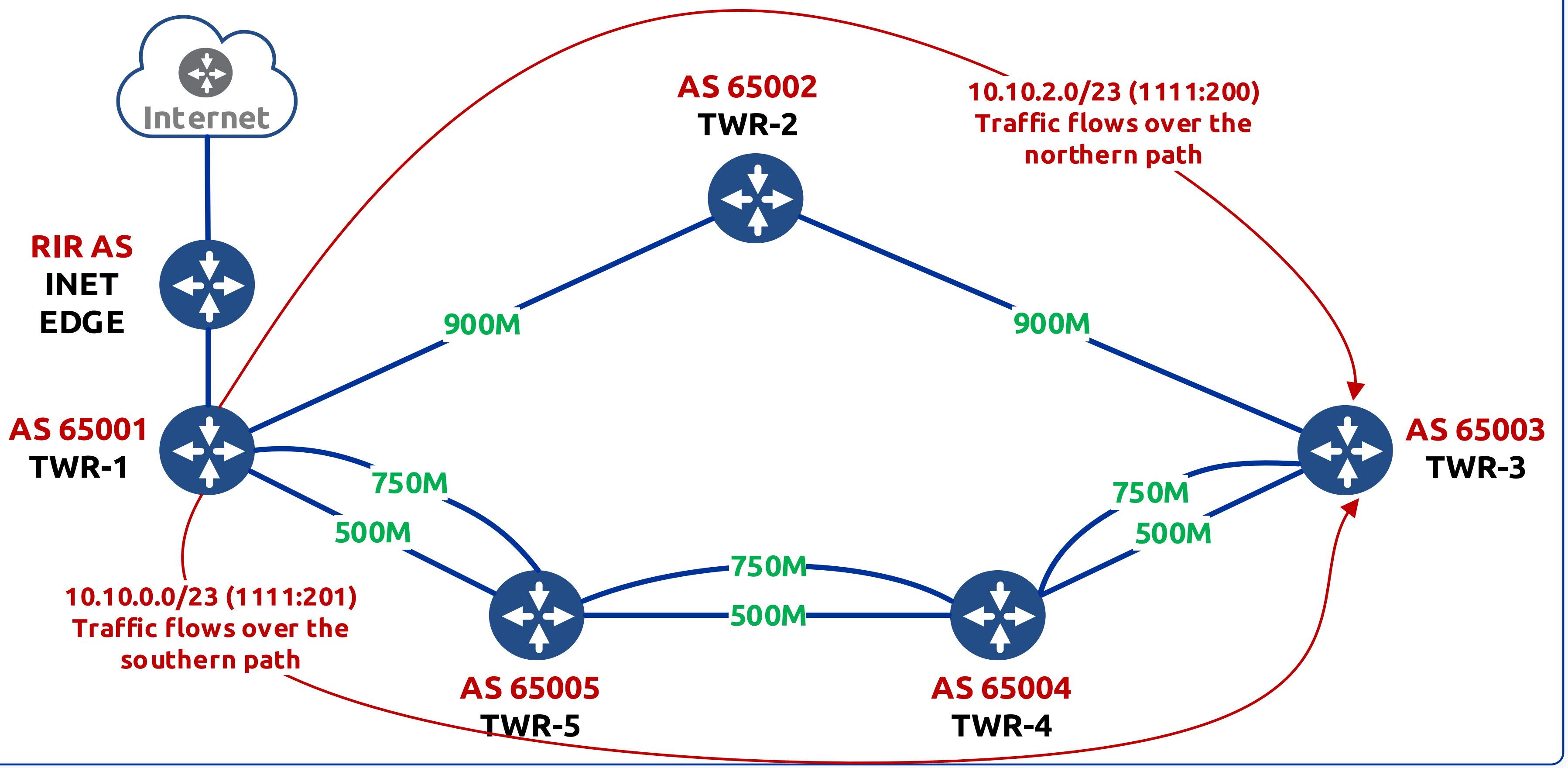


Design: **Utilizing eBGP at the tower to set traffic paths**

Influencing traffic down one path: By tagging the subnet for Sector 1 at Tower 3 (10.10.0.0/23) into community 1111:201, we are able to force the traffic along the higher capacity but less preferred southern tower path.



