



# Beginning BGP

**Peter J. Welcher**

---

## Introduction

At the time I'm writing this, it is time to register for Networkers / CCIE recertification. I just signed up for Denver. You'll probably be reading this around that time. I'm looking forward to both the talks and to once again seeing some of the many fine networking folks I've met!

This month I'd like to bring us back to a topic more closely related to router configuration and network design. I've been running into increasing numbers of people who are re-considering their Internet connectivity. BGP seems to be on peoples' minds. (Could this be a form of Spring fever?)

I'd like to look at some design options and then start looking into the things we can configure with BGP-speaking Cisco routers.

## When Do We Need BGP?

As I learn more about networking and BGP, it strikes me that the first question you've really got to ask yourself is "Why in the heck do I want to do BGP?" Now, this is not necessarily a question implying "Do I need my head examined?", but it certainly is a question that's worth pausing and answering.

The reason I'm suggesting this is that BGP raises the complexity level of your network. At the very least, you're going to be configuring and troubleshooting Yet Another IP Routing Protocol (YAIRP). You might also be doing some redistribution -- but I'm going to try to talk you out of that later in this article.

There are also some costs to running BGP. If you're receiving a fairly full Internet routing table, you're going to want adequate CPU and RAM in your router. That costs. Complexity costs, time and support hassle. Your ISP may also want to charge higher support fees for speaking BGP to you. First, it requires more work on their part to set it up and test it. Second, when something breaks, they'll need to work with you to resolve the problem. (Be nice to the person on the other end of the phone, or the day will come when something's broken and they don't return your calls.)

Now, many of us are aware that BGP comes in two flavors, external (eBGP) and internal (iBGP). We'll come back to this later. But another way to think about BGP is that it comes in two forms, corporate and ISP (concentrated and industrial strength BGP?). And when you're thinking about complexity, you might consider "corporate" BGP as being intermediate on the scale of difficulty. ISP "industrial

strength" BGP can range from advanced to nasty to worse. Since I have not yet served my MCI, WorldNet, UUNET apprenticeship, I'm going to keep quiet about ISP BGP (at least in this article).

All those readers who have taken the Mentor Technologies' Cisco-certified routing training courses indubitably know that BGP is an Exterior Gateway Protocol. As in, your routers speak BGP to routers in another Autonomous System (AS), generally because of policy and control issues. The other AS might be your ISP, but it also might be another company. You might be using BGP here to maintain an "arms-length" relationship.

Another use of BGP is to break a large network up into smaller pieces (hierarchical networks scale better). For instance, a company with a presence on each continent might consider using BGP between continents. If the company is running OSPF protocol, it might be better to make each continent an OSPF "autonomous system", and run BGP between them as a clean separation (vice dealing with OSPF redistribution into OSPF).

The big difference between talking to an ISP and to another company is that generally, the ISP has bigger routing tables. If they use a full Internet routing table, it might run to somewhere around 50,000 entries. When you're exchanging that much information, you generally want to do it incrementally and reliably. BGP does both. Because BGP speakers maintain a copy of information received from each peer, I've been thinking of BGP as in effect a BGP routing table synchronization protocol, intended to conserve WAN bandwidth (although stability of links may have an effect on that).

To get back to the original question, "**why BGP**" -- well, if you're thinking about control, about exchanging routing information with another entity, then BGP is one tool in your bag of tricks. It may be complex -- but then again, trying to solve the problem you're trying to solve may be equally complex if you use the wrong tool, and more error-prone in addition.

There is one pretty clear cut case where you probably do not want to be running BGP. In Diagram 1 below, the first network shown is a simple connection to an ISP. There is only one connection, and only one ISP. Why not just point default at the ISP, redistribute and advertise 0.0.0.0 /0 into your interior routing protocol (IGP), and have done with it? The picture intentionally shows a 25xx router: you can do this with just about any router model. Nice and simple!

There's another situation where you don't want to run BGP. If your ISP or business partner doesn't want to do it, don't! This is a two-way situation, and you cannot unilaterally decide to run BGP. In Diagram 1, the ISP would probably prefer to use statics routes on their side of the connection, for control and simplicity.

