White Paper

ixia

Border Gateway Protocol (BGP) Conformance and Performance Testing

26601 Agoura Road, Calabasas, CA 91302 | Tel: 818.871.1800 | Fax: 818.871.1805 | www.ixiacom.com

Contents

Ixia
26601 W. Agoura Road
Calabasas, CA 91302
Phone: (818) 871-1800
Fax: (818) 871-1805
Email: info@ixiacom.com
Internet: www.ixiacom.com

Border Gateway Protocol: Conformance and Performance Testing

Introduction

The relentless evolution of the Internet continues to transform the way individuals, as well as businesses, educational institutions, and government organizations access, share, and communicate information. Convergence of digital voice, video, and data, is further consolidating the Internet as a critical infrastructure. One of the main routing protocols in the Internet and current de facto standard is the *Border Gateway Protocol* (BGP). Presently ubiquitous, BGP is a critical component of the exponentially growing network of routers that constitutes our contemporary Internet. Carrier networks, as well as most large enterprise organizations with multiple links to one or more service providers use BGP.

The increasing popularity of BGP stems from its broad ability to distribute reachability information by selecting the best route to each destination according to policies specified by network administrators. To manage the complexity of BGP, however, a wide range of services, applications, and hardware must be tested and validated. Indeed, a comprehensive and well-designed conformance and performance testing solution is crucial to successful BGP deployment.

What Is BGP?

BGP is a protocol for facilitating communications between routers in different autonomous systems. An autonomous system (AS) is a network or group of networks under a shared technical administration and with common routing policies.

Network traffic in an AS is classified as either *local traffic* or *transit traffic*. Local traffic either comes from or terminates in that AS, where either the IP host source address or destination address reside. Any other traffic traversing that AS constitutes transit traffic. A major goal of BGP usage in the Internet is to reduce transit traffic.

BGP advertises routes as a “promise” to carry data to the address space indicated by the IP prefix of the announced route. Generally, all routes in a BGP routing table outline Internet network connections. When a BGP router advertises to a neighbor that it has a path for reaching a specific IP prefix, the neighbor can be confident that the advertising BGP speaker is actively using that path to reach the target destination. Route advertisements in BGP use the AS-Path attribute to announce current routing to neighbor BGP speakers, which includes a list of all transit ASs that must be used to reach the target network. By carrying path information associated with a given destination between autonomous systems, BGP enables loop-free inter-domain routing.

BGP conveys information about AS-Path topologies and achieves inter-AS routing without constraining the underlying network topology. An intra-AS routing protocol—that is, Interior Gateway Protocol (IGP), examples of which are Routing Information Protocol (RIP), Open Shortest Path First (OSPF), etc.—provides the routing within an autonomous system. In some circumstances, BGP is used to exchange routes within an AS. In those cases, it is called Internal BGP (I-BGP), as opposed to External BGP (E-BGP) when used between ASs.

Internal And External BGP

A BGP router can communicate with other BGP routers in its own AS or in other ASs. Both the I-BGP and E-BGP implement the BGP protocol with a few different rules. All I-BGP-speaking routers within the same AS

must peer with each other in a fully connected mesh. They are not required to be physical neighbors, just to keep a TCP connection as a reliable transport mechanism. Because there is no loop detection mechanism in I-BGP, all I-BGP-speaking routers must not forward any 3rd-party routing information to their peers. In contrast, E-BGP routers are able to advertise 3rd-party information to their E-BGP peers, by default.

Figure 1 shows routers R1, R2, and R3 using I-BGP to exchange routing information within the same AS, and router pairs R4-R2, R3-R5, and R4-R5 using E-BGP to exchange routing information between ASs.

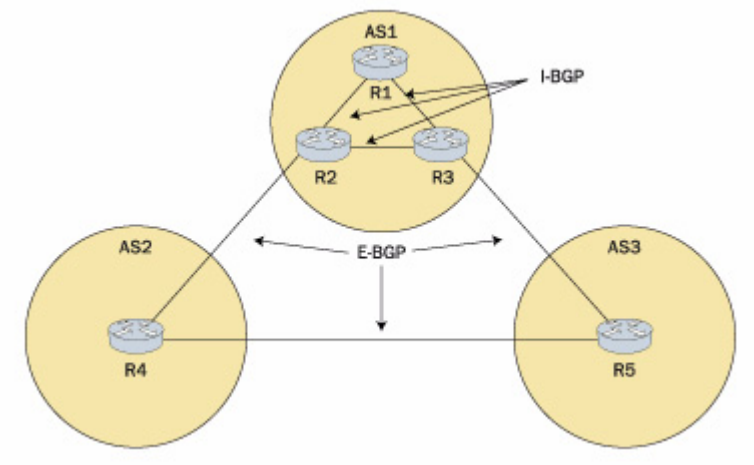

Figure 1. Internal BGP (I-BGP) versus external BGP (E-BGP).

Historical Perspective

Originally defined in RFC 1105, BGP became an Internet standard in 1989. It replaced the older Exterior Gateway Protocol (EGP) used on the ARPANET, becoming the EGP of choice for inter-domain routing. Since then, BGP has gone through several enhancement cycles and extensions. In 1990, it was updated to BGP-2 by RFC 1163; and in 1991, it was updated to BGP-3 by RFC 1267. The current version of the Border Gateway Protocol, BGP-4, was defined in RFC 1771 and adopted in 1995. All prior versions are now obsolete.

Several extensions for BGP-4 have been proposed since then. RFC 2283 defines BGP-4+, which includes IPV6 prefix advertisement and other important enhancements, increasingly supported by most Network Equipment Manufacturers. Another proposed extension is BGP's graceful restart capability, to improve recovery times and reduce the effect of software and equipment failures on IP routing. Multiple academic and industry-based contributions coordinated through IETF RFCs continue to modify and expand the scope of BGP.

How Does BGP Work?

The foundation of BGP is an asynchronous, distributed, preferred-path vector algorithm that uses TCP as its transport protocol. In contrast with OSPF and RIP, the use of TCP as BGP's transport protocol guarantees transport reliability (such as retransmission) and eliminates the additional complexity related to designing reliability into the protocol itself. BGP protocol data units are enclosed within TCP packets and the reliable transport layer protocol is used for acknowledgement, sequencing, fragmentation, and retransmission. After

setting up a BGP session and exchanging initial routes, BGP peers trade incremental routing and notification updates.

Protocol Overview

Routers that run a BGP routing process are often referred to as *BGP speakers*. A pair of BGP-speaking routers that form a TCP connection to exchange routing information between them are called BGP *neighbors* or *peers*. A single router can participate in many peering sessions at any given time. Each BGP session takes place exactly between two nodes, where two routers exchange routing information dynamically, over TCP port 179.

For any two BGP peers in a network to be able to send and receive traffic with each other, all intermediate BGP routers have to forward traffic such that the packets get closer to the destination. Because there can be multiple paths to a given target, BGP routers use a *routing table* to store all known topology information about the network. Based on its routing table, each BGP router selects the best route to use for every known network destination. That information is stored in a *forwarding table* together with the outgoing interface for the selected best path.

With BGP, it is not necessary to refresh routing information as with many other routing protocols. Instead, when a router advertises a prefix to one of its BGP neighbors, that information is considered valid until the first router explicitly advertises that the information is no longer valid or until the BGP session itself is lost or closed. It is assumed that the transport connection will deliver all data and eventually close properly in case of an error notification.

There are four possible message types used with BGP, all consisting of a standard header plus specific packet-type contents:

- **OPEN**: First message to open a BGP session, transmitted when a link to a BGP neighbor comes up. It contains AS number (ASN) and IP address of the router who has sent the message.
- **UPDATE**: Message embracing routing information, including path attributes. It contains Network Layer Reachability Information (NLRI), listing IP addresses of new usable routes as well as routes that are no longer active or viable and including both the lengths and attributes of the corresponding paths.
- **NOTIFICATION**: Final message transmitted on a link to a BGP neighbor before disconnecting. It usually describes atypical conditions prior to terminating the TCP connection, and provides a mechanism to gracefully close a connection between BGP peers.
- **KEEP-ALIVE**: Periodic message between BGP peers to inform neighbor that the connection is still viable by guaranteeing that the transmitter is still alive. It is an application type of message that is independent of the TCP keep-alive option.

The BGP protocol has four main stages:

1. **Opening and confirming a BGP connection with a neighbor router.** After two BGP peers establish a TCP connection, each one sends an OPEN message to the other.
2. **Maintaining the BGP connection.** A BGP router can detect a link or BGP peer host failure through the exchange of periodic keep-alive messages with the peer router. An error is assumed when no messages have been exchanged for the *hold timer* period. The hold timer period is calculated the smaller of its configured hold time setting and the hold time value received in the OPEN message. BGP utilizes periodic keep-alive messages to ensure that the connection between neighbors does not time out. Keep-alive packets are

small header-only BGP packets without any routing data.

3. **Sending reachability information.** Routing information is advertised between a pair of BGP neighbors in update messages. Each update message may simultaneously advertise a single feasible route to a neighbor and indicate withdrawal of several infeasible routes from service. Update messages contain NLRI with a list of <length, prefix> tuples designating reachable destinations, and path attributes, including degree of preference for each particular route, and the list of ASs that the route has traversed.
4. **Notifying error conditions.** Notification messages are sent to a neighbor router when error conditions (incompatibility, configuration, etc.) are detected. Notification messages consist of a main error code and a more detailed sub-code. Through the notification mechanism, a graceful close guarantees the delivery of all outstanding messages prior to closing the underlying TCP session.

