This chapter covers the following key topics:

- **Overview of routers and routing**—Provides a brief consideration of basic routing and interior gateway protocols (IGPs) as a point of contrast for the next chapter's more in-depth deliberation of exterior gateway protocols.
- **Routing protocol concepts**—This section provides an overview of the distance vector and link-state distributed routing algorithms.
- **Segregating the world into autonomous systems**—An autonomous system is a set of routers that shares the same routing policies. Various configurations for autonomous systems are possible, depending on how many exit points to outside networks are desired and whether the system should permit transit traffic.

CHAPTER 4

Interdomain Routing Basics

The Internet is a conglomeration of autonomous systems that define the administrative authority and the routing policies of different organizations. Autonomous systems are made up of routers that run Interior Gateway Protocols (IGPs) such as Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP), Open Shortest Path First (OSPF), and Intermediate System-to-Intermediate System (IS-IS) within their boundaries and interconnect via an Exterior Gateway Protocol (EGP). The current Internet de facto standard EGP is the Border Gateway Protocol Version 4 (BGP-4), defined in RFC 1771[1].

Overview of Routers and Routing

Routers are devices that direct traffic between hosts. They build routing tables that contain collected information on all the best paths to all the destinations that they know how to reach. The steps for basic routing are as follows:

Step 1 Routers run programs referred to as *routing protocols* to both transmit and receive route information to and from other routers in the network.

Step 2 Routers use this information to populate routing tables that are associated with each particular routing protocol.

Step 3 Routers scan the routing tables from the different routing protocols (if more than one routing protocol is running) and select the best path(s) to each destination.

Step 4 Routers associate with that destination the next-hop device's attached data link layer address and the local outgoing interface to be used when forwarding packets to the destination. Note that the next-hop device could be another router, or perhaps even the destination host.

Step 5 The next-hop device's forwarding information (data link layer address plus outgoing interface) is placed in the router's forwarding table.

Step 6 When a router receives a packet, the router examines the packet's header to determine the destination address.

Step 7 The router consults the forwarding table to obtain the outgoing interface and next-hop address to reach the destination.

Step 8 The router performs any additional functions required (such as IP TTL decrement or manipulating IP TOS settings) and then forwards the packet on to the appropriate device.

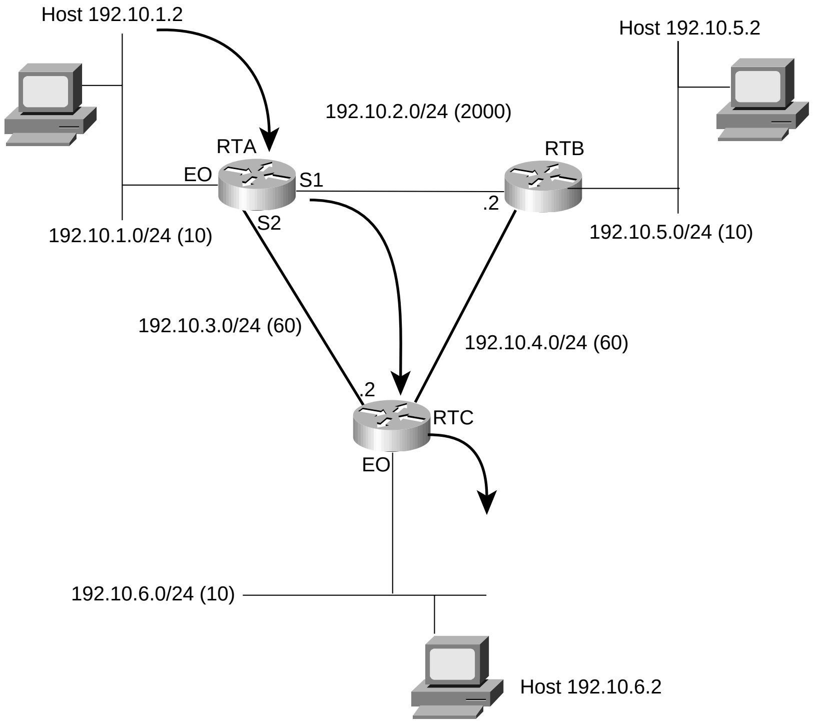
Step 9 This continues until the destination host is reached. This behavior reflects the hop-by-hop routing paradigm that's generally used in packet-switching networks.

EGPs, such as BGP, were introduced because IGPs do not scale well in networks that go beyond the enterprise level, with thousands of nodes and hundreds of thousands of routes. IGPs were never intended to be used for this purpose. This chapter touches on basic IGP functionality.