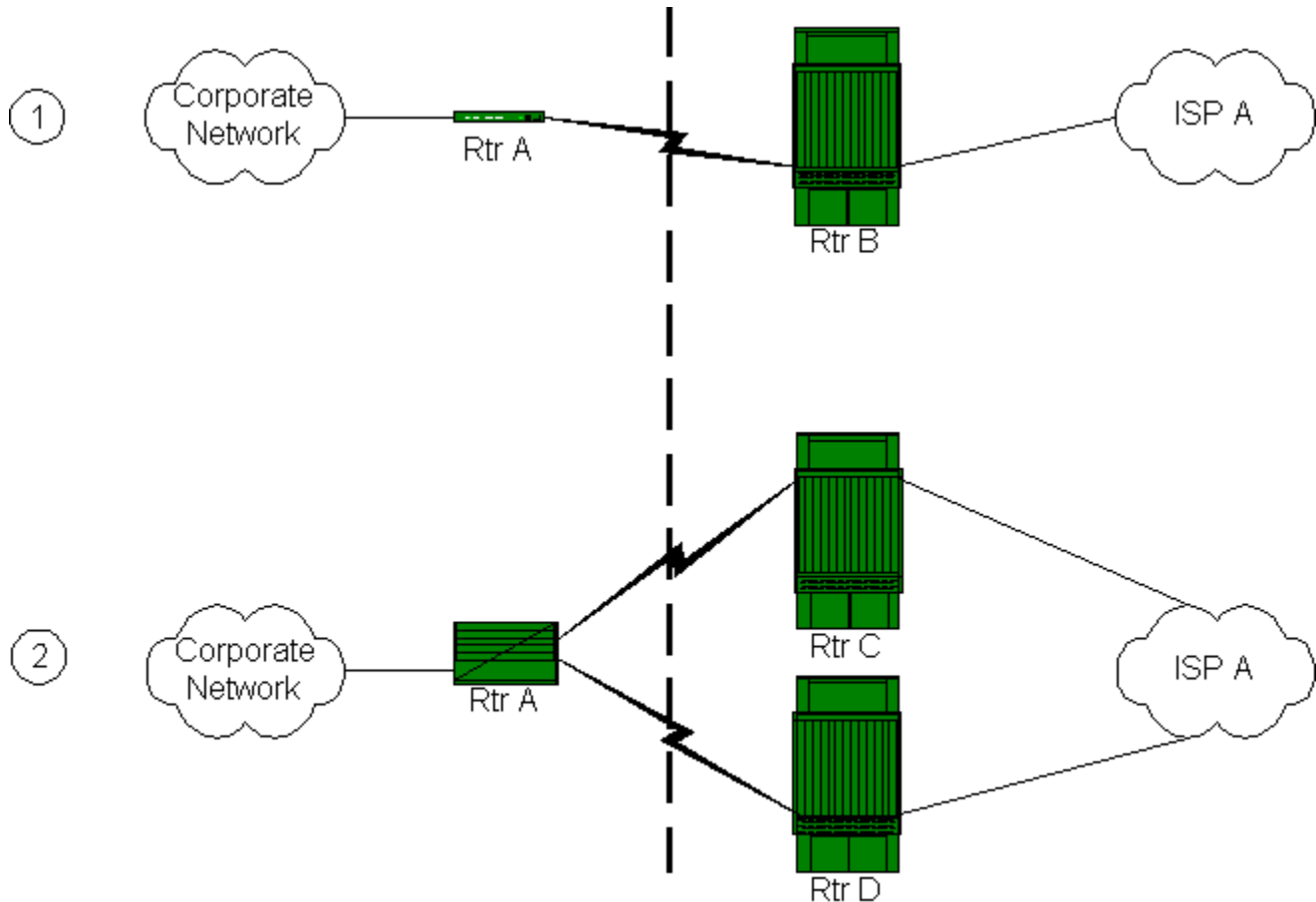
## Internal and External BGP

BGP uses a formally registered Autonomous System number, at least with entities that will advertise the AS number to the Internet. If your router and its BGP neighbor have different AS numbers, they're speaking external BGP, eBGP. When you're speaking BGP to your ISP, you're quite possibly doing eBGP. When two BGP neighbors have the same AS number, they're speaking internal BGP, iBGP. Inside your corporate network, two BGP speakers will probably be using iBGP, unless you have numerous Internet connections and have broken your network into multiple registered AS's.

The range of private BGP AS numbers is 64512-65535. If you're a "stub AS", your ISP may give you one of these.

## Taxonomy of Design

In the following diagrams, I've tried to show some options in order of increasing complexity.