AS Consistency

BGP mandates that each AS providing transit to other ASs expose the same view to all other AS's using its services. All BGP speakers of a given AS must be consistent in their representation of the topologies both intra-AS and inter-AS. However, BGP does not specify which method should be used to reach, maintain, and enforce the consistency. For example, OSPF can be used to synchronize router databases for intra-AS topology consistency, and rely on BGP itself can be used for inter-AS topology consistency.

BGP Route Advertisement

After establishing a TCP connection, two adjacent BGP speakers exchange full routing information. Each BGP router may receive multiple advertisements for the same route from multiple sources. Based on the described paths, the router filters them and selects only one as the best path, puts it in its IP routing table, and propagates the path to its neighbors. By sending a route announcement to a neighbor, the advertising BGP router is implicitly agreeing to forward IP traffic to the destination network on behalf of the neighbor. If a BGP router determines that a route is inaccessible, it informs all its BGP neighbors of the withdrawal of the route. When a BGP speaker determines that a route has changed or that a new path for the same prefix is chosen, it advertises the replacement route without requiring a route withdrawal.

For every neighboring BGP speaker, the administrator of a BGP router may set input policy filters to sort out route advertisements and perform attribute manipulation. For example, the filter could allow only advertisements such that paths going through a specific AS will not be used, or that include trustworthy ASs in the AS-Path, leaving out all other route notifications. The BGP routing table consists of only accepted routes that pass through the route-advertisement input filter; duplicates are not included.

A BGP router sends at most one route per destination to its BGP peers. It uses output filters to choose the destinations that will be advertised to each BGP neighbor, and leaves out routes that will not be advertised to one or more neighbors. BGP routers can be configured to modify route attributes before sharing routing information with a particular BGP peer. A BGP speaker can use a particular route while simultaneously choosing not to announce it to an external peer. If the peer has previously received an announcement for it, then the routing BGP speaker must report to the external peer that the previous route is now no longer available.

BGP routers use a table version number to keep track of their present routing table instance, incrementing the number every time the routing table changes. Rapid increase of table version numbers frequently indicates network instability. Therefore, *route flap damping* and other mechanisms have been implemented to cope with unstable networks that cause rapid table version number increases for any BGP speaker with access to the Internet routing tables. Nonetheless, such rapid increases are typical of large carrier networks connected to a great number of BGP speakers.

Route Flap Damping

Recommended in RFC 2439, BGP supports route flap damping (RFD) to reduce the impact of problems to a localized area in the network. RFD minimizes the instability caused by route flapping by suppressing the propagation of unstable BGP routes. The main parameters characterizing RFD are:

- **Penalty:** Metric that is incremented every time a route flaps. It is decremented over time at a given rate.
- **Half-life time:** Rate at which the penalty value is to be reduced to half the current value.
- **Suppress limit:** Threshold above which a route is suppressed.
- **Reuse limit:** Threshold below which a suppressed route is reclaimed.

RFD's goal is to reduce router-processing overhead due to instability without sacrificing convergence time for stable routes. Damping is not set up per path but per prefix. To achieve effective stabilization, BGP must distinguish between persistently unstable routes and those routes that only occasionally fail.

BGP Path Selection

BGP uses a preferred *path-vector* (PV) algorithm, described in RFC 1322, that itemizes the complete path to a destination. The PV routing algorithm supplements the advertisement of reachable destinations with information that describes various properties of the paths to these destinations. A path is the recorded sequence of ASNs through which the reachability information has passed. Each AS is considered equal, independent of its size and internal composition.

Different autonomous domains can have different route optimality notions. This is because PV only standardizes the results of route selection while allowing heterogeneous criteria across domains. Each AS can have its own policies for route selection. To prevent forming loops, BGP routers ignore any routing advertisement that contains their own ASN anywhere in the AS-Path. To originate a route, a BGP router creates an empty—null—path and advertises it to its neighboring BGP routers with its ASN prepended to the otherwise empty AS-Path.

BGP uses the shortest AS-Path routing criterion (lowest number of ASs that the route has traversed through) by default. However, "shortest" does not always mean "best" path to reach a destination prefix. Because the underlying network topology is unknown to BGP, a single AS hop could in fact correspond to multiple router hops. Further, default BGP routing is oblivious to network performance metrics, ignoring network parameters as congestion, packet loss, delay, and jitter. Tuning BGP for optimal or near-optimal routing depends on policies to modify the default behavior, and to allow for the best performing paths even when those are not the shortest ones.

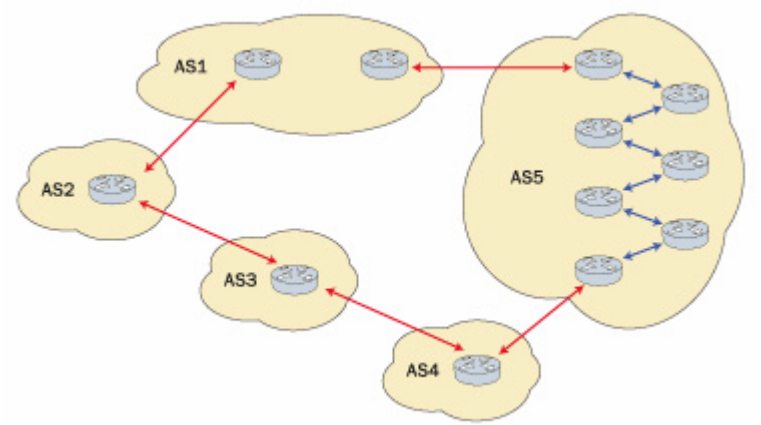

Figure 2. Which is the shortest path? BGP considers AS1-AS5-AS4 shorter than AS1-AS2-AS3-AS4 because it only counts AS hops and not internal router hops.

BGP routers use several route parameters, described by *BGP attributes*, to delineate routing policies. In addition to BGP attributes, BGP-4 introduced *route aggregation* mechanisms to reduce the size of the Internet routing tables. The aggregation technique is called classless inter-domain routing (CIDR) (or supernetting). CIDR represents IP addresses with common high-order bits by using shortened subnet masks. For routing purposes, only bits covered by the subnet mask are used, thus making all aggregated addresses to look like members of the same network.

Route calculations are influenced by re-configurable router settings that specify *route preferences*. For example, the preferences may specify that a destination not be advertised to some neighboring BGP speakers, or that a path through a given AS should not be used or should be edited when passing it to a specific neighbor.

BGP Policies And Traffic Engineering

BGP provides mechanisms for policy-based routing, which enables BGP routers to rank routes and control information redistribution according to their administrator's preference. BGP carries out policy routing by filtering certain routes, based on IP-prefix, AS-Path, or other attributes; or by adjusting selected attributes to influence the route selection process. Policies are not part of the protocol; they are decisions made by the AS administrator, and are specified to BGP by the AS administrator in configuration files. Routing policies are often coupled to security, economic, regulatory, or political considerations.

Through policy-based routing, BGP enables different implementations to specify path selection rules when many options are present and to control information distribution. Beyond using the routing table longest match criterion (which uses the routing table entry table that most specifically matches the target destination) and the shortest AS path, different vendors implement the BGP path selection criteria by checking BGP attributes in a slightly different order. In all cases, when a route is advertised, the ASN of the advertising router is added to the route. By stamping the sequence of ASNs, an AS-Path traces how the route became known to any of the routes in the trail.

One of the advantages of BGP's policy routing is that filtering is a local technique. Thus, changes can be applied promptly and without advertising the policy. However, while policy localization reduces the control overhead of the protocol, the absence of synchronized policies and lack

of global information often leads to sub-optimal route selection.

BGP implementations assume that there is a local method of managing a BGP router, constructing a function that takes as input all the information advertised in a BGP update message about a particular destination and outputting a number. After different possible routes are mapped to numbers, the routes can be compared. The preferred route is the one that maps to the smallest number.

Policy Convergence. BGP's lack of policy synchronization often leads to convergence concerns. Product specifications typically describe convergence time as a single numerical value. However, there are in fact two different kinds of convergence time: the time it takes a BGP router to build its full routing table after initialization, and the time it takes for a BGP router to react to a route announcement or withdrawal. It is important, therefore, to specify which convergence is being considered.

Because routers can have their own policies, the policies can tolerate convergence problems. There are policies that never converge, triggering ever-changing routes, which propagate adjustments in other routers as well. There are convergent policies that become non-convergent under some topology changes (e.g., when a router or link in the path goes down). Some policies may or may not converge, solely depending on message ordering. Lastly, the combination of some routing topologies and policies can result in scenarios where it is not computationally feasible to calculate policy convergence (for both convergence time definitions), requiring significant human intervention and rigorous testing to accurately measure convergence times.

Inter-Domain Traffic Engineering. BGP's default behavior is to attempt sending traffic over the route with the shortest AS-Path. Even though bandwidth is continuously getting cheaper, it is generally useful to balance traffic to take advantage of all the available bandwidth in a multi-homed setup. Traffic engineering is about getting network traffic to take the best route to a destination, and is performed by favoring one link over another to reach a given destination or to receive traffic from a given source.

In general, it is easier to engineer outgoing traffic than for incoming traffic because administrators only have control over what their own routers do.