Basic Routing Example

Figure 4-1 describes three routers—RTA, RTB, and RTC—connecting three local area networks—192.10.1.0, 192.10.5.0, and 192.10.6.0—via serial links. Each serial link is represented by its own network number, which results in three additional networks, 192.10.2.0, 192.10.3.0, and 192.10.4.0. Each network has a metric associated with it, indicating the level of overhead (cost) of transmitting traffic on that particular link. The link between RTA and RTB, for example, has a cost of 2,000, much higher than the cost of 60 of the link between RTA and RTC. In practice, the link between RTA and RTB might be a 56 Kbps link with much larger delays than the T1 link between RTA and RTC and the T1 link between RTC and RTB combined.

Figure 4-1 *Basic Routing Behavior*



RTA IP Routing Table (RIP)

Destination	Next Hop	Hop Count
192.10.1.0	Connected (E0)	-
192.10.2.0	Connected (S1)	-
192.10.3.0	Connected (S2)	-
192.10.4.0	192.10.2.2 (S1) 192.10.3.2 (S2)	1 1
192.10.5.0	192.10.2.2 (S1)	1
192.10.6.0	192.10.3.2 (S2)	1

RTA IP Routing Table (OSPF)

Destination	Next Hop	Hop Count
192.10.1.0	Connected (E0)	-
192.10.2.0	Connected (S1)	-
192.10.3.0	Connected (S2)	-
192.10.4.0	192.10.3.2 (S2)	120
192.10.5.0	192.10.3.2 (S2)	130
192.10.6.0	192.10.3.2 (S2)	70

Routers RTA, RTB, and RTC would exchange network information via some IGP and build their respective IP routing tables. Figure 4-1 shows examples of RTA’s IP routing table for two different scenarios; the routers are exchanging routing information via RIP in one scenario and OSPF in another.

As an example of how traffic is passed between end stations, if host 192.10.1.2 were trying to reach host 192.10.6.2, it would use its local manually installed default route to first send the traffic to RTA. RTA would look in its IP routing table for any network that matches this destination and would find that network 192.10.6.0 is reachable via next-hop 192.10.3.2 (RTC) on serial line 2 (S2). RTC would receive the traffic and would try to look for the destination in its IP routing table (not shown). RTC would discover that the host is directly connected to its Ethernet 0 interface (E0) and would send the traffic to 192.10.6.2.

In this example, the routing is the same whether RTA is using the RIP or OSPF scenario. RIP and OSPF, however, fall into different categories of IGP protocols—distance vector protocols and link-state protocols, respectively. For a different routing example in Figure 4-1, the results might be different depending on whether you are looking at the RIP or OSPF scenario. It is useful at this point to consider characteristics of both IGP protocol categories to see how protocols generally have evolved to meet increasingly sophisticated routing demands.

Routing Protocol Concepts

Generally speaking, most routing protocols used today are based on one of two types of distributed routing algorithms: link-state or distance vector. In the next few sections, we'll discuss the different properties of distance vector and link-state routing algorithms.

Distance Vector Routing Protocols

Distance vector protocols are sometimes referred to as Bellman-Ford protocols, named after the person who invented the algorithm used for calculating the shortest paths[2] and for the people who first described a distributed use of the algorithm[3]. The term *distance vector* is derived from the fact that the protocol includes a vector (list) of distances (hop counts or other metrics) associated with each destination prefix routing message.

Distance vector routing protocols, such as Routing Information Protocol (RIP), utilize a distributed computation approach to calculating the route to each destination prefix. In other words, distance vector protocols require that each node separately calculate the best path (output link) to each destination prefix.

After selecting the best path, a router then sends distance vectors to its neighbors, notifying them of the reachability of each destination prefix and of the corresponding metrics associated with the path it has selected to reach the prefix. In parallel, its neighbors also calculate the best path to each available destination and then notify their neighbors of the available path (and associated metrics) they've selected to reach the destination. Upon the receipt of messages from neighbors detailing the destination and associated metrics that the neighbor has selected, the router might determine that a better path exists via an alternative neighbor. The router will again notify its neighbors of its selected paths (and associated metrics) to reach each destination. This cycle continues until all the routers have converged upon a common understanding of the best paths to reach each destination prefix.

Initial specifications of distance vector routing protocols such as RIP Version 1 (RIP-1) had several drawbacks. For example, hop count was the only metric RIP-1 used to select a path. This imposed several limitations. Consider, for example, the RTA routing tables shown in Figure 4-1. One table represents routing information considered when using RIP, and the other when using OSPF. (OSPF is a link-state routing protocol that will be discussed in more detail in the following sections.)

When using RIP-1, RTA would select the direct link between RTA and RTB to reach network 192.10.5.0. RTA prefers this link because the direct path requires just one hop via the RTB path versus two hops via the RTC-RTB path. However, RTA has no knowledge that the RTA-RTB link is actually a very low-capacity, high-latency connection and that using the RTC-RTB path would provide a better level of service.