**Diagram 1:** one connection to ISP: use default routing!

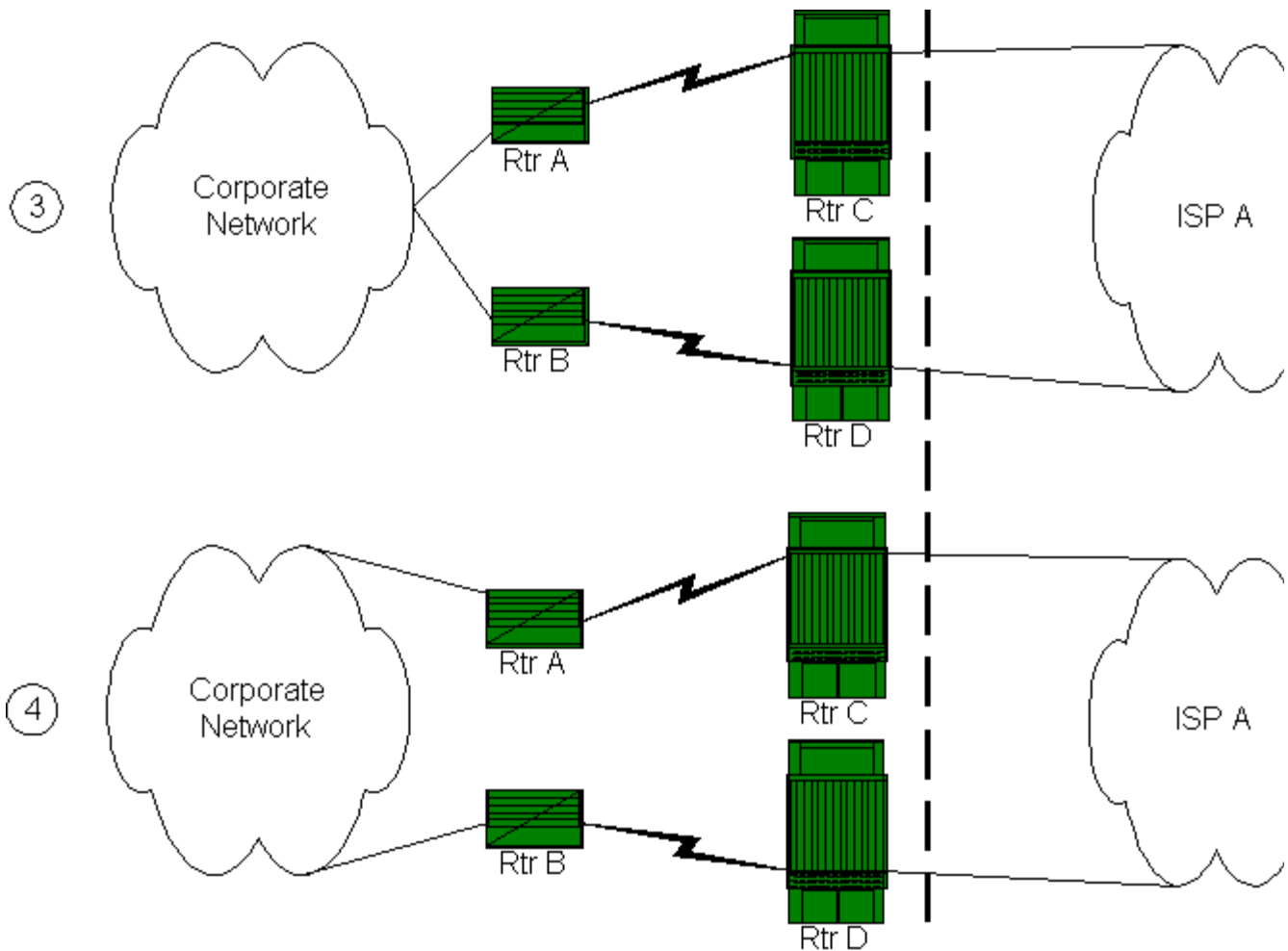
**Diagram 2:** one corporate gateway router, two connections to ISP (possibly geographically diverse).

If the two ISP routers are at the same site, there's no obvious gain to using BGP. Use two default routes on Router A. IP routing will load balance.

If the two ISP routers (C and D) are at different sites, then you have to decide how badly you want optimal routing. If you use dual default routes, traffic may get sent to D when C would have been a better choice, and vice versa. That's the price you pay for not knowing about routing metrics in the ISP's network. You don't know which of C or D is closer to a packet's destination. Sub-optimal routing can also cause extra latency.

