

# ISP TECH TALK



by [Avi Freedman](#)

---

## MULTI-HOMING WITH AND WITHOUT BGP4

**I**n this month's column we'll explore being "multi-homed" - connected to two or more "upstream" Internet providers. We'll look at why ISPs want to (most argue should) multi-home, and at some of the advantages and disadvantages.

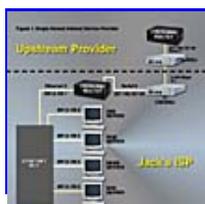
We'll also talk about the Border Gateway Protocol Version 4 (BGP4), the magic protocol that runs the Internet. The myth goes that you must run BGP4 to "multi-home," but this isn't quite true. We'll be talking quite a bit about BGP4 in future columns - for now, we'll cover just the basic concepts.

Before we can get to the meat of the discussion, though, we have to go over some basic topics.

### **BEING CONNECTED TO THE INTERNET: PART I**

To be connected to the Internet, you need to be able to send data (in the form of IP packets) to every valid IP address (host) on the Internet. Equally important is that all of the hosts on the Internet know how to send data to you.

### **BEING "SINGLE-HOMED"**



**Figure 1**

If you're "single-homed" you are connected to the Internet by one "upstream provider." All of your non-



local IP traffic (traffic destined to the Internet) is going to be sent to that provider, and all of your non-local IP traffic that comes from the Internet will come in from that provider.

See Figure 1 for a diagram of a single-homed ISP. Jack's ISP has an Ethernet and one Class C's worth of IP address space. Actually, it's called a /24, or slash 24 - IP Blocks that are 256 addresses in size in CIDR address space are called /24s. For more information on this, please see April's column.

All of Jack's ISP's machines are numbered in the 207.8.130.0/24. Jack's ISP has a router at .1; servers from .2 through .6; and a terminal server at .7. Furthermore, Jack's ISP's dial-up users get dynamically- assigned addresses in the 207.8.130.129 to 207.8.130.159 range, so the terminal server also sometimes responds to those IP addresses. The router in the network knows about this.

First we'll talk about the internal routing of Jack's ISP's network - and review the basics of IP routing - and then we'll talk about how the outside world can get data to Jack's ISP.

## AN IP ROUTING REFRESHER

Every machine that talks TCP/IP has an "IP routing table." This routing table tells the machine where to send IP packets. Each IP packet has a **source** address and a **destination** address.

When a machine's IP software sees a packet, it has to figure out where to send it. To do this, it examines the destination address of the packet and matches it against the "best" route in the routing table. To figure out which route is the best one, the IP software tries to find the "tightest fitting" - or "most specific" route. (See Figure 2). The "most specific" route is the smallest possible route which "contains" the destination IP address of the packet in question. The smallest route is the route which represents the fewest number of IP addresses (i.e. a Class C, or /24, is smaller than a Class B, or /16). The biggest route is the default route - also called **0.0.0.0**.

So let's look inside the router 207.8.130.1. There are four routes in the IP routing table:

("Network Number")	("Specificity")	Specificity	
0.0.0.0	0.0.0.0	(/0)	207.106.127.46 Serial0
207.106.127.44	255.255.255.252	(/30)	207.106.127.45 Serial0

207.8.130.0	255.255.255.0 (/24)	207.8.130.1	Ethernet0
207.8.130.128	255.255.255.192 (/27)	207.8.130.7	Ethernet0

Let's say the router gets a packet destined for the web server, 207.8.130.3. It looks at its routing table and finds 207.8.130.0/24 and 0.0.0.0/0 which match 207.8.130.3. 207.8.130.0/24 is the most specific route, so it sends the packet out the Ethernet0 interface.

Why did it get this packet? Probably, someone on the Internet is sending a request to the web server and the packet came in via the serial port. Note that it doesn't matter where the packet came **from** or what the **source** IP address of the packet is. Each packet is routed based only on the destination address.

