Trans-Oceanic Internet Backbones
Using Simplex Links - Configuration Options

Background

Today's trans-oceanic Internet backbones are more than just one lease line across the ocean. They are complex *systems* with multiple circuits using the best of terrestrial and satellite telecommunications systems. One of the most common examples is the combination of terrestrial and satellite telecommunications paths. This system relies on the ability of a satellite telecommunications system to send traffic in only one direction[1]. TCP/IP traffic request traverses the terrestrial path (Figure 1). Return traffic is forwarded over the simplex satellite circuit and passed to the requester. This allows the simplex satellite circuit to be a much higher bandwidth than the terrestrial link.
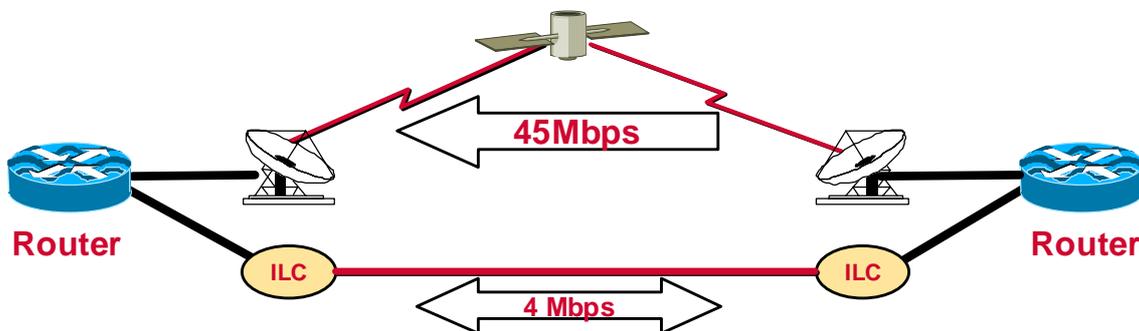

Figure 1 - Simplex Satellite Circuits Combined with Traditional Terrestrial Circuits

WHY USE THESE HYBRID SYSTEMS?

Two reasons are usually give by ISPs who select this combination of simplex satellite and terrestrial systems. First, the scarcity of terrestrial telecommunications infrastructure drives ISP to utilize all available bandwidth. In much of the world, 2 Mbps links are *high speed.* Once the Internet's exponential growth settles in, the local terrestrial telecommunications infrastructure is consumed faster than the most optimistic projections. In many places, the terrestrial infrastructure that is consumed the fastest is the trans-oceanic infrastructure - leaving little room for upgrade. Simplex satellite systems supplement the available trans-oceanic terrestrial systems ensuring the terrestrial path is reserved for the most latency sensitive applications.

---

[1] This Technique has been most commonly used in broadcast communications.

CISCO SYSTEMS

Second, simplex satellite circuits allow ISPs to create a backbone that matches their consumer oriented traffic patterns. Most ISPs outside of the US have consumer based traffic flow (i.e. consumers pull down more traffic than that create and push out to the Net). These consumption ratios range from 90:1 to 60:40 - depending on the country, ISP, and many other factors. This *asymmetric* traffic ratio leaves a lot of idle bandwidth on the trans-oceanic backbone. Combining simplex satellite circuits with traditional terrestrial circuits allow the ISPs to match their consumption oriented traffic ratio - ensuring effective utilization of their backbone.