BGP traffic engineering practices are intended to provide good scaling properties that result in predictable changes to traffic flows, and help limit the influence of neighboring domains. Alleviating congestion on edge links, adapting to provisioning changes (e.g., link capacity), and achieving good end-to-end performance are some of the goals for traffic engineering.

However, deep traffic engineering was not part of BGP's original design. BGP does not provide a very flexible and direct language, which can result in a restrictive decision process with limited control that requires many interactions with neighbors. However, by directing traffic to a different neighbor AS, and directing traffic with different links to the same neighbor, it is possible to control the influence of and on neighboring domains, achieving good scaling properties resulting in predictable changes to traffic flows.

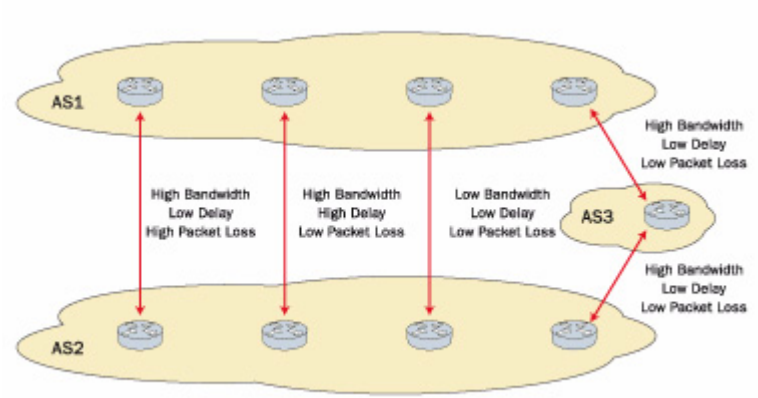

Figure 3. Which is the best route between AS1 and AS2? BGP's choice of the best path is relative to the administrator priorities, and it depends on the BGP attribute settings.

Figure 3 shows a scenario with multiple paths between AS1 to AS2. The definition of best route here is not always obvious, and depends on which parameter is more important to the network administrator. In this figure, bandwidth, delay, packet loss, and number of hops are considered.

BGP Attributes

BGP attributes are metrics that describe characteristics of routed prefixes in a BGP path. They are used to shape routing policy. For example, some of the attributes can be used in combination to equalize the distribution of inbound and outbound traffic among available multiple paths, and to prevent route-flapping while fine-tuning routing for load balancing. (By default, BGP does not load-balance traffic; it selects and uses the accepted "best" route.)

The attribute information is forwarded when BGP peers advertise routes using UPDATE messages. There are several types of BGP attributes:

- Well-known Mandatory.
- Well-known Discretionary.
- Optional (or partial) Transitive.
- Optional (or partial) Non-transitive.

A *well-known* attribute is one that all BGP implementations must be able to understand and are transmitted to all BGP neighbors. *Optional* attributes may not be supported by all BGP implementations. A *mandatory* attribute is one that must show in the description of a route. A *discretionary* attribute that does not have to appear. A *transitive* attribute is an optional attribute can be passed unmodified by a BGP speaker that does not have an implementation for it. After a transitive attribute has been passed, it is marked as a *partial* attribute. A BGP speaker that does not have an implementation for a *non-transitive* attribute must delete it, not passing it to other BGP peers. Commonly used attributes are listed in Table 1.

Table 1. Commonly used attributes.

Well-known, Mandatory Attributes	
AS-Path	List of ASs along the path to reach the destination. As the update passes through an AS, the ASN is inserted at the beginning of the list. The AS-Path attribute has a reverse-order list of ASs passed through to get to the destination. The AS-set object can be used to set an unordered set of AS's when aggregating routes and choosing to keep attribute information about the components of the aggregate.
Next-hop	IP address of the BGP router that should receive data packets with the intention of getting them closer to the target destination. While in most cases the advertising router is the same one that will receive and forward data packets, it is sometimes better to have one BGP speaker announce routes on behalf of another BGP peer that, in turn, will perform the actual routing of the data.
Origin	Indicates how BGP learned about a particular route. There are three possible types: • IGP, where the route and prefix are interior to the originating AS, thus information is considered trustworthy. • EGP, where the route and prefix are learned via eBGP. EGP is usually less preferable than IGP, because EGP does not work when topological loops exist. • Incomplete, which indicates either unknown origin or that the route was learned in a different way, and not via IGP or EGP, which, for example happens when a route is redistributed into BGP, or for a static route.
Well-known, Discretionary Attributes	
Local Preference	Indicates the degree of preference for an external route. It identifies the preferred exit point from the local AS for a specific route, where the route with the highest local preference value is preferred. It might override preferences from external ASs. The local preference path attribute is always advertised to I-BGP peers and neighboring confederations. It is never advertised to E-BGP peers.
Atomic Aggregate	Indicates that the aggregation of routes has caused some path attribute information to be lost.
Optional, Transitive Attributes	
Aggregator	Used together with the atomic aggregate attribute, specifies AS number and router ID of the router that executed the aggregation.
Optional, Non-transitive Attributes	
Multi-Exit-Discriminator (MED)	Also known as the *external metric attribute* of a route, it provides information about which path should be selected by external neighbors accessing an AS with various entry points. MED is advertised to external neighbors, suggesting to external peers the relative preference of entry points and defining a preferred path into the advertising AS. Because current RFCs do not require MED comparisons, vendor-specific implementations of path ordering can influence the routing decision process.

Table 1. Commonly used attributes. (Cont.)

Community	Group of destinations that share common properties so that policies can be applied at the group level. The community attribute is not restricted to one AS or network, each destination can belong to multiple communities. It indicates a set to which the destination belongs so that policy configuration can be done by group rather than by single prefixes.

BGP Extensions

There are multiple extensions to the original BGP-4 protocol. These extensions either fix problems or limitations of the original proposal, or add new functionality. For example, fully connected meshes cause scaling problems in I-BGP, which are fixed by the following two methods. One is *route reflection* (RFC 1996), where some BGP speakers called route reflectors in the AS are allowed to collect BGP information and forward it within the domain. The other solution, called *confederations* (RFC 1965), allows aggregation of many ASs within a bigger confederation-AS, as well as the possibility to subdivide an AS.

Route Reflectors

A route reflector (RR) is a concentration router acting as a focal point for I-BGP sessions, which adds a hierarchy level to I-BGP. Described in RFC 1996, a route reflector supports route re-advertisement between I-BGP neighbors to alleviate the need for a full mesh.

A route reflector client is a regular BGP speaking router that depends on a RR to re-advertise its routes within their AS and to learn about routes external to their AS. An AS can have more than one RR and each RR can receive AS-Paths from clients and non-clients. If the best path is from a client, it reflects to both clients and non-clients. If the best path is from a non-client, it reflects it only to clients, maintaining the behavior for re-advertisement between non-clients.

Route reflection is primarily recommended for ASs with large internal meshes, and is not recommended for every topology.

In Figure 4, R11 and R12 are a single cluster, where R12 is the RR and R11 is the client. Similarly, R16 is the RR of the cluster that also includes R15 and R17. R14 is a stand-alone RR, forming a full mesh with the other two RR's within AS1.

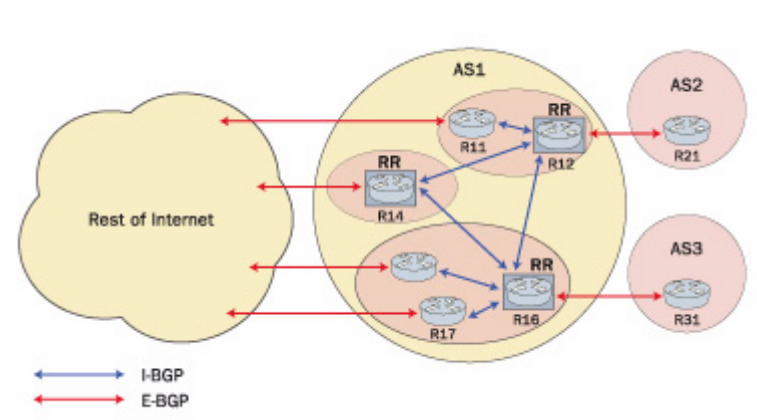

Figure 4. Example of route reflector utilization, where only route reflectors are connected in a full mesh.

AS Confederations

A confederation is a group of ASs that looks to outside routers as if they all were a single AS, with a regular ASN. Autonomous system aggregation makes it possible to simplify policies and traffic engineering tasks by using the confederation to represent multiple ASs in a path, or by blocking routes that go through the confederation, instead of explicitly listing all ASs. The addition of such a level of abstraction and hierarchy, however, impacts routing efficiency while lowering protocol overhead (such as storage, processing times, etc.). Simply stated, the more the aggregation the less optimal the routing.

To avoid loops, a confederation can only appear once in an AS-Path, possibly leading to sub-optimal routing in cases such as the one shown in Figure 5. While ASs inside the confederation do share their ASNs, they are invisible outside the confederation, and are replaced with the confederation identifier.