On the other hand, when using OSPF and metrics other than hop count alone for path selection, RTA will realize that the path to RTB via RTC (cost: 60 + 60 = 120; 2 hops) is actually more optimal than the direct path (cost: 2000; 1 hop).

Another issue with hop counts is the count to infinity restriction. Traditional distance vector protocols (for example, RIP-1) have a finite limit of hops, often 15, after which a route is considered unreachable. This would restrict the propagation of routing updates and would cause problems for large networks (those with more than 15 nodes in a given path). The reliance on hop counts is one deficiency of distance vector protocols, although newer distance vector protocols (that is, RIP-2 and EIGRP) are not constrained as such.

Another deficiency is the way that the routing information is exchanged. Traditional distance vector protocols work on the concept that routers exchange all the IP network numbers they can reach via periodic exchange of distance vector broadcasts—broadcasts that are sent when a “refresh timer” associated with the message exchange expires. Because of this, if the refresh timer expires and a fresh set of routing information is broadcast to your neighbors, the timer is reset, and no new information is sent until the timer expires again. Now, consider what would happen if a link or path became unavailable just after a refresh occurred. Propagation of the path failure would be suppressed until the refresh timer expired, thereby slowing convergence considerably.

Fortunately, newer distance vector protocols, such as EIGRP and RIP-2, introduce the concept of *triggered updates*. Triggered updates propagate failures as soon as they occur, speeding convergence considerably.

As you might have realized, in large networks, or even small networks with a large number of destination prefixes, periodic exchange of the routing table between neighbors might become very large and very difficult to maintain, contributing to slower convergence. Also, the amount of CPU and link overhead consumed by periodic advertisement of routing information can become quite large. Another property that newer distance vector protocols have adopted is to introduce reliability to the transmission of the distance vectors between neighbors, eliminating the need to periodically readvertise the entire routing table.

Convergence refers to the point in time at which the entire network becomes updated to the fact that a particular route has appeared, disappeared, or changed. Traditional distance vector protocols worked on the basis of periodic updates and hold-down timers: If a route is not received in a certain amount of time, the route goes into a hold-down state and gets aged out of the routing table. The hold-down and aging process translates into minutes in convergence time before the whole network detects that a route has disappeared. The delay between a route's becoming unavailable and its aging out of the routing tables can result in temporary forwarding loops or black holes.

Another issue in some distance vector protocols (for example, RIP) is that when an active route disappears, but the same route reappears with a higher metric (presumably emanating from another router, indicating a possible "good" alternative path), the route is still put into a hold-down state. Thus, the amount of time for the entire network to converge is still increased.

Another major drawback of first-generation distance vector protocols is their classful nature and their lack of support for VLSM or CIDR. These distance vector protocols do not exchange mask information in their routing updates and are therefore incapable of supporting these technologies. In RIP-1, a router that receives a routing update on a certain interface will apply to this update its locally defined subnet mask. IGRP does the same thing as RIP-1 but falls back to Class A, B, and C network masks if a portion of the transmitted network address does not match the local network address. This would lead to confusion (in case the interface belongs to a network that is variably subnetted) and a misinterpretation of the received routing update. Newer distance vector protocols, such as RIP Version 2 (RIP-2) and EIGRP, overcome the aforementioned shortcomings.

Several modifications have been made that alleviate deficiencies associated with traditional distance vector routing protocol behaviors. For example, RIP-2 and EIGRP support VLSM and CIDR. Also, IGRP and EIGRP have the capability to factor in composite metrics used to represent link characteristics along a path (such as bandwidth, utilization, delay, MTU, and so forth), which allows them to calculate more optimal paths than using a hop count alone.

The simplicity and maturity of distance vector protocols has led to their popularity. The primary drawback of traditional implementation of distance vector protocols is slow convergence, a property that can be a catalyst for introducing forwarding loops and/or black-holing traffic during topological changes. However, newer distance vector protocols—most notably, EIGRP—actually converge quite well.

This section wouldn't be complete without mentioning that BGP falls into the distance vector category. In addition to the standard distance vector properties, BGP employs an additional mechanism referred to as the *path vector*, used to avoid the count to infinity problem previously discussed. Essentially, the path vector contains a list of routing domains

(AS numbers) through which the route has traversed. If a domain receives a route for which its domain identifier is already listed in the path, the route is ignored. This path information provides a mechanism that allows routing loops to be pruned. It can also be used to apply domain-based policies. This path attribute, and many other path attributes, are discussed in detail in the following chapters.

Link-State Routing Protocols

Link-state routing protocols, such as Open Shortest Path First (OSPF)[4] and Intermediate System-to-Intermediate System (IS-IS)[5], utilize a replicated distributed database model and are considered to be more-complex routing protocols. Link-state protocols work on the basis that routers exchange information elements, called *link states*, which carry information about links and nodes in the routing domain. This means that routers running link-state protocols do not exchange routing tables as distance vector protocols do. Rather, they exchange information about adjacent neighbors and networks and include metric information associated with the connection.

