

BGP: Testing the Protocol of the Internet

Introduction

The Internet is everywhere. A farmer in rural Nebraska can check the local weather, inspect his New York stock portfolio, email his son overseas, and book an African safari all in a matter of minutes. The underlying technology that successfully connects the farmer to these destinations is a routing protocol, specifically, Border Gateway Protocol, or BGP. BGP is the routing protocol of the Internet, now in its 4th version; thus it is commonly called BGP-4.

BGP-4 is a very powerful tool. It facilitates connectivity throughout the labyrinth of the Internet, providing optimized communications by selecting the best path to a destination, and responds and adapts to frequent network changes. It accomplishes all of these functions seamlessly without the end user (our farmer, in this case) ever knowing that it exists.

Technology Overview

Background

The early Internet was not designed to support the massive growth experienced over the past decade. The first routing protocols (RIP and EGP) were certainly adequate in their time. However in the 1990s, as the Internet expanded at an exponential rate, it became apparent that a much more scalable solution was necessary. Hence, BGP was developed.

BGP was created in 1989, and is codified in RFC 1105, which is now obsolete. This RFC introduces the concepts of communication between Autonomous Systems (ASs), BGP state machines, and incremental updates. The protocol has been updated several times since then. Features have been added to improve performance or accommodate new technologies. The latest iteration, BGP-4, is specified by RFC 1771 (which supercedes several earlier RFCs) and was adopted in 1995. RFC 1771 has also been amended several times in order to introduce minor changes. The latest version is actually RFC 1771, revision 17.

Objectives

BGP was designed to resolve several challenges associated with the Internet. As the technology is discussed, it may be useful to refer back to these three primary objectives:

- **Scalability** The Internet was large in 1990, huge in 1995, and now is truly massive! And it is continuing to grow. A future-proof protocol had to be designed that could accommodate this extraordinary growth.

- **Responsiveness** The Internet is not a static entity. Instead, it is constantly changing. Circuits come and go, new ISPs are frequent occurrences, and new destinations are added daily, if not hourly. The routing protocol of the Internet must be able to rapidly incorporate these changes into its overall routing database.
- **Control** There are hundreds, if not thousands, of different Internet paths from San Jose to Boston. Selecting the optimal path requires a high level of policy control of the network. BGP offers network administrators many tools for directing traffic and controlling their portion of the Internet.

Basic Concepts

BGP is a distance vector routing protocol based upon a similar concept as its predecessors RIP and EGP – that is reachability is calculated based upon hop counts. A path with the smallest hop count is assumed to be the optimal route to a destination. BGP greatly enhances this fundamental concept in order to satisfy all of the objectives previously outlined.

Scalability

RIP calculates the best path to a remote destination based upon individual router hops. BGP calculates the path to a remote destination based upon Autonomous System (AS) hops. An autonomous system is an administrative domain (i.e. an ISP) that is comprised of a group of routers. This results in the aggregation of routers, simplifying BGP's route calculation process. The use of autonomous systems is the first step toward improving the scalability of the routing protocol.

BGP is also designed to minimize the network overhead that is specifically related to routing protocol (also known as “control plane”) functions. RIP, for example, is notorious for generating lots of overhead and control traffic. BGP greatly reduces this traffic, thereby improving the scalability of the protocol in several ways. The most significant of which is the use of “triggered updates.” This means that two adjacent BGP routers will only exchange their entire routing tables once – as they initiate a communications session. After a session is established, the two routers will exchange brief “hello” messages periodically to verify the integrity of the connection. They will only exchange routing data, or Network Layer Reachability Information (NLR), when a topological change takes place.

BGP-4 achieves further scalability by supporting network aggregation and summarization – this is a feature that was not included in the early versions of BGP. Advertisements for networks with contiguous IP addresses can be combined into single summary announcements. This reduces the overhead associated with network advertisements and, more significantly, also helps reduce the size of the forwarding databases contained within each router. The process of network summarization is based upon the principles of Classless Inter-Domain Routing (CIDR), which is a method for identifying IP networks based upon the desired length of the network prefix, rather than their traditional class-based (Class A, B or C) orientation. For example, take the following four contiguous network addresses:

- 200.200.4.0/24
- 200.200.5.0/24
- 200.200.6.0/24
- 200.200.7.0/24

Rather than having four separate routing table entries for these networks, a router can aggregate the addresses with a single entry of 200.200.4.0/22 simply by indicating a more robust subnet mask. If enough addresses are contiguous, the router may even be able to enter a summary route of 200.200.0.0/16 in its routing and forwarding table.