First Used

Telstra and Teleglobe are the first to commission a production trans-oceanic Internet backbone that used a combination of a 45 Mbps simplex satellite circuit and a 16 Mbps terrestrial circuit. Teleglobe was the US side of the trans-oceanic system. Geoff Huston [gih@telstra.net] developed a configuration technique based on eBGP and static routes.[2] Since two different ISPs controlled the routers on either side of the ocean, an eBGP configuration was used. Example 1 provides the configuration used to validate the feasibility of the simplex/terrestrial systems.
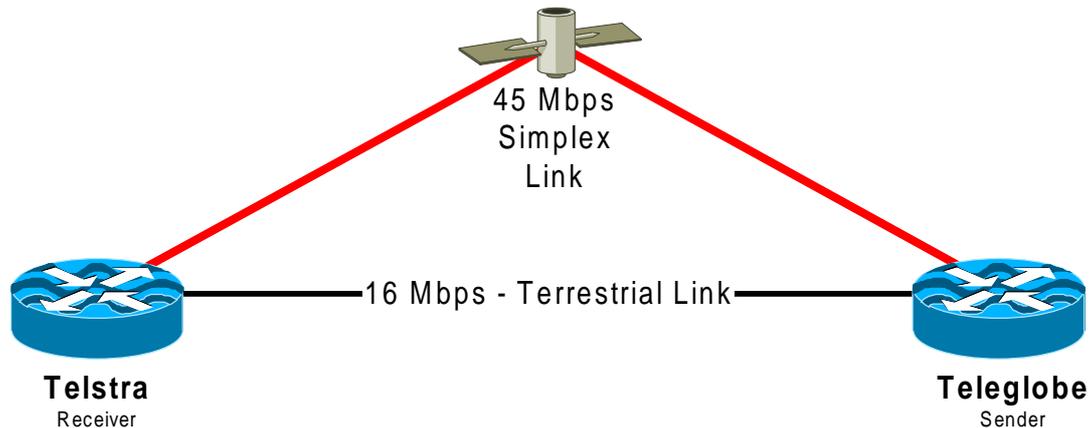

Figure 2 - Telstra/Teleglobe - The First Operational Simplex/Terrestrial System

---

[2] Geoff's lab test on the original design are available at http://www.telstra.net/ops/satellite.html

CISCO SYSTEMS

```
Receiver Configuration                     Sender Configuration

version 11.2                               version 11.2
!                                          !
interface Loopback3                        interface Loopback3
 ip address 3.3.3.1 255.255.255.224         ip address 3.3.3.2 255.255.255.224
!                                          !
interface Ethernet0                        interface Ethernet0
 ip address 2.2.2.1 255.255.255.224         ip address 2.2.2.2 255.255.255.224
!                                          !
interface Serial 0                         interface Serial 0
 transmit-interface Ethernet 0              ip address 1.1.1.2 255.255.255.252
 ip address 1.1.1.1 255.255.255.252         no keepalive
 no keepalive                               ignore-dcd
 ignore-dcd                                !
!                                          router bgp 50
router bgp 25                               timers bgp 5 30
 timers bgp 5 30                            redistribute static
 redistribute static                        neighbor 3.3.3.1 remote-as 25
 neighbor 3.3.3.2 remote-as 50              neighbor 3.3.3.1 ebgp-multihop 5
 neighbor 3.3.3.2 ebgp-multihop 5           neighbor 3.3.3.1 update-source Loopboack3
 neighbor 3.3.3.2 update-source Loopboack3 !
!                                          ip route 3.3.3.1 255.255.255.255 Serial0
ip route 0.0.0.0 0.0.0.0 2.2.2.2           ip route 15.0.0.0. 255.0.0.0 Null0
ip route 3.3.3.2 255.255.255.255 2.2.2.2   ip route 16.0.0.0 255.0.0.0 Null0
ip route 10.0.0.0. 255.0.0.0 Null0         ip route 17.0.0.0 255.0.0.0 Null0
ip route 11.0.0.0 255.0.0.0 Null0          ip route 18.0.0.0 255.0.0.0 Null0
ip route 12.0.0.0 255.0.0.0 Null0          ip route 19.0.0.0 255.0.0.0 Null0
ip route 13.0.0.0 255.0.0.0 Null0
ip route 14.0.0.0 255.0.0.0 Null0
```

Example 1 - Basic Simplex/Terrestrial System using EBGP

The core configuration principle is the manually pointing the routing information for the loopback interfaces down specific paths. The *sender* router (US side) points the *receiver* router's loopback interface to Serial 0, the simulated simplex satellite link. BGP and the forward table will use this information to forward packets down Serial0. The *receiver* router (Australia side) points the *sender* router's loopback interface to Ethernet 0, the simulated terrestrial link.

The key limitation to this technique is its use of eBGP vs iBGP. eBGP requires two Autonomous System numbers. Many Telco/ISPs in the world are now self provisioning their trans-oceanic links[3]. These Telco/ISPs would own both halves of an trans-oceanic circuit. It usually means that the Telco/ISP will set up a small operation in the US and provision the US half from those US facilities. Alternatively, many US North American NSPs are leasing routers and rack space in their own facilities – allowing the International ISP to manage the US side router. Optimally, the Telco/ISP would prefer to use their own AS number for the routers on both sides of the ocean.