One way to view link-state routing protocols is as a jigsaw puzzle. Each router in the network generates a piece of the puzzle (link state) that describes itself and where it connects to adjacent puzzle pieces. It also provides a list of the metrics corresponding to the connection with each piece of the puzzle. The local router's piece of the puzzle is then reliably distributed throughout the network, router by router, via a flooding mechanism, until all nodes in the domain have received a copy of the puzzle piece. When distribution is complete, every router in the network has a copy of every piece of the puzzle and stores the puzzle pieces in what's referred to as a *link-state database*. Each router then autonomously constructs the entire puzzle, the result of which is an identical copy of the entire puzzle on each router in the network.

Then, by applying the SPF (shortest path first) algorithm (most commonly, the Dijkstra Algorithm) to the puzzle, each router calculates a tree of shortest paths to each destination, placing itself at the root.

Following are some of the benefits that link-state protocols provide:

- **No hop count**—There are no limits on the number of hops a route can take. Link-state protocols work on the basis of link metrics rather than hop counts.

  As an example of a link-state protocol's reliance on metrics rather than hop count, turn again to the RTA routing tables shown in Figure 4-1. In the OSPF case, RTA has picked the optimal path to reach RTB by factoring in the cost of the links. Its routing table lists the next hop of 192.10.3.2 (RTC) to reach 192.10.5.0 (RTB). This is in contrast to the RIP scenario, which resulted in a suboptimal path.

- **Bandwidth representation**—Link bandwidth and delays may be (manually or dynamically) factored in when calculating the shortest path to a certain destination. This leads to better load balancing based on actual link cost rather than hop count.
- **Better convergence**—Link and node changes are immediately flooded into the domain via link-state updates. All routers in the domain will instantly update their routing tables (some similar to triggered updates).
- **Support for VLSM and CIDR**—Link-state protocols exchange mask information as part of the information elements that are flooded into the domain. As a result, networks with variable-length subnet masks can be easily identified.
- **Better hierarchy**—Whereas distance vector networks are flat networks, link-state protocols provide mechanisms to divide the domain into different levels or areas. This hierarchical approach better scopes network instabilities within areas.

Although link-state algorithms have traditionally provided better routing scalability, which allows them to be used in bigger and more complex topologies, they still should be restricted to interior routing. Link-state protocols by themselves cannot provide a global connectivity solution required for Internet interdomain routing. In very large networks and in case of route oscillation caused by link instabilities, link-state retransmission and recomputation will become too large for any single router to handle.

Although a more detailed discussion of IGPs is beyond the scope of this book, two excellent references that discuss the different link-state and distance vector routing protocols are *Interconnections, Second Edition: Bridges, Routers, Switches and Internetworking Protocols*[6] by Radia Perlman and *OSPF: Anatomy of an Internet Routing Protocol*[7] by John T. Moy.

Most large service providers today use link-state routing protocols for intra-AS routing, primarily because of its fast convergence capabilities. The two most common protocols deployed in this space are OSPF and IS-IS.

Many older service providers have selected IS-IS as their IGP, and some newer providers select OSPF or IS-IS. Initially, it might seem that older networks use IS-IS rather than OSPF because the U.S. Government required support of ISO CLNP by networks in order for the networks to be awarded federal contracts. (Note that IS-IS is capable of carrying both CLNP and IP Network layer information, while OSPF is capable of carrying only IP information.) However, Internet folklore suggests that the driving factor was that IS-IS implementations were much more stable than OSPF implementations when early providers were selecting which routing protocol to use. This stability obviously had a significant impact on which IGP service providers selected.

Today, both IS-IS and OSPF are widely deployed in ISP networks. The maturity and stability of IS-IS has resulted in its remaining deployed in large networks, as well as its being the IGP of choice for some more recently deployed networks.

Segregating the World into Autonomous Systems

Exterior routing protocols were created to control the expansion of routing tables and to provide a more structured view of the Internet by segregating routing domains into separate administrations, called *autonomous systems (ASs)*, which each have their own independent routing policies and unique IGPs.

During the early days of the Internet, an exterior gateway protocol called EGP[8] (not to be confused with Exterior Gateway Protocols in general) was used. The NSFNET used EGP to exchange reachability information between the backbone and the regional networks. Although the use of EGP was widely deployed, its topology restrictions and inefficiency in dealing with routing loops and setting routing policies created a need for a new and more robust protocol. Currently, BGP-4 is the de facto standard for interdomain routing in the Internet.

NOTE Note that the primary difference between intra-AS and inter-AS routing is that intra-AS routing is usually optimized in accordance with the required technical demands, while inter-AS usually reflects political and business relationships between the networks and companies involved.