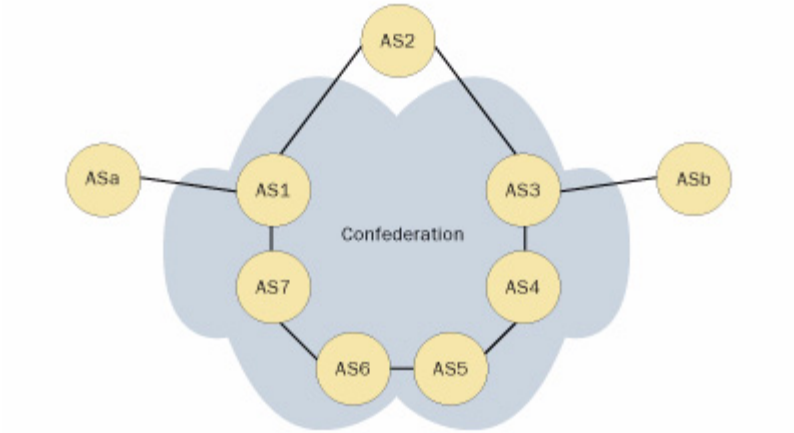

Figure 5. Confederations: To avoid loops, a "no route re-entry" rule is enforced.

Figure 5 suggests that routing from ASa to ASb would prefer the path ASa-AS1-AS2-AS3-ASb (if there would be no confederation, or if loops would not be a problem). Because it is not possible to enter a confederation more than once in a single AS path, the only possible path from ASa to ASb is ASa-AS1-AS7-AS6-AS5-AS4-AS3-ASB.

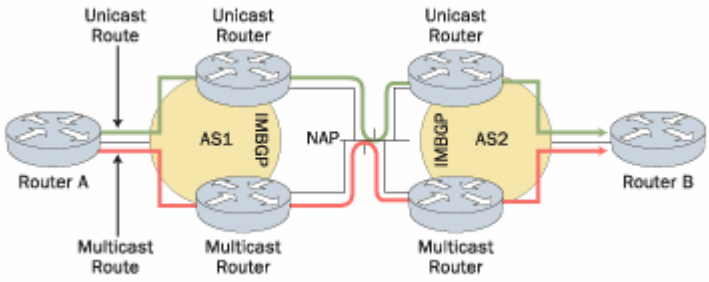

Figure 6. Multi-protocol BGP.

BGP Multi-Protocol Extensions

Multi-protocol BGP (MBGP), also called BGP-4+, is defined by RFC 2283. MBGP is an extension to BGP that expands routing support from IPv4 to other network layer protocols and is used mostly for ISP-to-ISP

peering and for multi-homed networks. Although currently IPv4 is the most common of these protocols, the transition to IPv6 is already underway. Of course, there are parts of the Internet that do forward other protocols, such as IPX, VINES, VPNv4, and others. The BGP extensions also allow multicast routing information between BGP peers to be forwarded. MBGP speakers keep two sets of routing databases, one for unicast and another one for multicast. These databases operate over BGP and use both TCP and UDP.

BGP-MPLS VPN Support

Internet service providers can offer VPN services to their customers using their existing IP backbone infrastructure. RFC 2547bis describes BGP/MPLS VPNs, where BGP is used to distribute VPN routing information, using MPLS to forward VPN traffic from one VPN site to another. The main goals are service simplification for customers, with scalability and flexibility of the service to facilitate large-scale deployment. Further, the objective is to allow policies in VPN creation that can be implemented by the service provider alone, or jointly by customer and service provider.

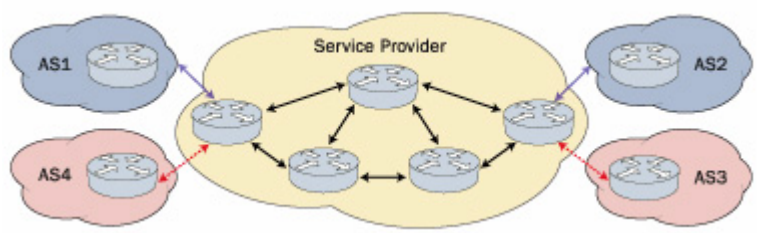

Figure 7. Customer edge-routers in AS1, AS2, AS3, and AS4 can distribute IP VPN routes to Service Provider routers, to reach other routers in other ASs using various tunneling techniques. In BGP/MPLS VPN, multi-protocol BGP is used to circulate VPN routes, and MPLS is utilized to forward VPN packets over the service provider backbone.

Figure 7 illustrates the fundamental building blocks of a BGP/MPLS VPN. Multi-protocol BGP is used to circulate VPN routes, and MPLS is utilized to forward VPN packets over the service provider backbone. Each provider edge (PE) router functions as a collection of virtual routers, one per VPN. The service provider institutes a mesh of MPLS Label Switched Paths between all the PE routers that have to communicate. To build a map of destinations and VPN labels, all PE routers qualify external IP addresses that they learn with a per-VPN identifier, sharing them with all other PE routers using multi-protocol BGP, and include MPLS labels for the destination route or destination port.

Extension For IPv6

There is no specific BGP for IPv6. There are IPv6 extensions for BGP defined in RFC 2545, which are based on the general multi-protocol extensions for BGP as defined in RFC 2858, and originally in older RFCs, now obsolete. The extensions for IPv6 are based on BGP's network layer protocol information exchange, beyond IPv4.

BGP Security

Some of the benefits of using such a ubiquitous transport protocol such as TCP come at the expense of associated vulnerabilities. To protect the BGP data stream from potential attacks, BGP can run over IPsec or use TCP MD5, a secure version of the transport protocol described in RFC 2385. The latter is most common in

current BGP implementations, and sets up a secure signature for the TCP packets based on a cryptographic protection. Every packet in BGP's TCP session contains a field with the secret key and the MD5 checksum of the packet content. BGP peers that use the TCP MD5 transport mechanism automatically discard any packet without the appropriate signature.

BGP Testing

Why test for BGP conformance and interoperability?

BGP standards and implementations are continuously adapting to the ever-changing needs of the industry. At the time of this writing, BGP and its evolving extensions have over 250 associated IETF drafts, and over 100 related RFCs. Various vendors present significantly different BGP implementations. In such a dynamic setting, the compliance of the equipment with accepted industry standards is crucial.

Service providers, network operators, and many enterprise organizations present multi-vendor environments. Conformance test tools, with a precise and thorough test methodology, can identify and isolate problems prior to deployment. Conformance testing results in increased product quality and customer confidence. For Network Equipment Manufacturers (NEMs), providing interoperable products is a key element to success in the introduction of any new product. Problems identified earlier in development reduce costly last-minute rework and post-deployment problems. Thus, NEMs must test interoperability and conformance between products in their own product lines, and in many cases test their interaction with relevant competitors.

A growing number of companies use network equipment from a primary single vendor, mainly to reduce support and management costs. In these more homogeneous environments, interoperability and system integration are easier to achieve. However, this approach relies on the capability of the strategic single vendor to provide technological innovations and product updates that will continue to serve the organization.

Testing interoperability is also important in homogeneous environments. Most large NEMs have multiple product series, with different groups that may even compete within the vendor's organization. For mission critical networks, the return on investment of assuring vital data cannot rely on the vendors' internal interoperability tests alone. Financial corporations, medical institutions, and a growing number of Fortune 1000 companies that have single-vendor environments protect mission critical data and equipment through preventive interoperability and conformance testing before and after deployment.

Lastly, homogeneous environments can also benefit from interoperability testing to certify that the equipment can work with other vendor's equipment, in the event that upgrades or new technologies are needed from other manufacturers. Since test cycles are short and require very frequent runs, these tests are often automated. To address these challenges, most vendors and service providers rely on third-party conformance testing products, maintained and supported by a dedicated organization.

Why test for BGP scalability and performance?

Scalability and performance tests are key for both vendors and end users alike. NEMs must understand the performance boundaries of their products both for engineering purposes and to generate accurate specifications. Customers must verify vendor claims within their own specific network settings. Network

managers must understand the scalability limitations and performance bottlenecks of each network element before deployment.

Scalability testing is critical to understanding network dynamics and their limits as new customers are added. Given the advances in hardware-based routing in recent years, the expectations for device performance have grown so that line rate traffic support is typically a given. To characterize BGP performance bottlenecks properly requires a test bed that can overrun the performance and scalability limitations of a device or system under test. It is critical to generate realistic BGP traffic for capacity testing, as well as randomized route instability to verify BGP speakers' ability to converge to stable routing, while measuring convergence times and the effects of chosen policies. Creating such a test bed from hundreds of routers or switches is prohibitively expensive and difficult to manage. NEMs and service providers need test tools that can simulate real-world network conditions affordably and manageably. To stress test both the control and data planes adequately, the test tool needs to emulate thousands of routers and generate wire-speed traffic, manipulating the mandatory and well known route attributes of one or many routes to create realistic Internet scenarios.

Both equipment vendors and network operators can benefit from a test methodology that can characterize data plane scalability and performance, including such metrics as:

- Throughput.
- Latency.
- Jitter.
- Packet loss.

and control plane performance metrics such as:

- Size of forwarding information base.
- Routing scalability.
- Route convergence.
- Routing stability.

Together, scalability and performance metrics can be competitive differentiators for equipment vendors. For service providers and network managers, they are a key selection criteria between vendors. Characterizing these elements is critical, since they directly impact the service quality that can be delivered to the end customer.

Ixia's approach to BGP testing

Conformance testing

Ixia has addressed the challenges of protocol conformance testing by developing IxANVL (Ixia Automated Network Validation Library), the industry standard conformance test suite.

IxANVL™

IxANVL™ is a data network testing solution that validates the protocol implementations and operational robustness of networking devices. For protocol conformance testing, IxANVL supports over 30 protocols overall, and the BGP conformance test suite contains more than 300 test cases to validate routers and hosts. IxANVL provides positive as well as negative test cases against the RFCs that specify these standards. Negative tests help validate device response to "killer packets."