Alternative Configuration using one AS Number

Ideally, the routers on both sides of a simplex/terrestrial link should use the same Autonomous System number. The main obstacle with iBGP is that any iBGP router in an Autonomous System (AS) must be fully meshed

---

[3] These Telco/ISPs would own both halves of an trans-oceanic circuit. It usually means that the Telco/ISP will set up a small operation in the US and provision the US half from those US facilities.

CISCO SYSTEMS

with every other router in the (AS). The one exception is if iBGP routers were peered with *an iBGP route reflector* (see Figure 3). This would allow one router to peer with the iBGP AS with out a full mesh - ideal for an trans-oceanic link.
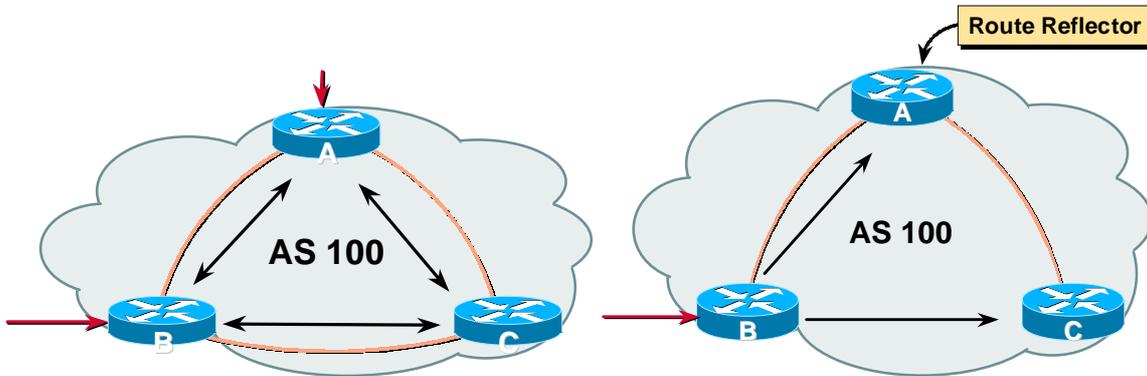

Figure 3 - iBGP Full Mesh vs iBGP Route Reflector

Additionally, many Emerging ISPs outside of the US cannot gain access to straight terrestrial links higher than 2 Mbps (E1). Hence, some type of *inverse multiplexing* (iMux) is required to scale the backbone links (see Figure 4. In many cases, the eBGP Multi-Hop load balancing technique would be used to *iMux* the terrestrial links. In this case, iBGP is used, eliminating eBGP Multi-Hop as an option. Luckily, CEF has two load balancing options that will work together with iBGP.
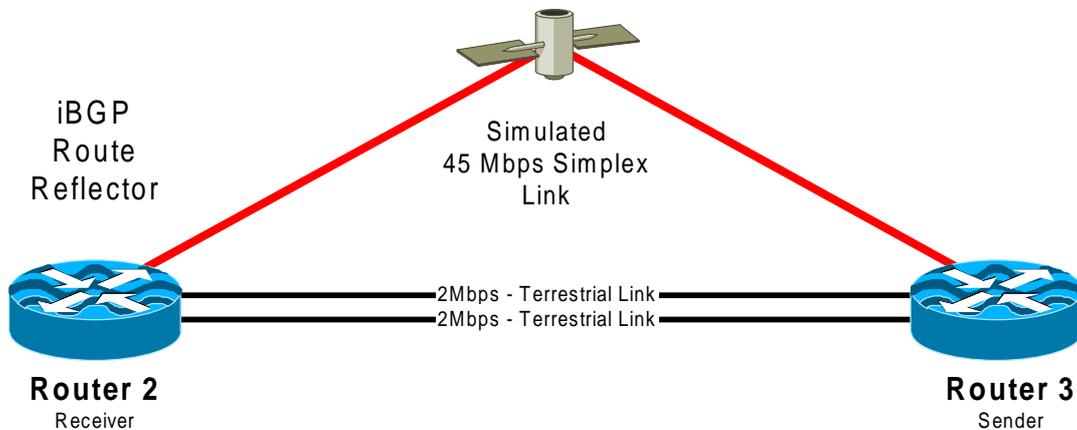

Figure 4 - Simplex System with one AS Number via iBGP

The example in Figure 4 shows the use of the iBGP RR +CEF technique of an trans-oceanic simplex system. The next section will cover some of the aspect in detail.