Responsiveness

The Internet is a constantly changing organic entity. By design, BGP is much more responsive to these network changes than some of its predecessors such as RIP. When a BGP router discovers a topology change either by a status change in a locally attached network or via an update message from another router, it immediately floods this information to all other adjacent routers. These routers in turn update their local routing and forwarding tables with the new information, and then forward the change to their peers. As a result of this procedure, topology changes are propagated very rapidly throughout the Internet.

BGP also has a couple of safeguards associated with the update procedure. First of all, BGP uses the Transmission Control Protocol (TCP), a reliable transport-layer protocol, for all of its update messages. TCP requires that the recipient acknowledge all messages. If an update message is not acknowledged, it will be retransmitted. Thus, a degree of reliability is added to the update procedure. Secondly, many BGP implementations provide “hold down” timers. These are designed to dampen the effect of frequently flapping routes. In essence, if a route “flaps” – that is, it comes and goes, perhaps due to a loose connection – the repeating changes will be propagated throughout the Internet, forcing all of the routers to constantly recalculate their routing and forwarding tables. A hold down timer will prevent the generation of multiple advertisements for the same route.

In sum, BGP is designed to quickly react to the ever-changing Internet. BGP’s responsiveness, coupled with a pair of protection mechanisms, helps ensure up to date network connectivity throughout the Internet.

Control

The third objective of an Internet routing protocol is control, or policy-based routing. The essence of this concept is to allow service providers to direct traffic based upon their own unique policies and service requirements. For example, if there are multiple routes to a given destination (and this is usually the case), the service provider may want to direct traffic to the path with the highest bandwidth. Alternatively, the cheapest path may be preferable. Or perhaps the service provider will want to optimize their network traffic based upon time-of-day routing decisions. BGP provides a set of tools to exercise extremely granular levels of control.

Network routing policies are administrative decisions made by network managers for directing network traffic. Policies are implemented by specifying values for variables, also called “BGP Path Attributes,” within a routing update. These policies can be as simple as “Accept” or “Deny” access to a particular network. Or they can be as complex as “forward all trans-Atlantic traffic via AT&T’s network trunk, except for the U.K.’s traffic which will be forwarded to British Telecom, and then use Sprint as a backup link for both streams.”

Additional control is available in most implementations of BGP, which also permit routing update policies. These policies direct whether individual network routes can be accepted or advertised by a local router.

BGP Path Attributes

The seven primary BGP path attributes are listed below. Several others have been added for very specific extended functionality. Some of these attributes provide the router with additional information concerning a particular route or prefix – this data can then be used as the basis for advanced policy decisions. Others are variables that can be “tweaked” for policy or control purposes.

- **Origin** States whether the route was learned by EGP or IGP
- **AS-Path** Contains a list of autonomous systems through which the announcement for that particular route has passed

- **Next-Hop** The advertised address of the next hop node toward a particular route or destination
- **Multi-Exit-Discriminator** (Commonly called the “MED”) Used when there are multiple links between the same autonomous systems; this is a metric expressing the degree of preference for each connection
- **Local-Pref** A configurable metric used to select among multiple paths to the same prefix; works in conjunction with the MED, but only applies to decisions within a single autonomous system
- **Atomic-Aggregator** Indicates that a prefix has been aggregated and must not be disaggregated
- **Aggregator** Indicates that the router has aggregated the advertised prefix

Internal vs. External BGP

The border gateway protocol, as described above, is primarily used for interconnecting autonomous systems. This kind of operation is more accurately called External BGP or E-BGP, since it is used for establishing communications paths external to the originating autonomous systems. BGP can also be used as a routing protocol within a given AS. This implementation is known as Internal BGP, or I-BGP. The principles associated with E-BGP and I-BGP are subtly different from each other.