IxANVL performs its tests as a dialog: it sends packets to the router being tested, receives the packets sent in response, and then analyzes the response to determine the next action to take. This allows IxANVL to test complicated situations or reactions in a much deeper and flexible way than can be done by simple packet generation and capture devices.

IxANVL can run on standalone workstations or via Ixia's optimized test platforms. IxANVL can be completely automated using a scripting interface. IxANVL source code is also available to users for customization, allowing a great degree of testing flexibility.

By incorporating IxANVL into the development and test processes, users save valuable time and money. Whether testing protocol interoperability or regression testing with new releases, IxANVL has proven to be an indispensable tool for numerous leading Network Equipment Manufacturers, Internet Service Providers, and Embedded Stack Developers for Communications Processors.

Protocol Emulations

As routers become increasingly complex, so must the analysis equipment designed to assess their performance. Such sophisticated analysis systems must incorporate powerful applications for routing protocol analysis that are flexible, highly scalable, and easy to use. Ixia's routing emulation software gives users the flexibility to customize protocol operation and meet a wide range of application requirements to test complex routing topologies consisting of thousands of routers advertising millions of routes. Sophisticated configurations can be created using Ixia's IxExplorer interface, and automated tests can be run using the IxScriptMate application.

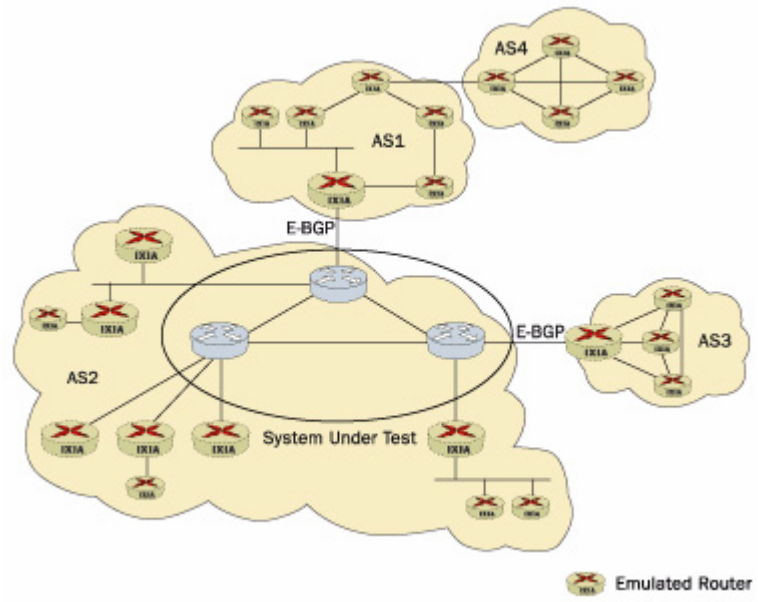

Figure 8. System test using Ixia's router emulation.

IxExplorer™

Ixia's BGP Emulation Software within IxExplorer offers an extensive set of features for testing the performance and scalability of BGP routers running over IPv4 and IPv6 protocol stacks. All mandatory and many additional BGP attributes are exposed to facilitate complex network configurations. Any combination of peers can be defined as I-BGP routers and inter-autonomous E-BGP routers. Thousands of BGP routers can be simulated and millions of routes can be advertised from a single Ixia test port. Multiple Ixia protocol emulations can be run simultaneously on each test port in conjunction with wire-speed data traffic to simultaneously test the data and control planes. Configurations can be created or changed while the BGP state machine is running, providing on-the-fly testing scenarios.

Traffic streams can be automatically generated for sending data across advertised BGP routes. Customized traffic streams can be configured with Ixia's IxExplorer, well-renowned for its traffic generation and analysis flexibility.

Automated scripts can be quickly created using the Tcl scripting environment. Alternatively, the IxExplorer GUI can be used to set up a test configuration, then Ixia's ScriptGen utility used to translate the GUI settings to Tcl code with minimal commands. Tcl support is available on Windows and UNIX platforms.

BGP statistics in IxExplorer describe the state of the BGP session with a configured peer. A BGP session can be in any one of six states:

- **IDLE:** The BGP session cannot be established and is idle.
- **CONNECT:** The BGP Session is attempting to connect to the configured peer.
- **OPEN SENT:** The "Open" packet has been sent.
- **OPEN CONFIRM:** A response to the "Open" packet has been received from the peer.
- **ACTIVE:** The BGP Session is actively attempting to connect to the configured peer.
- **ESTABLISHED:** The BGP Session is up and routes are being shared between peers.

IxScriptMate™

IxScriptMate provides a framework for running automated test scenarios. Numerous test suites have been developed within the IxScriptMate environment for testing BGP traffic throughput performance, latency, tunneling and routing performance, and scalability. IxScriptMate simplifies the configuration process by defining a configuration for the test and displaying the relevant parameters for user input. Tests then run automatically, and the results are presented to the user.

Conclusion

BGP is a core component of the Internet, connecting virtually all autonomous systems across the globe. It has prevailed due to its continuous adaptation to varying requirements, and will continue to be the standard protocol of inter-domain routing.

Equipment vendors, carriers and service providers, as well as enterprise customers depend on the interoperability, scalability, and performance of their network equipment to perform multiple services, critical to their communications and core infrastructures. Handling BGP's dynamic complexity requires proficient testing tools and methods like Ixia's family of products:

- IxANVL, the industry standard conformance test suite.
- IxExplorer, providing flexibility and functionality in protocol emulation, traffic generation, and analysis.
- IxScriptMate, providing the efficiency of automated testing.

Appendix: BGP Testing Examples

1. BGP Conformance Test

Objective. Verify the Device Under Test's (DUT's) compliance with the capabilities defined in various BGP specifications: RFC 1771, RFC 1772, draft-ietf-idr-bgp4-12, and draft-ietf-idr-bgp4-17.

Setup. A minimum of two network connections is required from the test tool to the DUT—one for request packets and one for response packets. Ixia's IxANVL conformance test solution is run from a Linux workstation either connected directly to the DUT, or via Ixia test hardware (see Figure 9). IxANVL emulates various BGP topologies, depending on the configuration of each test case.

Figure 9. BGP conformance test setup.

Input parameters. Two sets of parameters are required prior to running conformance tests: one for test tool configuration and one for DUT configuration. The test tool configuration describes the interface and protocol configuration of the tester, while the DUT configuration describes the BGP commands sent to the DUT using Expect scripts (see Table 2).

Table 2. Conformance test input parameters.

Parameter	Description
Test Tool Configuration	Tester Test IP Addresses, DUT IP Address, BGP protocol parameters (AS number, authentication, and timer values).
DUT Configuration	BGP features (TOS Routing, timers, AS number, peer configuration, etc.) via Expect scripts.

Methodology. Conformance testing is an important tool to verify how a DUT complies with specific protocol standards. Conformance test tools perform their tests as a dialog: They send packets to the router being tested, receive the packets sent in response, and then analyze the response to determine the next action to take. This methodology allows conformance test tools to test complicated scenarios much more intelligently and flexibly than achievable by simple packet generation and capture devices. Conformance testing also includes negative test cases to help validate device response to "killer packets".

For BGP conformance testing, a number of test cases are run against the DUT, based on the direct interpretation of various BGP RFCs. Ixia conformance testing consists of the following tasks:

1. Enter parameters to describe both the Conformance Tester and DUT configuration.
2. Select all or a set of test cases to run (see Figure 10).
3. Run the conformance tests from the user interface, or in a batch mode via

command scripts, reconfiguring the DUT automatically between test cases to match the test setup.

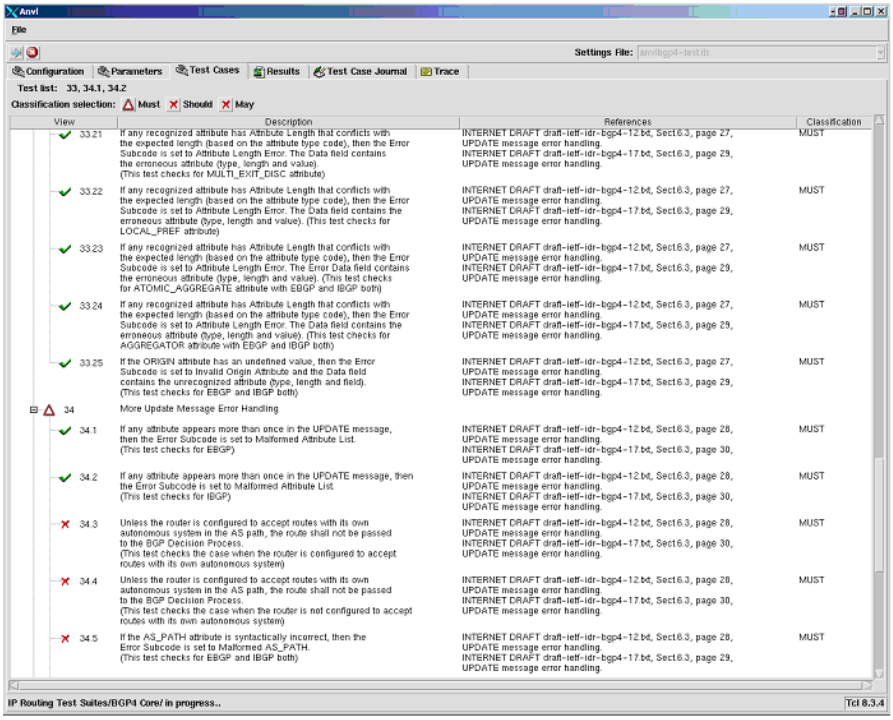

Figure 10. BGP test case selection.