Static Routing, Default Routing, and Dynamic Routing

Before introducing and looking at the basic ways in which autonomous systems can be connected to ISPs, we need to establish some basic terminology and concepts of routing:

- *Static routing* refers to routes to destinations being listed manually, or statically, as the name implies, in the router. Network reachability in this case is not dependent on the existence and state of the network itself. Whether a destination is active or not, the static routes remain in the routing table, and traffic is still sent toward the specified destination.
- *Default routing* refers to a “last resort” outlet. Traffic to destinations that is unknown to the router is sent to that default outlet. Default routing is the easiest form of routing for a domain connected to a single exit point.
- *Dynamic routing* refers to routes being learned via an interior or exterior routing protocol. Network reachability is dependent on the existence and state of the network. If a destination is down, the route disappears from the routing table, and traffic is not sent toward that destination.

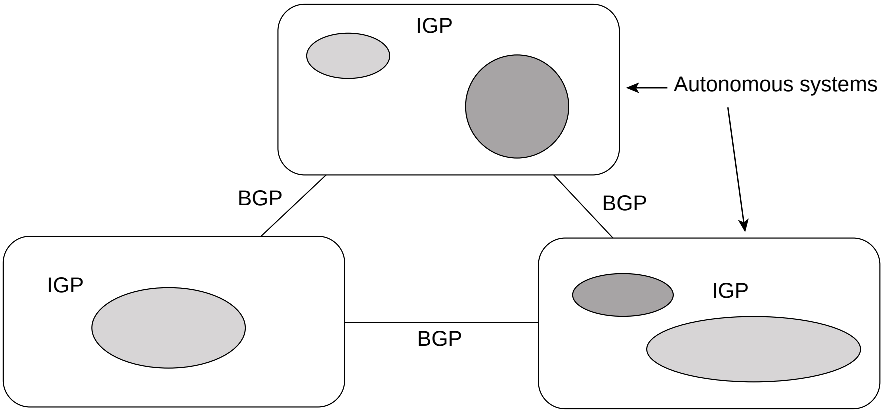
These three routing approaches are possibilities for all the AS configurations considered in forthcoming sections, but usually there is an optimal approach. Thus, in illustrating different autonomous systems, this chapter considers whether static, dynamic, default, or some combination of these is optimal. This chapter also considers whether interior or exterior routing protocols are appropriate. However, a more detailed exploration of routing choices for different AS topologies will not be discussed until Chapter 6, “Tuning BGP Capabilities.”

Always remember that static and default routing are not your enemy. The most stable (but sometimes less flexible) configurations are based on static routing. Many people feel that they are not technologically up to date just because they are not running dynamic routing. Trying to force dynamic routing on situations that do not require it is a waste of bandwidth, effort, and money. Recall the KISS principle introduced in the preceding chapter!

Autonomous Systems

An *autonomous system* (AS) is a set of routers that has a single routing policy, that run under a single technical administration, and that commonly utilizes a single IGP (the AS could also be a collection of IGPs working together to provide interior routing). To the outside world, the entire AS is viewed as a single entity. Each AS has an identifying number, which is assigned to it by an Internet Registry, or a service provider in the instance of private ASs. Routing information between ASs is exchanged via an exterior gateway protocol such as BGP-4, as illustrated in Figure 4-2.

Figure 4-2 *Routing Information Exchange Between Autonomous Systems*



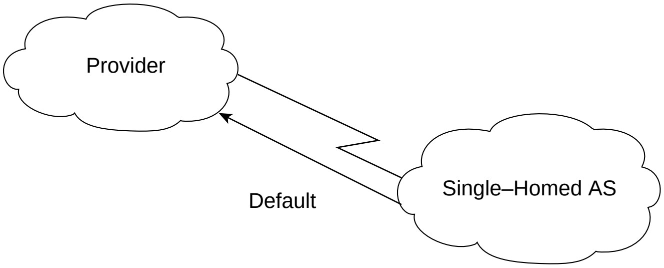
What we have gained by segregating the world into administrations is the capability to have one large network (in the sense that the Internet could have been one huge OSPF or IS-IS network) divided into smaller and more manageable networks. These networks, represented as ASs, can now implement their own set of rules and policies that will uniquely distinguish their networks and associated service offerings from other networks. Each AS can now run its own set of IGPs, independent of IGPs in other ASs.

The next few sections discuss potential network configurations with stub (single-homed) networks, multihomed nontransit networks, and multihomed transit networks.

Stub AS

An AS is considered stub when it reaches networks outside its domain via a single exit point. These ASs are also referred to as *single-homed* with respect to other providers. Figure 4-3 illustrates a single-homed or stub AS.