The most significant difference between I-BGP and E-BGP is the manner in which network advertisements and updates are handled. When an E-BGP router receives an update message from an adjacent router, it will typically flood that update to all of its peers. I-BGP routers try to reduce the amount of routing overhead (and potential looping) within an autonomous system. Hence, when they receive an update from an I-BGP peer, they will process the information and update their routing tables, but they will not forward the information to any other I-BGP routers. This means that the I-BGP router, which learns of a topology change, must bear the responsibility to directly pass that information on to all other I-BGP routers within a given autonomous system. This requires direct connectivity between all I-BGP routers within an AS; in other words, I-BGP routers must be configured in a full mesh topology.

Routers can, of course, be both I-BGP and E-BGP speakers simultaneously. For example, one link may be used to connect a particular router to an adjacent autonomous system (via E-BGP) and the remaining router links can be connected to routers within the same AS (I-BGP).

BGP Extensions – the Future is Now

Since BGP is the protocol of the Internet, it has been modified, enhanced, adapted, and sometimes even distorted to accommodate additional features, concepts, and standards. Some of the more significant extensions and enhancements are discussed below.

One of the most significant extensions of BGP is the support for multiple protocols. BGP was specifically designed for IPv4 unicast traffic. However, multiprotocol extensions are specified in RFC 2858. This provides the ability for BGP to support multicast traffic, IPv6, VPNs, and other protocols.

Other enhancements have improved the scalability of Internal BGP. Instead of requiring a full mesh topology, some autonomous systems can now use a process known as route reflection. This operates by having a router within the AS designated as a route reflector. All of the other routers in the AS will be clients to the route server. The clients will be adjacent (I-BGP peers) to the route reflector, and they will forward topology updates to the designated route reflector. This router then will forward all routing updates to all of the other routers within that AS. It is necessary for the route reflector to be adjacent to all other routers within the AS, but this alleviates the requirement for all routers to be fully meshed.

A second approach to resolve the scalability challenge of fully meshed I-BGP is the use of confederations. These are simply divisions within an autonomous system (sometimes called Sub-ASs). Each confederation will be fully meshed and a protocol very similar to E-BGP, known as EIBGP, is used to interconnect the confederations.

An additional policy and control feature was added to BGP. This is an optional path attribute known as a BGP Community (see RFC 2197). This provides network administrators with the ability to group routes based upon common administrative policies. Communities allow common routing policies to be established for specific groups of networks. In short, this is designed to simplify network policy management.

And finally, many features and options exist in order to ensure that BGP interoperates with other routing protocols – typically IGP. Many of these features are specific to the BGP implementations of individual vendors, but some fundamental principles have been established for importing and exporting (or redistributing) routes to and from other protocols.

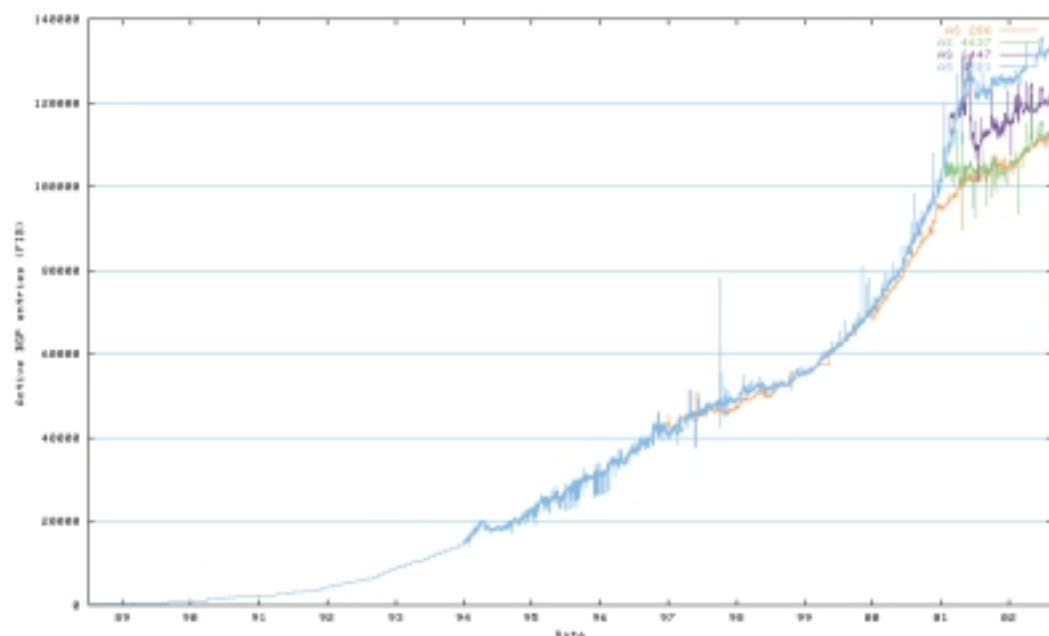
Many of these supplemental BGP features are optional, so it is necessary for two adjacent routers to agree upon which features and attributes they will support over a given network link. Therefore, there is also an optional “capabilities” field in some BGP update messages. This field is used to identify the capabilities associated with a given router, thus facilitating a capabilities negotiation process between two BGP peers.