If you want optimality, you have two further decisions to make. First, what protocol will your routers speak to your ISP. And second, how much optimality you want. Whether or not you speak BGP, the choice is learning routes and metrics to sites within your ISP (a partial feed of your ISP's networks), or also learning some or many Internet routes as well.

I broke diagram 2 out as a separate diagram because with only one router, if it speaks BGP, it probably only speaks eBGP. It has no (and needs no) iBGP peers.



We then come to three progressively more complex situations.

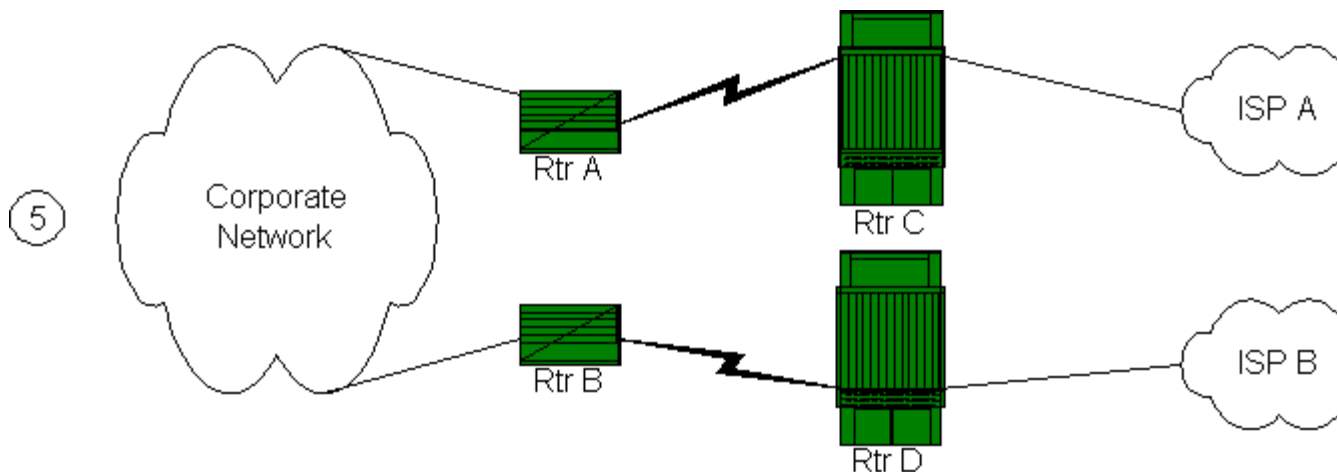
**Diagram 3:** two corporate gateway routers (in close proximity), two connections to ISP (possibly geographically diverse).

**Diagram 4:** two corporate gateway routers, geographically diverse on both ends.

**Diagram 5:** two geographically diverse corporate routers, different ISP's on the other ends of the links.

Not shown in the diagrams:

- Speed of the links. They might be equal, they might not.
- The routers A and B in diagrams 3 - 5 might be talking iBGP to each other. They might be doing iBGP, they might not.
- WAN connections between A and B. Depending on your strategy, it might be a good idea to have a high bandwidth connection between A and B.



Perhaps the first thought here is, "what are we trying to do?"

Presumably the geographic diversity and differing carriers are there to try to increase availability, and to minimize the effects of outages. If the link speeds are equal, you probably want load balancing behavior, with some traffic going via A to C, and the rest via B to D. I'm generally a fan of equal speed links and load-balancing, since then when something fails you're using links that are already in regular use, and you probably won't get too many surprises. If the link speeds are unequal, then you probably want most or all of your traffic using the higher speed link, perhaps with some traffic using the other link. This is perhaps popular when the connections come out of two different budgets, with two organizations acting as mutual backup link.

Once you've settled "what are we trying to do?", then there's "how do we do it?".

Do you use static and floating static routes? In effect, this tells the router "go this way, unless it fails, in which case go this other route." Alternatively, do you run a dynamic routing protocol to the ISP(s)? If you do, just how much do you want to know about routes in the Internet? Do you just learn information about some of the ISP network(s)? Or all about each ISP network? Are some or all of a full Internet routing feed as well? It still comes down to trading small routing tables and simplicity for optimal routing. Optimal routing costs. Unless you have a valid business requirement for optimal routing, don't bother.

If you do run a dynamic routing protocol between your routers and the ISP(s), that makes routers A and B more or less well-informed. Do they pass this information on to your corporate internal network? Why do other router want to send packets to A and B, and how do is the exit router chosen?