Now let's say that the router gets a packet destined for one of the dial-up users, 207.8.130.134. It looks at its routing table and finds 207.8.130.128/27, 207.8.130.0/24, and 0.0.0.0/0 as matching routes. 207.8.130.128/27 is the most specific route, so it looks at the next-hop of that route, which is 207.8.130.7. Then it looks and finds that the best route for 207.8.130.7 is 207.8.130.0/24, which has a next-hop of Ethernet0, so it sends the packet out Ethernet0 to 207.8.130.7.

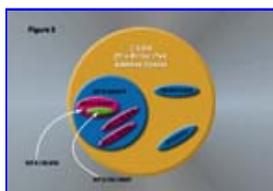
## DEFAULT ROUTING

Each machine on Jack's ISP's network also has a routing table which probably looks like:

### Starting IP

address Interface ("Network Number")	Network Mask ("Specificity")	CIDR Specificity	Next-Hop
0.0.0.0	0.0.0.0	(/0)	207.106.127.46 Serial0
207.106.127.44	255.255.255.252 (/30)		207.106.127.45 Serial0
207.8.130.0	255.255.255.0 (/24)		207.8.130.X Ethernet0
207.8.130.128	255.255.255.192 (/27)		207.8.130.7 Ethernet0

*Where X is the address of the particular local machine.*



**Figure 2**

Note that you **could** omit the 207.106.127.44/30 route from the local routing tables, but that would force all data destined to the terminal server to "bounce" off of the router at 207.106.127.1. If

each local machine isn't specifically told how to get to the addresses 207.8.130.129 to 207.8.130.159 (the 207.8.130.128/37 route), then packets destined to the local modem users will be sent to the router, which will then send them to the terminal server. This means that those packets must traverse the Ethernet twice, which chews up bandwidth on the Ethernet. So it's always better to put all of your local routes into all of your local machines somehow - but as your network grows you may find that this is a pain to do.

Now, for a local server to get to any other machine or modem user on your network, it will have a **specific** route (some route other than the default route) telling it how to get there (the next-hop). But there won't be a route in the routing tables on any of your machines for an IP packet destined to a host out on the Internet - so the route that will "match" packets destined to the Internet is the "default" route - 0.0.0.0/0.

## STATIC AND CONNECTED ROUTES

How do the routes **get** into the IP routing tables? Some are "connected" routes - they are routes that are associated with the interfaces on the machine. In the case of the router, 207.106.127.44/30 is a connected route associated with Serial0, which has a local IP address of 207.106.127.45. 207.8.130.0/24 is a connected route associated with Ethernet0, which has a local IP address of 207.8.130.1.

The other routes in the routing table had to be inserted "statically" as part of the startup procedure of the router. These "static routes" are "installed" whenever the associated interface is up. 0.0.0.0/0 is a static route with a next-hop of 207.106.127.46 (the provider's router at the far end of the T-1); and 207.8.130.128/27 is a static route with a next-hop of the terminal server on the local network, 207.8.130.128.

A very important note: Most well-behaved routers will delete any routes associated with interfaces that are down. For example, the default route 0.0.0.0/0 and the connected route 207.106.127.44/30 will go away if the connection to the remote provider over the T-1 goes down.

Some of the outbound load-balancing we will talk about later relies on this behavior. But most UNIX machines - and some routers - don't obey this fundamental rule of routing. If you intend to multi-home yourself with a particular router, try pulling the serial cord from one interface. Make sure that the routes associated with that interface go away.

## BEING "CONNECTED" TO THE INTERNET: PART II



*Figure 3*



To be connected to the Internet usefully, every machine on your network must have an IP address that is "reachable" globally. If you have a default route to your router's WAN interface, then it's your provider's job to figure out how to deliver traffic destined to the world.

But an equally (and arguably **more**) important job is to announce to all of the key routers on the Internet that they know how to send packets to you. If even one key router on the Internet doesn't know how to get to you, you do NOT have global Internet connectivity. A router knows how to get to you because it has a route in its IP routing table that tells it how to send data to you - it's the same principle as the IP routing going on in your local network.

## **ROUTE ANNOUNCEMENTS WITH BGP4: BEING "SEEN" BY THE INTERNET**

The key routers on the Internet are the routers in the "defaultless core." These routers have no default route. If you want to be reachable by all of the hosts on the Internet, there must be a matching route for your IP space in all of these routers.