Results. Number of tests passed/failed, including reasons for failed cases. IxANVL also keeps the history of each pass or fail test case in the Test Journal (Figure 11).

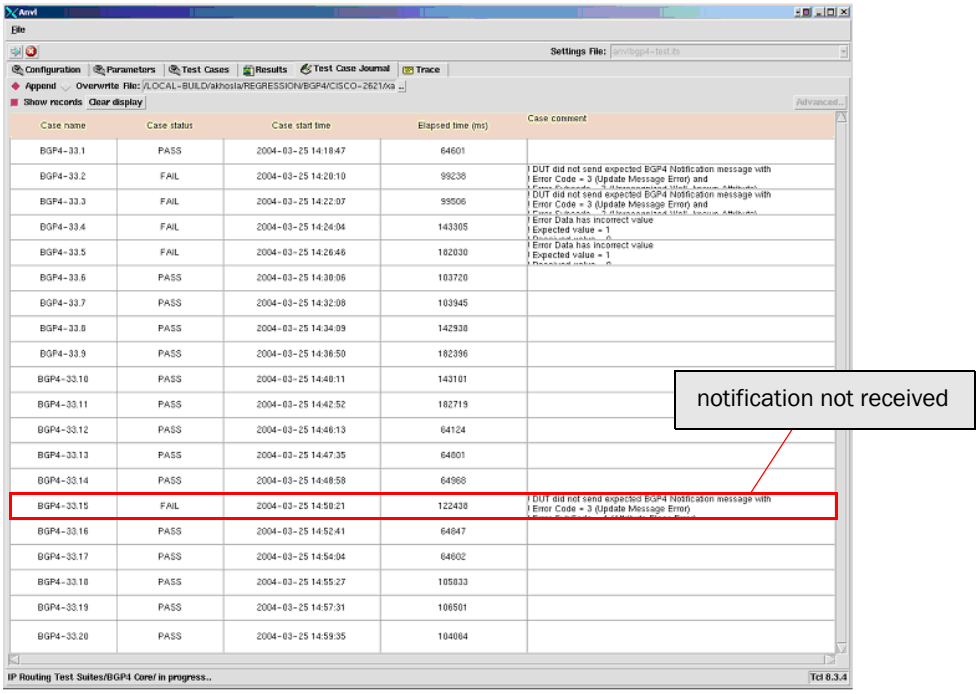

Figure 11. BGP conformance result.

2. BGP Route Capacity Test

Objective. Determines the number of routes that a BGP-enabled DUT can sustain at a single time. This scalability test is designed to help network and test engineers:

- Evaluate devices to be purchased or used in a network, based on the ability to scale with BGP.
- Test capacity and understand network limitations before actual implementation or deployment of live networks.

Setup. The test requires two tester ports—one to transmit traffic and one to receive. The transmit direction of traffic is unidirectional. Test port 2 is used to advertise the BGP4 routes, while test port 1 sends traffic to verify the advertised prefixes (Figure 12). During the test, tester port 2 increases the number of advertised routes with a "route step" until the maximum sustainable route capacity can be determined. Ixia's IxScriptMate application can be used to configure, control, and execute this test. IxScriptMate also provides comprehensive test results showing frame loss percentage based on the device's ability to forward under maximum route capacity.

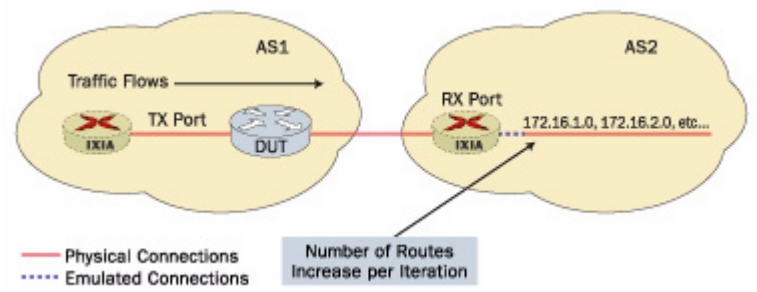

Figure 12. BGP route capacity test topology.

Input parameters. See Table 3.

Table 3. BGP route capacity test input parameters.

Parameter	Description
Max Rate	Rate at which frames will be sent to advertised routes.
Tolerance	Percentage of traffic loss tolerance.
Route Step	Number of routes to increase per iteration.
Routes Per Peer	Number of route prefixes to generate at the beginning of the test.
Delay	Maximum time in seconds the router is allowed to absorb the advertised routes.

Methodology. Route capacity testing can be summarized as follows:

1. Test port 2 advertises the initial number of routes set in "Routes Per Peer".
2. After waiting an amount of time specified by the "Delay" parameter, which is the time allowed for DUT to learn routes, test port 1 sends traffic targeting each advertised route by test port 2. The traffic throughput rate is set by the parameter "Max Rate".
3. Test port 2 verifies packets received within the defined loss "Tolerance".
4. Test port 2 incrementally advertises more routes, increasing the number by the amount defined by "Route Step".
5. Repeat step 2 through step 4 until

port 2 receives no packets or packet loss is above the “Tolerance” level.

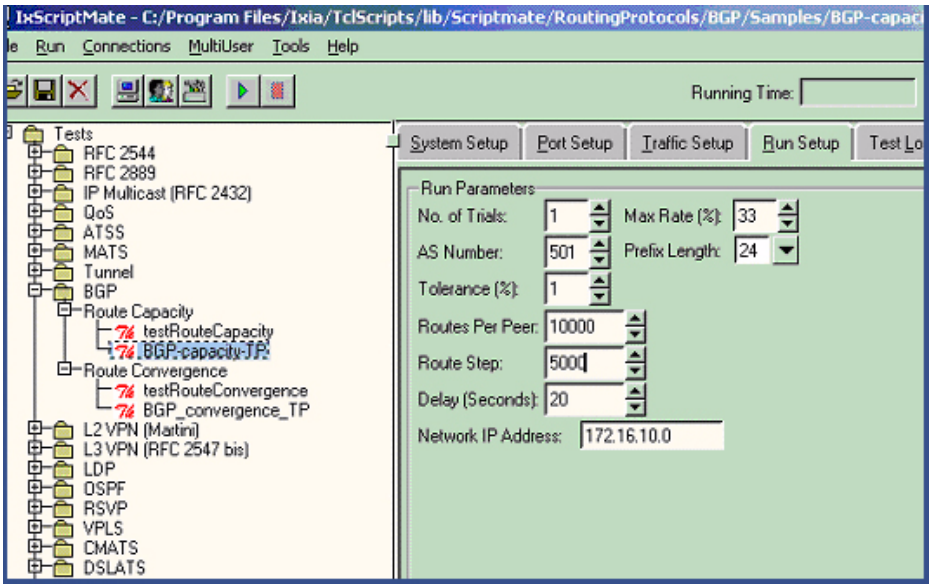

Figure 13. BGP route capacity test example configuration.

Results. When the test completes and the tolerance has been exceeded, the test results will show the maximum number of routes learned by the DUT. Figure 14 shows an example results page created by IxScriptMate. The results are broken down per frame size and show results for “Max Routes Verified”, “Total Loss Percentage”, and “Tolerance”. The “Max Routes Verified” value shows the maximum number of routes that could be sustained at that particular traffic rate and frame size. This test can be executed manually with Ixia’s IxExplorer application, but automation with IxScriptMate helps to simplify and speed the testing process.

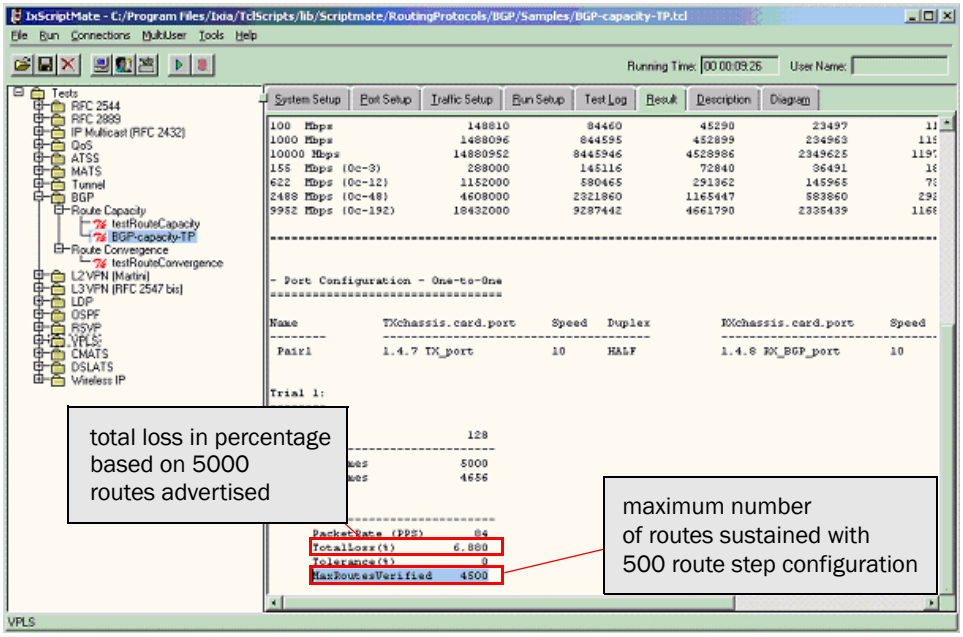

Figure 14. BGP route capacity test results.