Figure 4-3 *Single-Homed (Stub) AS*



A single-homed AS does not really have to learn Internet routes from its provider. Because there is a single way out, all traffic can default to the provider. When using this configuration, the provider can use different methods to advertise the customer's routes to other networks.

One possibility is for the provider to list the customer's subnets as static entries in its router. The provider would then advertise these static entries toward the Internet via BGP. This method would scale very well if the customer's routes can be represented by a small set of aggregate routes. When the customer has too many noncontiguous subnets, listing all these subnets via static routes becomes inefficient.

Alternatively, the provider can employ IGPs for advertising the customer's networks. An IGP can be used between the customer and provider for the customer to advertise its routes. This has all the benefits of dynamic routing where network information and changes are dynamically sent to the provider. This is very uncommon, however, primarily because it doesn't scale well because customer link instability can result in IGP instabilities.

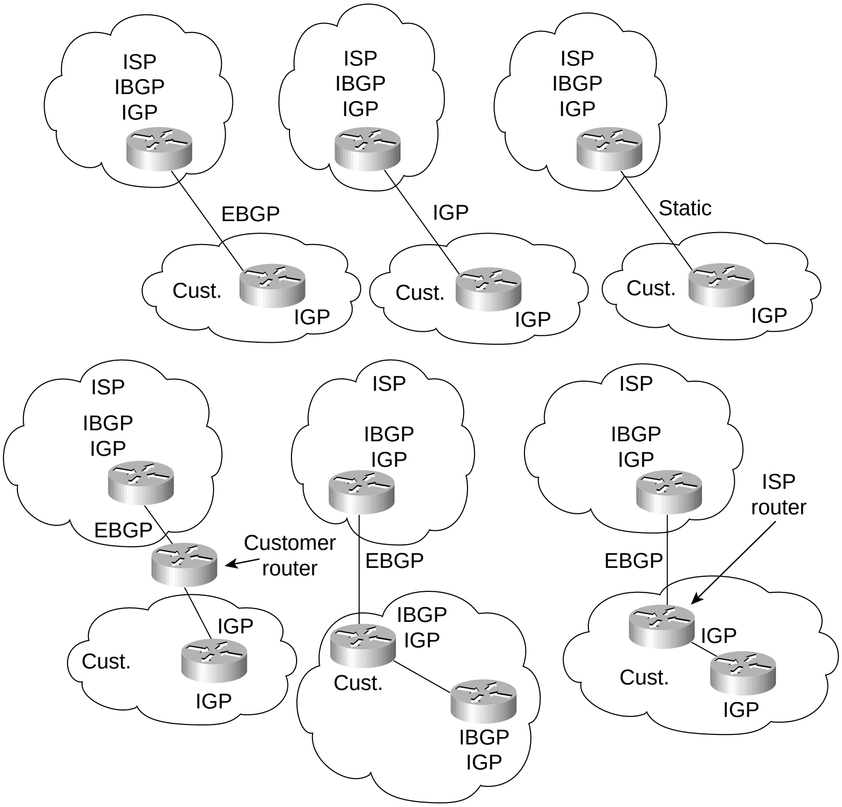
The third method by which the ISP can learn and advertise the customer's routes is to use BGP between the customer and the provider. In the stub AS situation, it is hard to get a registered AS number from an IRR because the customer's routing policies are an extension of the policies of a single provider.

NOTE RFC 1930[9] provides a set of guidelines for the creation, selection, and registration of autonomous system numbers.

Instead, the provider can give the customer an AS number from the private pool of ASs (65412-65535), assuming that the provider's routing policies have provisioned support for using private AS space with customers, as described in RFC 2270[10].

Quite a few combinations of protocols can be used between the ISP and the customer. Figure 4-4 illustrates some of the possible configurations, using just stub ASs as an example. (The meaning of EBGP and IBGP will be discussed in upcoming sections.) Providers might extend customer routers to their POPs, or providers might extend their routers to the customer's network. Note that not every situation requires that a customer run BGP with its provider, as mentioned earlier.

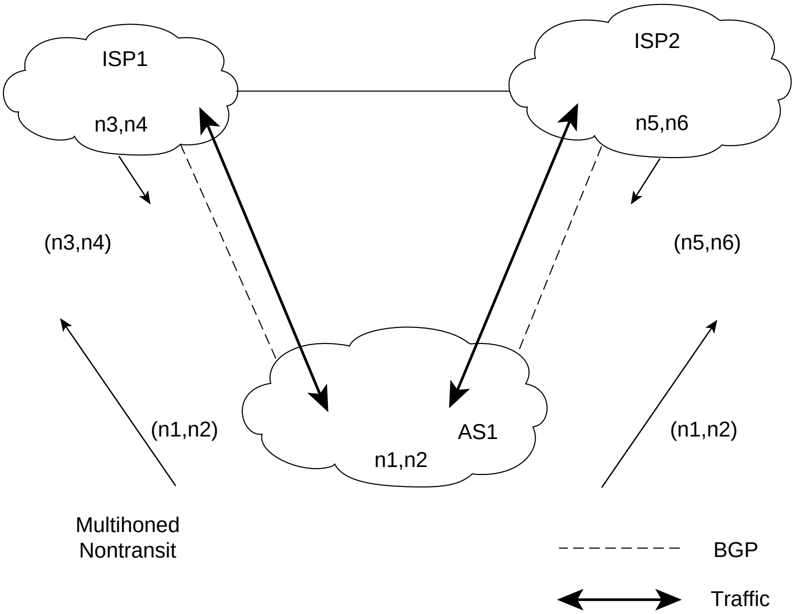
Figure 4-4 *Stub ASs: Sample Protocol Implementation Variations*