How do those routes get into those key routers? Different providers "peer" with each other. They have private and public "Exchange Points" where they exchange "route advertisements." These advertisements are "promises" to carry traffic to various sections of the IP space.

Providers use the BGP4 protocol to advertise routing information to each other. Routers at the "border" of various networks ("Autonomous Systems" in BGP-speak) exchange routes with each other via "peering sessions." (See Figures 3 and 4).

Route announcements filter upwards from the edge ISPs, through various transit providers, to the top-level providers. By the time all of this occurs, all of those key routers know about your routes.

## **WHAT ROUTES GET ADVERTISED FOR YOU WHEN YOU'RE SINGLE-HOMED?**

If you have your own address space (old Class C blocks) - or more recent CIDR allocations from the NIC - your provider(s) will advertise those blocks for you as BGP route announcements. If you are using IP space that you got from your ISP, chances are good that you're part of one of their "aggregates" (or "IP blocks"). Aggregate routes are larger routes that cover multiple /24s. If you have been allocated 207.8.130.0/24, a "whois" on 207.8.130.0 will show you who owns the underlying IP block,

and what the specificity is. In the case of 207.8.130.0, a "whois" will show you that it's part of the 207.8.130.0/17 block (half of a Class B in size). If you were to look at the "global routing table" you'd see that the owner of the 207.8.128.0/17 block, Net Access, advertises 207.8.128.0 "as" a /17 (meaning, with the specificity of /17).

If you are single-homed, the only way that anyone in the world can send data to you is through your one provider. So it makes no sense for your provider (Net Access, in this case) to advertise 207.8.130.0/24 as well as 207.8.128.0! Let's say that Net Access advertised 207.8.128.0/17 and 207.8.130.0/24. The more specific route for any of your IP addresses would be 207.8.130.0/24, so that is the route that would get used. If that route were not advertised, however, 207.8.128.0/17 is the route that would get used. The net effect is the same, and having the more specific route for 207.8.130.0/24 out there is just one more route for every core router to know about and have to spend memory and CPU time on.

If you're multi-homed, however, at least one of your providers will need to announce a specific route to match your blocks. More on this shortly.

## **MORE BGP TERMINOLOGY**

As we said, BGP4 as we're looking at it, is a protocol spoken between Autonomous Systems to advertise routes. BGP4 is spoken between two routers and routes are exchanged via "peering sessions." Each Autonomous System has an "Autonomous System Number" (ASN). Net Access's is 4969. UUNET's is 701. There are hundreds of ASNs in use on the Internet.

In addition to the usual information (starting IP address, specificity, next-hop information), each BGP route has a property called the AS\_PATH.-the path of Autonomous Systems that the route took from the very first advertising router. It's just a list of the ASNs in the reverse order of advertisement.

Every time a route is advertised by one router to another over a peering session, the receiving router "stamps" the incoming route with the ASN of the remote router by prepending the remote ASN to the beginning of the AS\_PATH. When UUNET hears a route from Net Access, for example, it notes that it got the route from AS 4969 (Net Access's ASN).

Remote routers will generally decide which is the best route of multiple BGP routes by picking the route with the shortest AS\_PATH (that is, the route that has traveled few the fewest providers) to get from start to finish.

## **REVISITING THE SINGLE-HOMED PROVIDER**

Jack's ISP (the average single-homed provider) will typically have one route announcement "out there" for each of its routes. This makes sense - you can only get into Jack's ISP's network by one path - sending data to Net Access.

Jack's ISP's upstream provider must either peer with all of the other key networks on the Internet (so that routes for Jack's ISP get into all of the default-free core routers), or must get "transit" (be a complete or partial customer) of some other providers.

For example, AGIS (ASN 4200) will peer with very few other networks, so Net Access (ASN 4969) must get some Autonomous System who does peer with ASN 4200 to advertise Net Access's routes to AGIS, or AGIS's routers will not know about Jack's ISP's (Net Access's) routes.