3. BGP Route Convergence Test

Objective. Verifies the ability of a router to switch between preferred and less-preferred routes when the preferred routes are withdrawn and re-advertised. The test calculates convergence by taking an average convergence latency of multiple topological changes.

Setup. This test uses three test ports—one to transmit and two to receive (see Figure 15). Both receive ports emulate BGP networks. The transmit direction of traffic is unidirectional. The DUT must have three ports utilized with two enabled for BGP. All three ports should be configured for IP and have unique subnets in which to communicate with the tester ports. Ixia's IxScriptMate application can be used to configure, control, and execute this test.

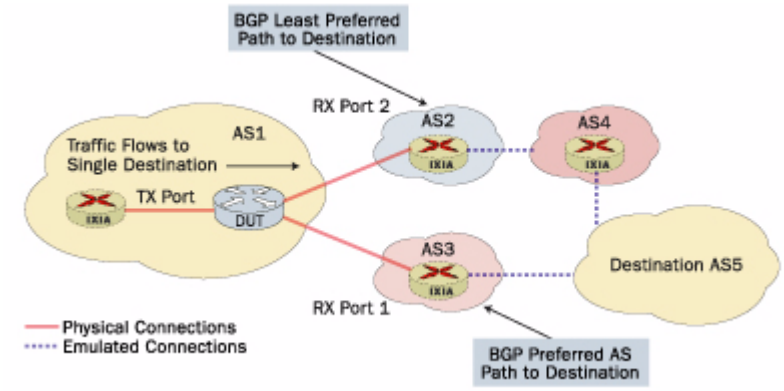

Figure 15. BGP convergence test topology.

Input parameters. See Table 4.

Table 4. BGP convergence test input parameters.

Parameter	Description
Max Rate	The rate at which frames are transmitted.
Routes Per Peer	The number of route prefixes to generate at the start of the test per peer.
Delay	Maximum time in seconds the router is allowed to absorb the advertised routes.
Advertise Delay Per Route	The maximum time, in seconds, to allow the router to absorb each route.

Methodology. The key to determining an accurate convergence time is in understanding the DUT's capabilities and manipulating the test parameters properly (see Figure 16). This methodology can be executed manually or by script:

1. RX ports 1 and 2 advertise the same BGP prefixes with one path preferred with a lower AS-Path count. The path via RX port 1 is used as the preferred route, while the path via RX port 2 is used as the alternate route.
2. After waiting an amount of time indicated by "Delay", the TX port sends one packet to each advertised route. The DUT should route the traffic via the preferred AS-Path to RX port 1.
3. Routes are withdrawn from test RX port 1 (the preferred path). Traffic should reroute to arrive at test RX port 2 (the alternate path).
4. The number of packets lost or transmitted in the incorrect direction is measured after the routes are withdrawn for each route. The packet

loss is converted to time.

5. Repeat step 3 and 4 to obtain convergence time results for all withdrawn routes. Calculate the average convergence across all routes.

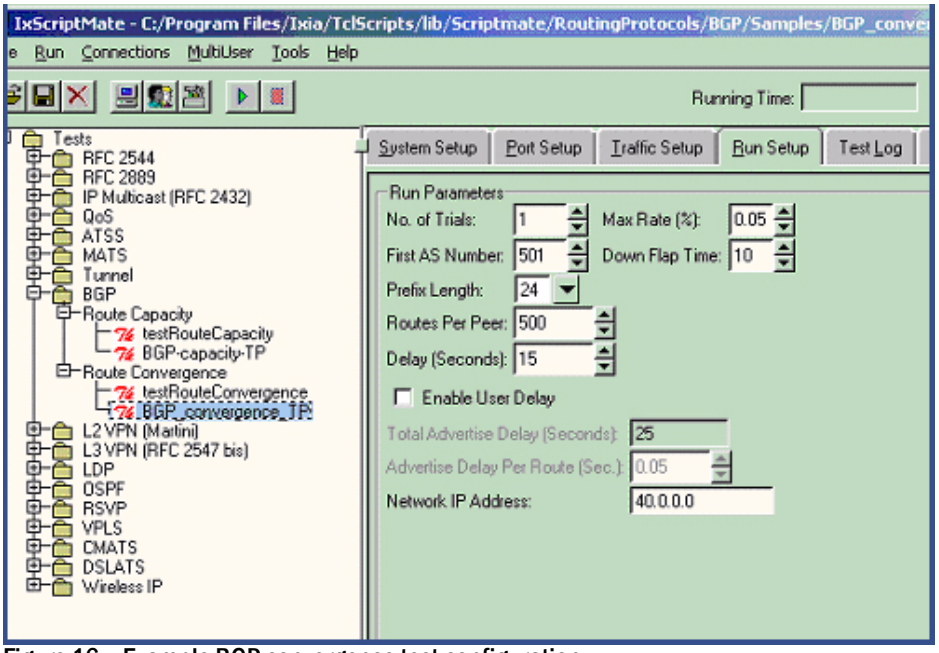

Figure 16. Example BGP convergence test configuration.

Results. The test results provide an average convergence time for all routes. Figure 17 displays example results for the automated BGP convergence test in IxScriptMate. In addition to convergence time, this test also indicates the amount of lost packets caused by the controlled flap in BGP.

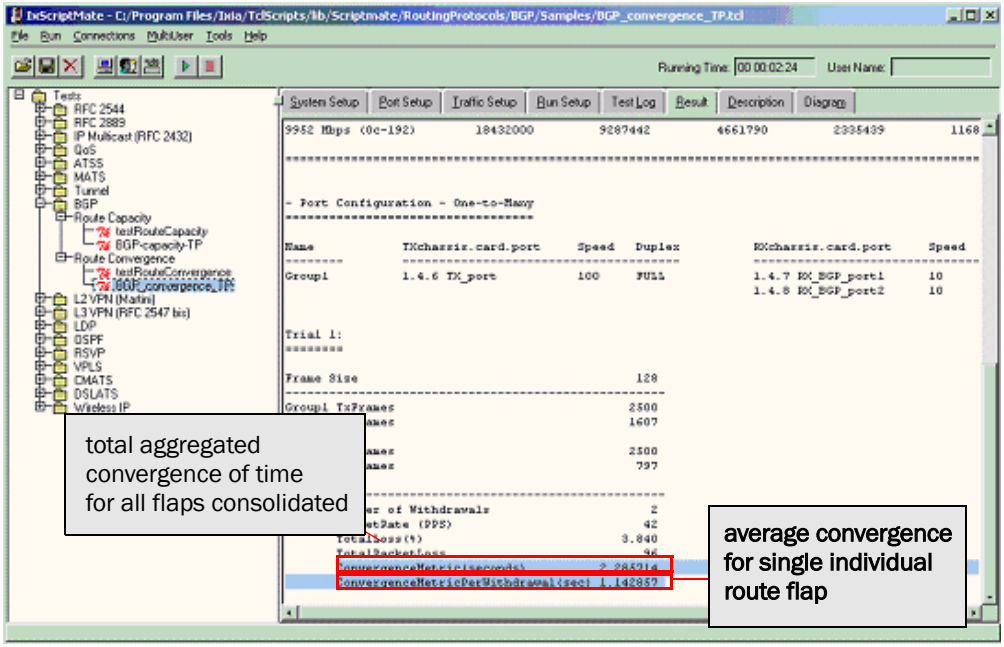

Figure 17. BGP route convergence test results.

4. BGP Damping Test

Objective. This test verifies a router's policy for BGP damping of unstable routes. This test confirms the BGP policy on a customer-specific basis, and is tailored to exact prefixes for perfect accuracy.

Setup. This test requires two test ports—one to transmit and one to emulate BGP routes with flapping capabilities. At least two prefixes are advertised, one which is stable and the other which is unstable. The device's BGP damping policy is tested for proper suppression.

Note: This test is a good verification for ISP and enterprise organizations contending with BGP damping policies affecting their networks.

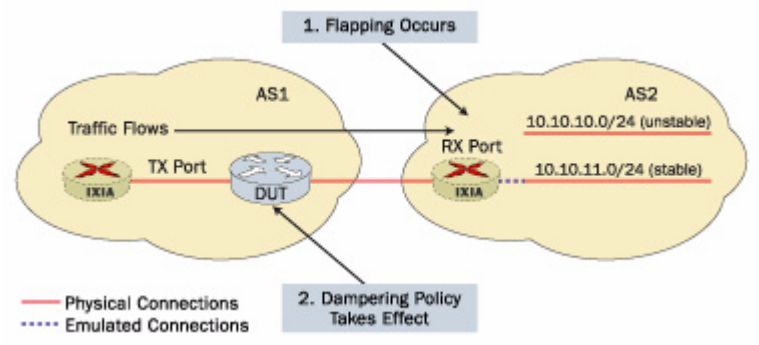

Figure 18. BGP damping test topology.

Input parameters. See Table 5.

Table 5. BGP damping test parameters.