The Internet Today

BGP is used predominantly in the Internet. Nothing precludes using BGP in other environments. In fact, some large corporations do take advantage of the many benefits of BGP within their own infrastructures. However, since the primary focus of the protocol is the Internet, it is appropriate to review the current status of the BGP-based Internet.

Today’s Internet is an intricate web of routers. Very little structure exists between the various ISPs and other Internet resources. Hierarchical architecture, redundancy, performance metrics, if they are implemented at all, are executed on a random or even haphazard basis. Untangling this web is a substantial challenge for the Internet routers.

The Internet currently contains approximately 120,000 routes. The chart below shows three views from three different autonomous systems of the historical growth of the Internet routing tables (source: <http://bgp.potaroo.net>). Keep in mind that this is only a fraction of the total quantity of Internet addresses that have been issued, since the current routers will tend to view aggregated or summarized networks.



HISTORICAL GROWTH OF THE INTERNET ROUTING TABLES

Currently 24,000 different autonomous systems have been registered worldwide. This quantity, like everything else that is associated with the Internet, continues to grow on a weekly basis!

BGP routers successfully juggle the size and complexity of the Internet. The basic fact that the Internet is operational may be credited to BGP.

Testing BGP

BGP is a powerful tool. It is a complex protocol with lots of optional features. BGP implementations vary considerably between vendors. It therefore is absolutely essential that the equipment be fully tested. This includes testing by the network equipment manufacturers, the service providers, and even some end-users. Testing includes ensuring adherence to the established standards, as well as testing the practical and functional aspects of BGP – scalability, responsiveness, and control.

Conformance Testing

The most academic level of testing is conformance testing. The purpose of conformance testing is to verify the strict adherence to the prevailing BGP RFCs. Conformance testing helps ensure multi-vendor interoperability. This will also identify any potential shortcomings, as well as value-added features associated with a particular vendor's BGP implementation. Service providers will benefit from conformance tests since this is a good way to validate vendors' BGP implementations. In fact, they should execute a full batch of conformance tests in a lab environment every time their chosen vendor releases a new software update.

BGP conformance tests consist of a whole suite of tests, each focusing on a specific portion of the protocol. These tests can typically be automated for simplicity and expediency.

Scalability Testing

BGP was designed to support scalability. All vendors understand this fundamental axiom. However, "scalability" can be a very subjective term. Quantifying this scalability can be a significant challenge. It is absolutely essential that network equipment manufacturers and service providers understand the limits of the products that they are building or deploying.

Scalability testing has several facets. The first item to test is the routing table capacity – after all, the Internet currently has 120,000 routes and is continuing to grow. A BGP tester should generate IPv4 routes, and then validate the capacity of the local router to correctly process and store these routes. The upper limits should be discovered and communicated to the appropriate network architects. The current Internet routing table should also be loaded into the router so that "real world" scalability can be verified.