## **WHY BE MULTI-HOMED?**

Additional bandwidth and redundancy are the two major reasons that ISPs multi-home. Multi-homing (adding another upstream provider) is expensive, costing you at least \$1,000 to \$2,000 extra per month, so most providers put it off as long as possible. This is a mistake! In even the smallest areas there is now generally competition among ISPs, and the ISP with the most reliable network can attract customers willing to pay for quality and reliability.

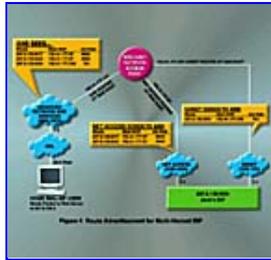
If you've got a T-1 to the Internet and it's running at 70 to 80 percent capacity, then you'll start to see slowness. Or, worse, you could have the dreaded packet loss on your line due to congestion. Your customers will start leaving you in droves if this happens. Additional bandwidth is a major reason that many finally break down and multi-home. But consider this: If you are single-homed you are **completely** dependent on the up-time and quality of your one upstream provider. Also, you're completely dependent on your one border router and your one link to that upstream provider. If any of those components fails you're completely off the Net!

There's only one route out there for you. If your router or T-1 goes down, then your provider's not going to be able to send data to you. And if your provider becomes disconnected from the Internet or has some major internal routing problem, then you're also disconnected from some or all of the Internet.

## **MULTI-HOMING: ROUTE ADVERTISEMENTS**

When you're multi-homed, your routes should be advertised by **both**

providers. If you have two or more routes out there for each one of your IP blocks, you can sustain a complete loss of a T-1 - or severe problems with one of your upstream providers - and not have your customers notice a thing.



**Figure 4**

Let's look at Figure 4. In this example, we'll use our friend Jack's ISP, now connected to both Net Access and to UUNET, and a dial-up user on AOL who wants to query a web server on Jack's ISP's network, 207.8.130.3. AOL's Internet provider is

ANS, so AOL sends the packet to somewhere in ANS's network. ANS sees three BGP route advertisements for 207.8.130.0:

207.8.130.0/24, from Net Access (ASN 4969); 207.8.128.0/17, from Net Access (ASN 4969); and 207.8.130.0/24, from UUNET (ASN 701).

If those routes weren't there, ANS (and thus AOL) would have no connectivity to Jack's ISP's network. If there was only one route of specificity /24, then there'd only be one "primary" path to Jack's ISP's network. But since there are two routes for 207.8.130.0/24 (also called two views of this one route), ANS will choose one of those as the best path and will use it. Incoming load-balancing works on the principle that roughly half of the networks out there will choose one of multiple paths for any network and that the rest will pick one of the other paths. We'll go into more detail on this in future columns, and we will discuss the actual algorithms used by routers and providers to make and tune these decisions.

## **MULTI-HOMING WITH BGP**

If you speak BGP to your providers, then **you** are the one originating the route advertisements that put your Autonomous System on the Internet. You have some degree of control over how those routes are advertised - we'll be talking about methods like AS\_PATH padding and changing route specificity in the future.

And you can also take "full routes" from one or both providers and let your router use BGP "selection algorithms" to pick the best path for each outgoing packet. This will generally give you a moderate bump in "quality" of connectivity - perhaps as much as 5 or 10 percent less packet loss to certain destinations.

Also, if you have nothing better to do than tune your connectivity, you can start to fine-tune outgoing data flow. For example, if you notice that one provider has poor PSI connectivity although it's got the best BGP

routes to PSI, then all of your PSI traffic is sent out to your other provider.

## **MULTI-HOMING WITHOUT BGP**

But you don't need to speak BGP to have your routes be advertised by both providers. They can do it for you, just as your single upstream provider does it for you when you are single-homed. In this case you are part of their Autonomous System for routing purposes, and you have to call them on the phone and ask them to change or add route advertisements for you.

Providers are generally happier to do the work of injecting routes into BGP for you, since it's a bit of work for them to set up a BGP peering session for you and to carefully filter the routes they hear from you over that session.

To control outgoing data flow from your network, all you have to do is add another default route - there's no need to use BGP to take a list of all 45,000 or more routes on the Internet. It's true that you have more control and flexibility if you have those routes to work with, but the difference is usually not that great between selecting the best BGP route and round-robin default-routing.