This interacts with your design. If you have a moderate to high bandwidth (WAN) direct connection between routers A and B, then that may make your strategy simpler. Just send the packets to A or B, whichever is closer, and let them figure out which is the better exit point. They can use the link between them if the packet is at the "wrong" router of the two. If you're willing to use suboptimal routes, just push packets out through the closer of A or B, and let your ISP(s) deliver them. That costs less, since you're not paying for a WAN link between A and B, and is simpler too.

To add to the fun, we're seeing folks trying to have two geographically separated gateway routers with two CheckPoint (or other) firewalls. Talking routing to or through a firewall product means you'd better be good friends with the firewall administrator! So add to our list of questions: just how does all of this interact with the firewall(s)?

Tip: in thinking through and planning BGP routing, I find it useful to think in terms of the following traffic flows:

- Corporate to Corporate
- Corporate to Internet
- Internet to Internet
- Internet to Corporate.

Two of these are obvious, I hope. Your Corporate to Internet traffic needs a route to and through the gateway and out into the world. And traffic coming from anywhere in the Internet needs to be able to find your gateway and reach your network. The other two are sometimes sources of problems or surprises: do you want internal corporate traffic traversing the Internet? And do you want your network to possibly become a transit network for Internet traffic, either due to an outage in your ISP or due to an error on your part?

Chances are, you want to be a "multi-homed nontransit AS".

## Default Routes

In general, default routes are A Good Thing. If you're working with situations like diagrams 3 through 5, you might seriously consider having A and B originate default. One way to do this is to have a static route pointing out the serial port, and redistribute it into your IGP (with an access list filter if you're being careful). The route will not be advertised if the serial link fails. If you do this, traffic will be drawn to the closest exit point. With a modicum of care in your design, this provides geographical load balancing -- it's fairly easy to arrange things so each exit point is "close" to half of your network.

OSPF, IGRP, and EIGRP allow other approaches (originate default, default network) that have the same net result.

I'm not a fan of configuring static default routes into **all** your routers. As things change, that's more maintenance work to keep current. If you have two exit points, this approach is not dynamic in responding to link outages. You do not want to have to reconfigure lots of routers when an outage occurs!

See also <http://www.netcraftsmen.net/welcher/papers/default.htm> .

## Redistribution with BGP

Don't.

If you are running BGP to ISP's, and you redistribute BGP into your IGP, then you run the risk of one filter list slipup causing your internal network to fail, because of routers that can't handle a full Internet-scale routing table. Unless you have the time and the talent to maintain such a setting (and only the largest companies do), don't go this way. Advertise default from A and B. And unless you'd prefer to send packets from one location to another over your wires rather than the Internet, leave it at that. If you do care, run iBGP between A and B, and for simplicity use a direct link between A and B. If A and B aren't directly linked, then you have to think through how packets headed for an Internet might transit the intermediate routers.

You generally don't want to redistribute your IGP into BGP either. The network configuration statement in BGP is much different than with other IP routing protocols. It acts much more like a permit statement in a redistribute filter list than anything else: it tells BGP which routes it's okay to advertise. With such powerful configuration statements at your fingertips, redistributing into BGP is at best redundant and could cause problems.

## Configuring BGP

Configuring BGP is going to have to wait for another (next?) month, as we've exhausted the available time and space.

## References

Howard Berkowitz has an interesting draft document on multihoming. See <http://search.ietf.org/internet-drafts/draft-berkowitz-multirgmt-01.txt> .

The ultimate reference on BGP is of course the book Internet Routing Architectures, Bassam Halabi, Cisco Press, 1997.

---

Dr. Peter J. Welcher (CCIE #1773, CCSI #94014) is a Senior Consultant with Chesapeake NetCraftsmen. NetCraftsmen is a high-end consulting firm and Cisco Premier Partner dedicated to quality consulting and knowledge transfer. NetCraftsmen has nine CCIE's, with expertise including large network high-availability routing/switching and design, VoIP, QoS, MPLS, network management, security, IP multicast, and other areas. See <http://www.netcraftsmen.net> for more information about NetCraftsmen. Pete's links start at <http://www.netcraftsmen.net/welcher> . New articles will be posted under the Articles link. Questions, suggestions for articles, etc. can be sent to [pjw@netcraftsmen.net](mailto:pjw@netcraftsmen.net) .

---