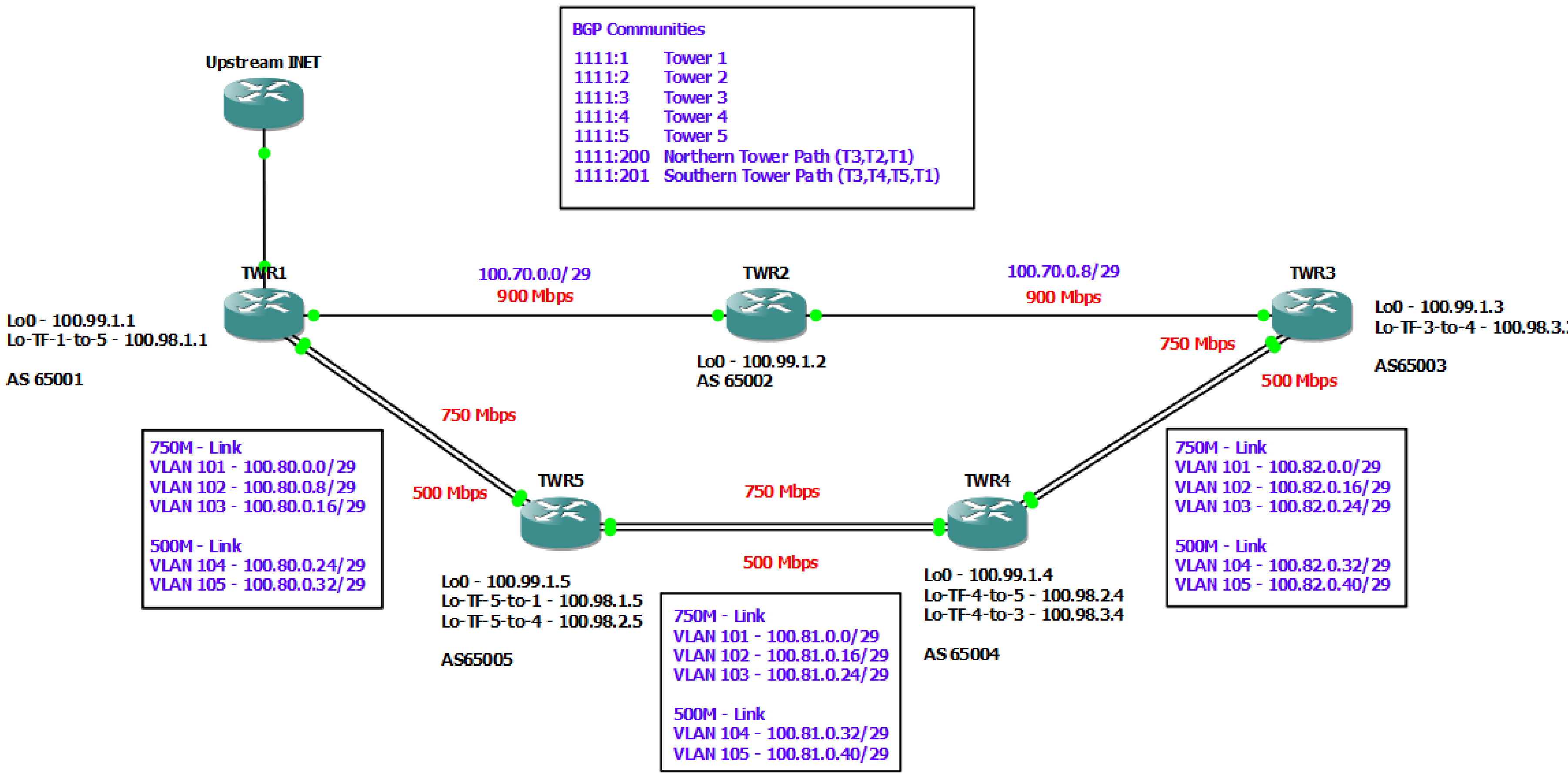
Influencing traffic down two paths: Now we can break the sectors up at Tower 3 and send one subnet down the Northern path and one subnet down the Southern path to utilize multiple paths to get back to one tower.



Now let's talk about your Questions!

Design: **Utilizing eBGP at the tower to set traffic paths**

Live Demo : See the design in action at the IPA booth! Come by and learn more about eBGP and OSPF Transit Fabric



BGP Party at 6:30!

MAY
25

BGP Party - Billards, Grub & Planning your network

Public · Hosted by IP Architechs

✓ Going ▾ | Share ▾ | ···

Today at 6:30 PM - 10 PM
Starts in about 4 hours · 72° Cloudy

Tarantula Billiards Bar & Grill
1520 Stout Street, Denver, Colorado 80202
Show Map

Please RSVP by going to www.facebook.com/iparchitechs