## **SO WHICH IS BETTER?**

You can get 90 percent of the benefit of multi-homing without speaking BGP yourself. But as long as you're paying for two T-1s, why not take full advantage of them by speaking BGP?

The first problem is that running BGP on your own requires a lot of knowledge (and hopefully experience). Screw-ups in BGP route advertisements can be felt all over the Internet, and are treated very seriously. If you don't know what you're doing you'll wind up with a less reliable network instead of a more reliable one.

The second problem is that it takes a fairly expensive router to take full routes from your providers. Each route takes a certain amount of CPU and memory, and the current Internet routing "load" of 45,000+ routes is too much for almost all low-end routers; however, you can take less than full routes - or even just use BGP to announce your routes and still just default-route for outbound traffic - with a smaller router like a Cisco 2501.

So the decision is yours. Most pro-viders will let you speak BGP to them, but will be somewhat unsympathetic if you misconfigure your end and break your connectivity. On the other hand, there are some advantages to

speaking BGP at your end.

For the rest of this column, we'll finish explaining how multi-homing without BGP works.

Next month's column will go into more BGP details and will show you how to safely announce your routes via BGP with router configurations. This can be done even with a low-end router. We will also discuss some of the options that you have for managing the incoming routing information which determines how you send your outbound traffic to the Net.

Now we'll explain both sides of being multi-homed without speaking BGP to your provider. The first is load-balanced default routing, which is the art of getting data out of your network. The second is how your providers advertise your routes for you, which controls how other providers hear your routes - and thus how data flows in to your network.

Keep in mind that these two things (sending data out with default routes and getting others to send data to you based on route advertisements from your providers) are the key to being connected to the Internet.

## **DEFAULT ROUTING TO MULTIPLE PROVIDERS**

If you go from one provider to two, then you're going to already have a default route installed. You have to decide how you want to send data out from your network. Once you've decided what you want to do, you can go about implementing it.

Remember, the default route controls where you send data that is not destined for your local network. So adding default routes is the way you both increase your outbound bandwidth to the Internet and add the redundancy that multi-homing provides.

Let's say that Jack's ISP has his original default route installed. In the configuration mode on his Cisco, he enters:

```
ip route 0.0.0.0 0.0.0.0 137.39.34.65  
int s0  
ip route-cache  
int s1  
ip route-cache
```

Now there are two routes (0.0.0.0/0 through 137.39.34.65 with metric 0; and 0.0.0.0/0 through 207.106.127.46 with metric 0) of equal specificity (both are /0 routes) and of equal metric (both have a metric of 0).

Both routes will now be installed in the Cisco's IP routing table, and the Cisco will now use both T-1s for outbound data. If one of the serial

interaces goes down (because one of the T-1s goes down), one of the default routes will go away, but one will still be there to make sure your router has a way of sending data out to the Internet.

Also, this example turns on **ip route cache** on both interfaces to ensure that the load balancing is done on a **per-connection** basis rather than by simply sending every other packet out each interface. Note that sending the packets out in a round-robin fashion makes TCP/IP unhappy.

## **MULTI-HOMING TO THE SAME PROVIDER**

Instead of multi-homing to different providers, you can multi-home to the same provider. Net Access, for example, allows all of their dedicated-line T-1 customers to establish a backup link via the local Frame Relay "cloud." Customers install a backup default route so they'll only use the Frame Relay path if the primary T-1 goes down (since they're not paying for additional bandwidth, only redundancy).

Customers who are happy with their upstream providers might want to run multiple links for added bandwidth or reliability. Sadly, few IP providers are so reliable that customers are tempted to do this, but it does happen. Also, paranoia dictates that an Act of God can always strike a particular upstream provider or that provider's POP in a particular city. You almost always get better reliability by multi-homing to multiple providers.

If you do decide to multi-home to the same provider and run two equal-bandwidth pipes into either the same router at the provider's side, or routers at the same POP, you can set **no ip route-cache** on your serial interfaces and things should work well. Why? Because all of your outgoing packets travel almost the same path and have almost the same latency this way.