Another valid scalability test – especially for I-BGP implementations – involves verifying the quantity of BGP peers that an individual router can manage. This is accomplished simply by emulating multiple BGP peer routers, each with a full Internet-sized routing table, and testing the router's ability to maintain the peering sessions, as well as processing the routing data.

Packet throughput and bandwidth capacity are also fundamental scalability and performance tests.

Responsiveness Testing

Responsiveness testing should be conducted both for devices (routers) and for systems (networks). This will validate the agility of the product and/or network and its ability to react to topological changes.

The first step toward responsiveness testing is simply measuring the convergence times for the router(s). To accomplish this, use a BGP tester to establish a peer session with the router under test, and load a full BGP routing table into the device. Measure the amount of time required to process the routing table and the latency associated with forwarding the first data packet.

After the router has established the appropriate peering sessions, test its responsiveness to network changes. While forwarding traffic, start withdrawing and adding routes, including the routes which are already in use. Measure the latency associated with processing new routing information. Also determine if any datagrams are lost, or if any incremental latency is created by these changes. Flap a few routes at a time and observe the reaction. Then flap large quantities of routes and examine the results. Next, flap some of the network links, including the ones being used for data forwarding. Again, observe and measure the data plane and control plane reactions.

Policy and Control Testing

The most complex functions of BGP are the policy and control processes. Testing these processes will validate a vendor's ability to make the correct logical choices regarding route selection. To test these functions, the test equipment will need to fully support all of the mandatory and optional BGP path attributes.

Testing will require multiple ports on the router and the test equipment. The test equipment ports will emulate BGP peers. These peers should each generate large quantities of routes. Some of the routes, or even all of the routes, should be redundant. However, the path attributes should be varied so that the router is required to select the optimal path to an end destination. Verify that traffic is being forwarded to the correct link. While traffic is being forwarded to a given destination, inject a preferable route using another physical router port to the same destination. Verify that the router redirects the traffic appropriately. Measure the response time.

Conclusion

BGP is the routing protocol of the Internet. It is designed to support the current Internet, as well as the Internet of the future. All major router vendors support it, and nearly every ISP and service provider throughout the world uses it. In short, it truly is a universal protocol.

BGP has a big job to do. It must keep track of the ever-changing, ever-growing Internet; respond to changes as they occur; be able to support the future growth of the Internet; support the ability to intelligently select the correct path between a source and destination; and accommodate new protocols, concepts, and features as they are developed.

The Internet expects a lot from BGP. As a protocol, BGP is designed to deliver the desired results. However, these results are entirely dependent upon the manner in which router vendors implement BGP. As the cliché goes: Your mileage may vary! Some vendors excel in scalability; others strive to conform to every exact detail of the specifications. All have their own unique implementations and tradeoffs. The only way one can truly understand each of these unique implementations is by fully testing the products.

References

Internet Standards and Drafts

- RFC 1771 – A Border Gateway Protocol, March 1995
- RFC 1772 – Application of the Border Gateway Protocol in the Internet, March 1995
- RFC 1773 – Experience with the BGP-4 Protocol, March 1995
- RFC 1774 – BGP-4 Protocol Analysis, March 1995
- RFC 2197 – BGP Communities Attribute, August 1996
- RFC 2796 – BGP Route Reflection, April 2000
- RFC 2842 – Capabilities Advertisement with BGP-4, May 2000
- RFC 2858 – Multiprotocol Extensions for BGP, June 2000
- RFC 2918 – Route Refresh Capability for BGP-4, September 2000
- RFC 3065 – Autonomous System Confederations for BGP, February 2001

Text Books

- Halabi, Sam Internet Routing Architectures San Francisco, Cisco Press, 2000.
- Stewart, John W. BGP4: Inter-Domain Routing in the Internet San Francisco, Addison-Wesley, 1999.

Spirent

Communications

27349 Agoura Road
Calabasas Hills, CA
91301 USA
E-mail: productinfo
@spirentcom.com

Sales Contacts:

North America

+1 800-927-2660

Europe, Middle East, Africa

+33-1-6137-2250

Asia Pacific

+852-2166-8382

All Other Regions

+1 818-676-2683

www.spirentcom.com