Multihomed Nontransit AS

An AS is multihomed if it has more than one exit point to the outside world. An AS can be multihomed to a single provider or multiple providers. A nontransit AS does not allow transit traffic to go through it. *Transit traffic* is any traffic that has a source and destination outside the AS. Figure 4-5 illustrates an AS (AS1) that is nontransit and multihomed to two providers, ISP1 and ISP2.

Figure 4-5 *Multihomed Nontransit AS Example*



A nontransit AS would only advertise its own routes and would not propagate routes that it learned from other ASs. This ensures that traffic for any destination that does not belong to the AS would not be directed to the AS. In Figure 4-5, AS1 learns about routes n3 and n4 via ISP1 and routes n5 and n6 via ISP2. AS1 advertises only its local routes (n1,n2). It does not pass to ISP2 the routes it learned from ISP1 or to ISP1 the routes it learned from ISP2. This way, AS1 does not open itself to outside traffic, such as ISP1 trying to reach n5 or n6 and ISP2 trying to reach n3 and n4 via AS1. Of course, ISP1 or ISP2 can force its traffic to be directed to AS1 via default or static routing. As a precaution against this, AS1 could filter any traffic coming toward it with a destination not belonging to AS1.

Multihomed nontransit ASs do not really need to run BGP with their providers, although it is recommended and most of the time is required by the provider. As you will see later in this book, running BGP-4 with the providers has many advantages as far as controlling route propagation and filtering.

Multihomed Transit AS

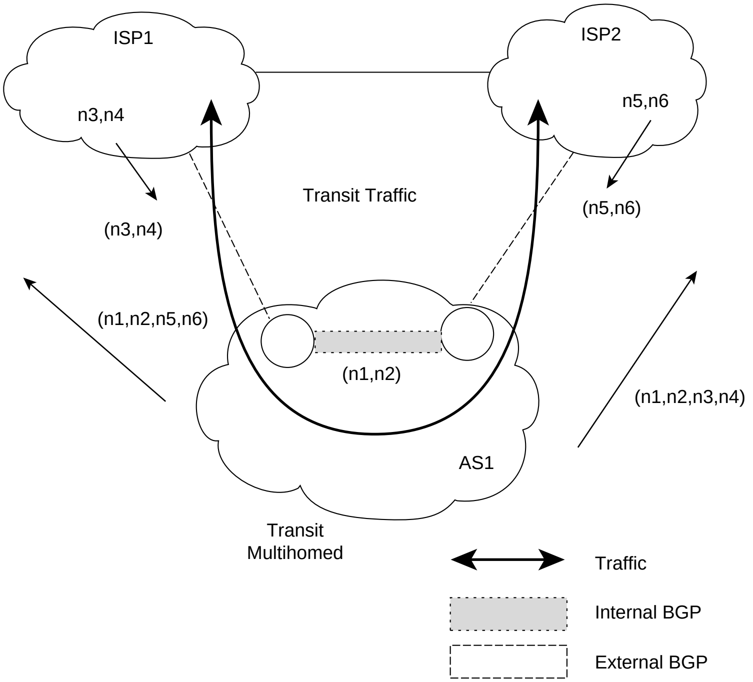
A *multihomed transit AS* has more than one connection to the outside world and can still be used for transit traffic by other ASs (see Figure 4-6). Transit traffic (relative to the multihomed AS) is any traffic that has an origin and destination that does not belong to the local AS.

Although BGP-4 is an exterior gateway protocol, it can still be used inside an AS as a pipe to exchange BGP updates. BGP connections between routers inside an autonomous system are referred to as *Internal BGP (IBGP)*, whereas BGP connections between routers in separate autonomous systems are referred to as *External BGP (EBGP)*. Routers that are running IBGP are called *transit routers* when they carry the transit traffic going through the AS.