CISCO SYSTEMS

DETAILED EXPLANATION OF THE TECHNIQUE

Figure 4 provides a visual topology of the iBGP + CEF Load Balancing Technique while Example 2 provides detailed configuration from lab test.

LOOPBACK INTERFACES, STATIC ROUTES, AND BGP

Loopback interfaces as the BGP source in a router are a common practice in the ISP community.[4] In this technique, the loopback interface is used to point traffic down the proper link. The BGP Neighbor command *update-source Loopback0* is used to advertise to the other router that its next hop information uses the loopback interface:

```
router bgp 65000
 neighbor 192.168.10.2 remote-as 65000
 neighbor 192.168.10.2 description Trans-Oceanic Link to Router #2
 neighbor 192.168.10.2 update-source Loopback0
 neighbor 192.168.10.2 next-hop-self
```

This provides a simple tool for static routes to manipulate the direction of next hop information. On the US side, the next-hop information should be pointed down the simplex satellite connection (HSSI 2/1/0 in this example):

```
ip route 192.168.10.2 255.255.255.255 Hssi2/1/0
ip route 192.168.10.2 255.255.255.255 Serial2/0/2 100
ip route 192.168.10.2 255.255.255.255 Serial2/0/3 100
```

By point the next hop to an interface, traffic may be manipulated asymmetrically. Back-up static routes are included in case the satellite link goes down. When the satellite side of a simplex system fails, judicious filtering should be applied to drop all but the most critical traffic.

The static routes on the international router are very similar – pointing traffic to the terrestrial link(s):

```
ip route 192.168.10.3 255.255.255.255 Serial2/2
ip route 192.168.10.3 255.255.255.255 Serial2/3
```

One static route for each of the parallel links is needed to activate CEF load balancing (see section on IP CEF Load Balancing).

iBGP ROUTE REFLECTOR

iBGP Route Reflector was created as a means to scale the full meshed iBGP topologies in an ISP's network. It allows routers to be clustered by logical topologies, leaving the full iBGP mesh to the core of the backbone. In this case we're stretching the original objective of iBGP Route Reflector to enable the trans-oceanic link to remain within the International ISP's Autonomous System.

The configuration is only one line on the BGP configuration for the router on the non-US side of the link:

```
router bgp 65000
 neighbor 192.168.10.3 remote-as 65000
 neighbor 192.168.10.3 description Trans-Oceanic Peer to Router #3
 neighbor 192.168.10.3 route-reflector-client
```

---

[4] See a detailed explanation in *Essential IOS Features Every ISP Should Consider* by Barry Raveendran Greene and Philip Smith.

CISCO SYSTEMS

The BGP Neighbor command *router-reflector-client* changes the iBGP peering session into a route reflector mode. No additional configuration changes are needed on the router in the US. It just does not have to be meshed with every other router in iBGP autonomous system.

IP CEF LOAD BALANCING

CEF addition to IOS has opened the door to two new load balancing techniques for parallel links. One technique load balances per packet – the other based on per flow. Check the following documents for specific details on how IP CEF load balancing works:

>**Load balancing with Cisco Express Forwarding**
>http://www.cisco.com/warp/public/cc/cisco/mkt/core/adap/multi/prodlit/loadb_an.htm

>**Cisco Express Forwarding**
>http://www.cisco.com/univercd/cc/td/doc/product/software/ios112/ios112p/gsr/cef.htm

In the trans-oceanic simplex systems, IP CEF load balancing allows parallel links to be utilized without the use of BGP (i.e. as in eBGP Multi-Hop). Yet, there are limits to the maximum number of T1s/E1s this technique can work. Lab work has demonstrated 8 x T1/E1 with a VIP2-40 and maxed out memory. This would provide up to 16 Mbps iMuxed. Careful thought should be given beyond 8 x T1/E1. Dedicated iMux equipment (i.e. like the Larscom Orion 4000) may provide a more robust iMux solution.[5]

---

[5] The original Telstra solution used Larscom Orion 4000 for iMuxing with a HSSI connection from the Orion to the router.

