Jonathan Frederickson
BGP Design

Background

Every device on the Internet must have an IP address to communicate on the network. IP address uniquely identifies a network interface within a computing device, which in turn also identifies that computing device. Within a local network, IP addresses are distributed using some bootstrap protocol such as Dynamic Host Configuration Protocol (DHCP) or Bootstrap Protocol (BOOTP), or configured manually. In the network core, the backbone of the Internet, device addressing is conducted using a more sophisticated protocol such as Border Gateway Protocol (BGP).

Generally, the Internet backbone is divided into a set of autonomous systems (AS) or domains, each managed by a separate authority. ASs use two variants of BGP: one for traffic within the same AS (internal BGP, or iBGP) and one for traffic that must traverse routers outside the confines of the AS (external BGP, or eBGP). BGP is the only inter-domain routing protocol currently in use on the Internet, and it is responsible for maintaining and updating internal inter-domain routing tables [Forouzan].

Autonomous systems are identified by an AS number (ASN), a 32-bit number assigned to the authority responsible for operating routers for that AS. ASNs are assigned in blocks by the Internet Assigned Numbers Authority (IANA) to regional Internet registries (RIRs), who then assign these to individual ASs. 33523, for example, is the ASN assigned to Rowan University by the American Registry for Internet Numbers (ARIN).

It is important to note that an ASN does not identify a single router, but rather the entire collection of routers controlled by that AS. Routers within an AS will typically run a combination of iBGP and an interior gateway protocol (IGP). The interior gateway protocol, commonly either Routing Information Protocol (RIP) or Open Shortest Path First (OSPF) is responsible for exchanging routing information between routers in the same AS. For example, if traffic is sent to a router that is destined for a host on a different network but within the same AS, the router will make this decision based on its IGP routing tables. The details of the operation of IGPs is beyond the scope of this document.

When a router receives traffic, it must consult its routing table to determine which of its network interfaces to forward the traffic through to bring the traffic closer to its destination. As a result, to be able to route traffic destined to any host on the Internet, a router must know where to send the traffic to reach any network on the Internet. BGP is the protocol used to distribute this information among routers on the Internet.

Initially, a router does not have any information about other routers on the network. To begin with, the router opens connections to its direct peers (which are typically configured manually) on the well-known TCP port 179. Having done so, it will download their entire routing tables as well as send an UPDATE message containing its ASN. When its peers receive the UPDATE, they determine if it is a “better” route to that destination than any existing route, and that they do not already appear in the path (to avoid routing loops). If so, they record it and propagate it to their own peers, adding their own ASNs. This process continues until all routers in the network have knowledge of the new router, and the “best” path to reach it. The ranking of routes as better or worse in this context is done according to the router’s configurable decision policy. In the end, each router will have full paths that they can traverse to reach any destination, such as:

AS1001-AS2345-AS2345-AS12437-AS8888

The process above is known as a path-vector routing algorithm. If the Internet as a whole is represented as an undirected graph, this algorithm will effectively select a spanning tree for that graph based on its own configurable routing policy. This tree can be used to determine the preferred route for traffic to any destination on the Internet.

These routes are, of course, not created instantaneously across the entire network. The process of updating neighbor routers, and as a result the network has a whole, takes some finite amount of time. Once all routers agree on the state of the network, the network is said to have converged, but in the meantime there can be some temporary instability in the network. The route to a destination can change while packets are in transit, for example, which can lead to lost or delayed packets.

In many cases, convergence time is not a major issue, as it is only relevant when a node in the network is brought up or down. However, in the event of an unstable link between routers, a router may appear to rapidly join and disconnect from the network. Each time it joins, it will advertise itself as reachable; conversely, each time it disconnects, its peer will advertise that it is no longer reachable. If this happens rapidly enough, absent any protective measures against it, the network would never converge while this was occurring. This issue is commonly known as route flapping.

Additionally, all of the activities previously mentioned do not mandate any form of authentication. The ASN used by a particular router is configured manually by its operator and can be set to any desired value. Most larger providers perform filtering of their customers' BGP sessions to ensure that they only advertise routes that belong to them, however this is enforced at the ISP level and not within the network core. If an ISP does not filter their customers' activity, it is possible for a malicious AS to inject routes and redirect traffic from anywhere on the Internet.

Experiment Design

Riverbed Modeler Academic Edition includes sample BGP networks that I will modify to meet my needs. By advertising routes and disconnecting rapidly in the simulated BGP network, I will attempt to induce route flapping in the network. Riverbed Modeler allows events to be started at specific times during the simulation, which will facilitate demonstration of this issue.

As BGP does not have authentication for route updates, it is possible for a malicious peer to advertise address space that they do not own, causing routing issues for the legitimate user of that space. I will be attempting to model this in Riverbed Modeler as well. However, due to the limitations of the Academic Edition of Riverbed Modeler, it may not be possible to do this, as doing so requires one peer to be configured incorrectly.

Riverbed Modeler Academic Edition supports simulation of the BGP protocol, however the simulation itself is opaque; in full editions of Modeler, it is possible to view and edit the protocol state machine, but in Academic Edition it is not. However, it does allow limited control of the simulated routers through a Cisco-like syntax. This will mostly restrict my simulation of events in the BGP network to those that were preconfigured in Riverbed Modeler, though I will have limited control of the simulation during runtime.

References

[1] Alexander Probst, “Simulating The Bgp With Opnet Guru 10.5”, http://faculty.kfupm.edu.sa/coe/ashraf/RichFilesTeaching/COE081_540/BPG_OPNET/Studienarbeit.pdf, last visited on 3/13/2015

[2] OPNET Modeler/Release 17.5 User Manual - BGP

[3] Y. Rekhter, T. Li, S. Hares, “A Border Gateway Protocol 4 (BGP-4)”, http://tools.ietf.org/html/rfc4271, last visited on 3/19/2015