A transit AS would advertise to one AS routes that it learned from another AS. This way, the transit AS would open itself to traffic that does not belong to it. Multihomed transit ASs are advised to use BGP-4 for their connections to other ASs and to shield their internal nontransit routers from Internet routes. Not all routers inside a domain need to run BGP; internal nontransit routers could run default routing to the BGP routers, which alleviates the number of routes the internal nontransit routers must carry. In most large service provider networks, however, all routers usually carry a full set of BGP routes internally.

Figure 4-6 illustrates a multihomed transit autonomous system, AS1, connected to two different providers, ISP1 and ISP2. AS1 learns routes n3, n4, n5, and n6 from both ISP1 and ISP2 and in turn advertises all that it learns, including its local routes, to ISP1 and ISP2. In this case, ISP1 could use AS1 as a transit AS to reach networks n5 and n6, and ISP2 could use AS1 to reach networks n3 and n4.

Figure 4-6 *Multihomed Transit AS Using BGP Internally and Externally*



Looking Ahead

The Border Gateway Protocol has defined the basis of routing architectures in the Internet. The segregation of networks into autonomous systems has logically defined the administrative and political borders between organizations. Interior Gateway Protocols can now run independently of each other, but networks can still interconnect via BGP to provide global routing.

Chapter 5, "Border Gateway Protocol Version 4," is an overview of how BGP-4 operates, including detailed discussions of its message header formats.

Frequently Asked Questions

Q — *What is the difference between a domain and an autonomous system?*

A — Both terms are used to indicate a collection of routers. The domain notation is usually used to indicate a collection of routers running the same routing protocol, such as a RIP domain or an OSPF domain. The AS represents one or more domains under a single administration that have a unified routing policy with other ASs.

Q — *My company is connected to an ISP via RIP. Should I use BGP instead?*

A — If you are thinking of connecting to multiple providers in the near future, you should start discussing the option of using BGP with your provider. If your traffic needs do not require multiple provider connectivity, you should be okay with what you have.

Q — *I have a single IGP connection to a provider. I am thinking of connecting to the same provider in a different location. Can I connect via an IGP, or should I use BGP?*

A — This depends on the provider. Some providers will let you connect via IGP in multiple locations; others prefer that you use BGP. Practically speaking, when you use BGP, you will be in better control of your traffic, as you will see in the following chapters.

Q — *I thought that BGP is to be used between ASs. I am a bit confused about using BGP inside the AS.*

A — Think of BGP inside the AS (IBGP) as a tunnel through which routing information flows. If your AS is a transit AS, IBGP will shield all your internal nontransit routers from the potentially overwhelming number of external routing updates. On the other hand, even if you are not a transit AS, you will realize as this book progresses that IBGP will give you better control in choosing exit and entrance points for your traffic.

Q — *You talk about BGP-4, but is anybody still using BGP-1, -2, or -3? What about EGP?*

A — BGP-4 is the de facto interdomain routing protocol used on the Internet. EGP and BGP-1, 2, and 3 are obsolete. BGP-4's support of CIDR, incremental updates, and better filtering and policy-setting capabilities have prompted everybody to shift gears into using this new protocol.

Q — *I'm planning to install a second connection to my current Internet service provider. Should I get an AS number from my RIR?*

A — Getting an AS number is indeed an option, although you might first see if your provider has provisions in place to support the use of private ASs for customers multihomed to a single provider. In addition, you should check with your RIR to ensure that it will allocate AS numbers to networks connected to only a single provider.

References

[1] RFC 1771, “A Border Gateway Protocol 4 (BGP-4),” www.isi.edu/in-notes/rfc1771.txt

[2] Bellman, R. *Dynamic Programming* (Princeton University Press, 1957)

[3] Ford, L. R., Jr. and D. R. Fulkerson. *Flows in Networks* (Princeton University Press, 1962)

[4] RFC 1583, “OSPF Version 2,” www.isi.edu/in-notes/rfc1583.txt

[5] ISO 10589, “Intermediate System to Intermediate System”; RFC 1195, “Use of OSI IS-IS for Routing in TCP/IP and Dual Environments,” www.isi.edu/in-notes/rfc1195.txt

[6] Perlman, Radia. *Interconnections, Second Edition: Bridges, Routers, Switches, and Internetworking Protocols* (Boston, Mass.: Addison-Wesley Longman, Inc., 1999)

[7] Moy, John. *OSPF: Anatomy of an Internet Routing Protocol* (Boston, Mass.: Addison-Wesley Longman, Inc., 1998)

[8] RFC 904, “Exterior Gateway Protocol Formal Specification,” www.isi.edu/in-notes/rfc904.txt

[9] RFC 1930, “Guidelines for creation, selection, and registration of an Autonomous System (AS),” www.isi.edu/in-notes/rfc1930.txt

[10] RFC 2270, “Using a Dedicated AS for Sites Homed to a Single Provider,” www.isi.edu/in-notes/rfc2270.txt