ISP-LAB-7505-2
```
!
ip cef distributed
!
interface Loopback0
 description Router's Loopback Interface for
Peering
 ip address 192.168.10.2 255.255.255.255
 no ip redirects
 no ip directed-broadcast
 no ip proxy-arp
!
interface Hssi1/1/0
 description Link to Router #3 -
Unidirectional
 ip address 192.168.2.9 255.255.255.252
 no ip route-cache optimum
 ip route-cache distributed
 hssi internal-clock
!
interface Serial2/2
 description internal link to router 7505-3
 ip address 192.168.2.1 255.255.255.252
 ip load-sharing per-packet
 no ip route-cache optimum
 clockrate 2000000
!
interface Serial2/3
 description Link to Router #3
 ip address 192.168.2.5 255.255.255.252
 no ip redirects
 no ip directed-broadcast
 no ip proxy-arp
 ip load-sharing per-packet
 no ip route-cache optimum
 clockrate 2000000
!
router bgp 65000
 no synchronization
 redistribute connected
 neighbor 192.168.10.1 remote-as 65000
 neighbor 192.168.10.1 description Peer
Connection to Router #1
 neighbor 192.168.10.1 update-source Loopback0
 neighbor 192.168.10.1 next-hop-self
 neighbor 192.168.10.1 version 4
 neighbor 192.168.10.3 remote-as 65000
 neighbor 192.168.10.3 description Trans-
Oceanic Peer to Router #3
 neighbor 192.168.10.3 route-reflector-client
 neighbor 192.168.10.3 update-source Loopback0
 neighbor 192.168.10.3 next-hop-self
 neighbor 192.168.10.3 soft-reconfiguration
inbound
!
ip classless
ip route 0.0.0.0 0.0.0.0 192.150.42.1
ip route 192.168.10.3 255.255.255.255
Serial2/2
ip route 192.168.10.3 255.255.255.255
Serial2/3
!
```

ISP-LAB-7505-3
```
!
ip cef distributed
!
interface Loopback0
 ip address 192.168.10.3 255.255.255.255
 no ip redirects
 no ip directed-broadcast
 no ip proxy-arp
!
interface Serial2/0/2
 description Link to Router #2
 ip address 192.168.2.2 255.255.255.252
 no ip redirects
 no ip directed-broadcast
 no ip proxy-arp
 ip load-sharing per-packet
 no ip route-cache optimum
 ip route-cache distributed
!
interface Serial2/0/3
 description Link to Router #2
 ip address 192.168.2.6 255.255.255.252
 no ip redirects
 no ip directed-broadcast
 no ip proxy-arp
 ip load-sharing per-packet
 no ip route-cache optimum
 ip route-cache distributed
!
interface Hssi2/1/0
 description Link to Router #2 -
Unidirectional
 ip address 192.168.2.10 255.255.255.252
 no ip route-cache optimum
 ip route-cache distributed
 hssi internal-clock
!
router bgp 65000
 no synchronization
 network 169.223.0.0
 neighbor 169.222.10.1 remote-as 100
 neighbor 169.222.10.1 next-hop-self
 neighbor 169.222.10.1 version 4
 neighbor 169.222.10.1 soft-reconfiguration
inbound
 neighbor 192.168.10.2 remote-as 65000
 neighbor 192.168.10.2 description Trans-
Oceanic Link to Router #2
 neighbor 192.168.10.2 update-source Loopback0
 neighbor 192.168.10.2 next-hop-self
 neighbor 192.168.10.2 version 4
 neighbor 192.168.10.2 soft-reconfiguration
inbound
!
ip classless
ip route 169.223.0.0 255.255.0.0 Null0
ip route 192.168.10.2 255.255.255.255
Hssi2/1/0
ip route 192.168.10.2 255.255.255.255
Serial2/0/2 100
ip route 192.168.10.2 255.255.255.255
Serial2/0/3 100
!
```

Example 2 - Simplex/Terrestrial System using one AS number