## **BEING ADVERTISED BY MULTIPLE PROVIDERS WITHOUT PI-SPACE**

We'll use Figure 4 again. Remember April 1997's document on getting provider-independent (PI) space? The reason it's so important to have "your own" ip space is that without it multi-homing is quite tricky and requires a lot of cooperation from your original provider.

Let's say that Jack's ISP is using 207.8.130.0/24. Jack's ISP's provider (Net Access) has 207.8.128.0/17. So Net Access announces only 207.8.128.0/17 to the world. There is no global advertisement for 207.8.130.0/24. In this case, any packet destined to 207.8.128.0/17 will

be picked up by the less specific (more general) route 207.8.128.0/17. We talked about this before, when examining how Jack's ISP's routes are seen by the outside world when single-homed.

Now Jack's ISP wants to multi-home. So Jack's ISP buys a T-1 from newprovider. Jack's ISP sets up BGP with both Net Access and newprovider. Suddenly, the world sees two routes for Jack's ISP:

207.106.0.0/16, advertised by Net Access; and 207.8.130.0/24, advertised by **newprovider**. Remember, the most specific route always wins, so newprovider will wind up carrying almost all, if not all, of Jack's ISP's incoming traffic! Certain parts of Net Access's network may actually prefer newprovider's T-1 to get to Jack's ISP! In fact, this won't happen with Net Access's network, but some providers will prefer externally-heard more specific routes to get to dual-homed customers.

This state of affairs gets you **backup** (if newprovider goes down, your data will once again flow through Net Access based on the aggregate advertisement), but doesn't **load-balance** your inbound traffic. To have incoming load-balancing, you need to advertise two routes of **equal** specificity. Even so, inbound load-balancing is tricky and depends on who peers with whom and how different providers "hear" each other - but unless there are two "competing" equal-specificity routes out there for you, there's no hope if inbound load-balancing.

The problem is that most largeish providers use something called **aggregate-address statements** - and they certainly have some sort of filter to keep the more specific routes floating around inside of their networks from being advertised to the world. Remember, the world only wants to hear about 207.8.128.0/17 if the little, more specific, routes inside of 207.8.128.0/17 are not multi-homed.

So what does Net Access have to do? Blow holes in their aggregation statement. One way or another, it's going to take modifications in Net Access's "border" routers to make incoming load-balancing work properly for Jack's ISP. Net Access may not want to do this. (In fact, Net Access does do this, but some providers don't.) Basically, everywhere that Net Access peers with anyone else (and this is usually at least 5 to 10 places), they have to modify their aggregation statements or other filters to "allow" your more specific route announcement to pass through.

If Net Access blows holes in their aggregation filter using "suppress maps," which we'll talk about in a few columns, then Net Access can announce:

**207.8.128.0/17**  
**207.8.130.0/24**

And newprovider will announce:

### **207.8.130.0/24**

Thus, Net Access's other customers who "live" inside of 207.8.128.0/17 will still be advertised, but 207.8.130.0 will now have two advertisements at a specificity of /24, so some load-balancing may happen for inbound traffic to Jack's ISP.

This is one reason that it's important to choose a primary provider based on how cooperative they'll be when you want to multi-home.

## **BEING ADVERTISED WITH PI-SPACE**

If you have Provider-Independent (PI) address space, announcing your routes is easier, since you're not a part of anyone else's aggregate block. Examples include old Class C or Class B blocks that you or your customers might have, or newer CIDR allocations that you (or, if you sell to ISPs, your customers) might have.

## **SUMMARY**

We've talked about how the world sees you when you're single-homed and multi-homed, and about what BGP does. We've also talked about how you can be multi-homed effectively without using BGP.

Next month we'll go into more detail on BGP. We'll go through the details of configuring a router to speak BGP to multiple providers - and to announce and receive BGP routes.

---

Copyright 1998 Mecklermedia Corporation.  
All Rights Reserved. [Legal Notices.](#)  
[About Mecklermedia Corp.](#)

**Colorado Offices**  
13949 W Colfax Ave Suite 250, Golden, CO 80401  
Voice: 303-235-9510; Fax: 303-235-9502



 [Fable Of Contents](#)