Parameter	Description
Penalty	Increasing number assigned to a route every time it flaps.
Half Life	Amount of time that must pass to reduce the Penalty by one half.
Suppress Limit	Numeric number compared to the Penalty. If the Penalty is larger than the Suppress Limit, the route is suppressed.
Reuse Limit	Numeric number compared to the Penalty. If the Penalty is less than the Reuse Limit, the route will be unsuppressed.

Methodology . damping testing consists of the following steps:

1. Configure the BGP test port to advertise two different prefixes, one that is stable and the other to represent unstable (flapping) behavior.
2. Configure the damping policy on the DUT. The object is to test timed parameters such as "Penalty" and "Half Life" for route damping.
3. Configure the tester transmit port with two traffic flows. The first flow transmits at a configured rate destined to the stable prefix 1. The second slow transmits to the unstable flapping route of destination prefix 2. The stable flow is established for comparison during the damping.
4. Bring up the EBGP session on the DUT and verify the establishment of both advertised prefixes in its IP forwarding table.
5. Initiate the flapping sequence to start the damping process. Ensure that the flaps occur shortly after the "Penalty" expiration. This should continue until the route is suppressed. Figure 19 shows a flapping configuration using Ixia's IxExplorer BGP emulation.

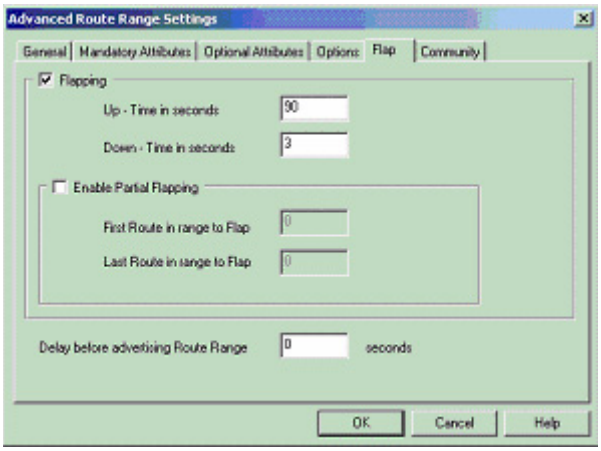

Figure 19. BGP flapping configuration example.

Results. This test results in a display of traffic received on the BGP test port. The results should reflect the effect of route flaps on received traffic. Figure 20 is an example of test results that show the damping policy being applied right after the first 5-second flap. Two more flaps then take place, which result in the traffic shown by the green graph being suppressed. The red graph compares traffic being received on the other stable prefix.

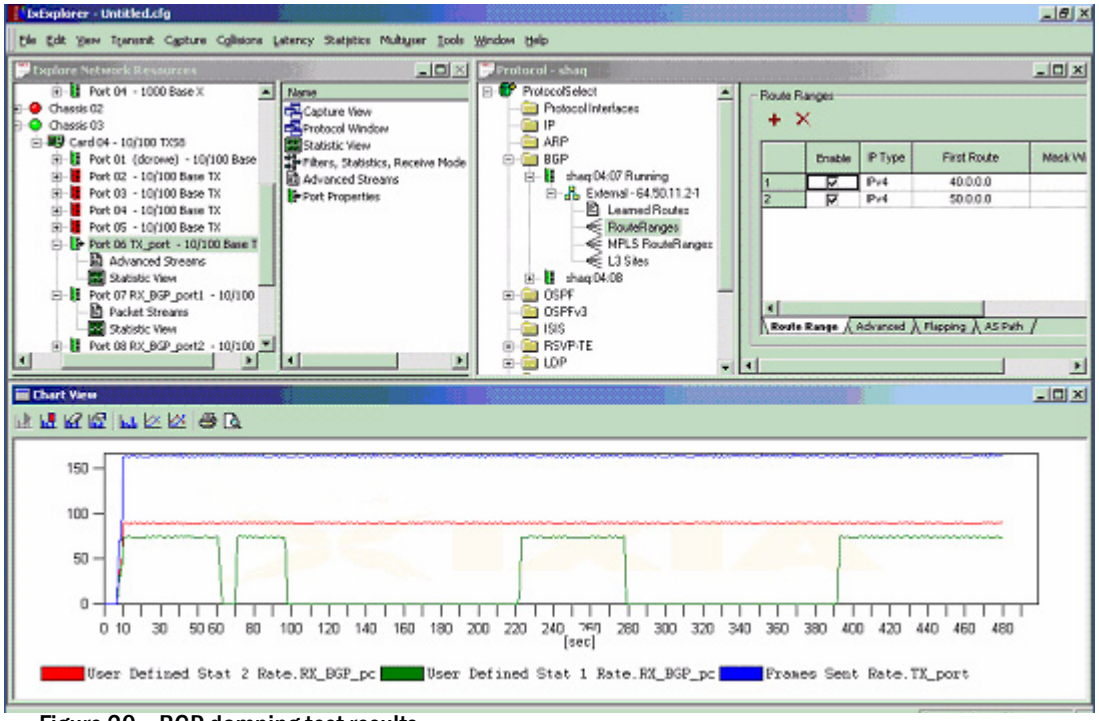

Figure 20. BGP damping test results.

5. BGP Graceful Restart Test

Objective. This test is used for verification of BGP graceful restart capabilities. The test verifies graceful restart functionality using traffic flows. The flows are received on a receive port when neighbor flapping is introduced.

Setup. This test requires a minimum of two ports—one to transmit and the other to receive and represent a BGP neighbor adjacency. The BGP peer will advertise graceful restart capabilities to include restart and stale timers. The path advertised by the neighbor peer will contain a single prefix to represent an IP destination. As traffic enters the DUT, it will flow to the BGP receive tester port. The test then introduces the neighbor flap and traffic continuity is verified. Ixia's IxExplorer BGP emulation can be used to produce timed neighbor flaps and support of extended capabilities for graceful restart.

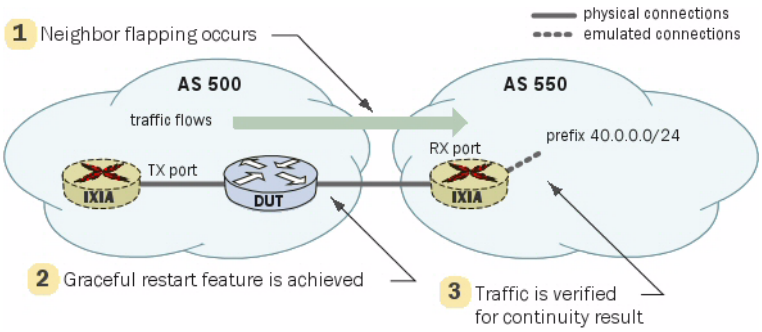

Figure 21. BGP graceful restart test topology.

Input parameters. See Table 6.

Table 6. BGP graceful restart test parameters.

Parameter	Description
Restart Time	The estimated time, following a restart operation, allowed re-establishing a BGP session.
Stale Path Time	The maximum time to maintain stale paths of a gracefully restarted peer. All stale paths are deleted after the expiration of this value.

Methodology. Testing graceful restart consists of the following tasks:

1. Test port 1 emulates the control plane for BGP and establishes an adjacency to the device being tested. The BGP graceful restart function is configured on both the tester and DUT ports.
2. Confirm the DUT forwarding table has learned the received route via BGP. This route receives the traffic and represents the IP destination.
3. Configure test port 2 with a single traffic flow sending to the advertised BGP prefix.
4. Using a test tool, construct a graph or statistical view showing either frames received, frames received rate, and/or peer up/down status. Begin traffic flows to the IP destination.
5. Produce the neighbor flap in BGP from the tester port maintaining the state machine.

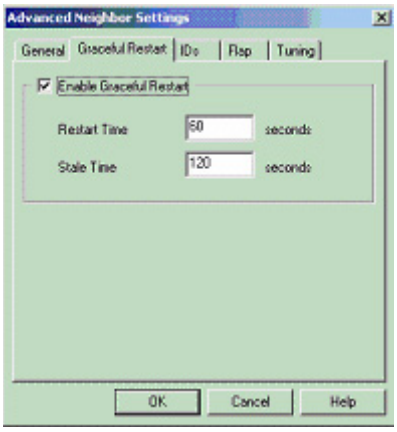

Figure 22. Ixia graceful restart configuration.

Results. The primary goal of this test is to verify traffic flows in BGP even though adjacencies flap. The key is in the proper measurement of the time that traffic continues to flow (the "Restart Time"). Traffic should then cease after the "Stale Path Time" is reached and the peer is dropped or notification sent. Figure 23 shows a graph representing both the received rate of traffic and peer flapping.

Figure 23. BGP graceful restart statistical results.

Glossary

Autonomous System (AS) A set of IP networks under control of a single technical administration.

To the outside world, an AS appears to be a single entity. It uses one or many Interior Gateway Protocols (IGP) and shared metrics for intra-AS routing, and uses one Exterior Gateway Protocol (EGP)—such as BGP—to exchange routing information with other Autonomous Systems.

Each AS has a unique 16-bit integer AS Number, allocated by the same authorities that assign IP addresses. Numbers ranging from 1 to 64511 are public numbers. Numbers from 64512 to 65535 are private, for internal use within organizations, thus should never appear in any Internet routing table.

Classless Inter-domain routing (CIDR) Technique that applies route aggregation mechanisms to reduce the size of the Internet routing tables. Also called “supernetting”, CIDR represents IP addresses with common high-order bits by using shortened subnet masks. For routing purposes, only bits covered by the subnet mask are used, thus making all aggregated addresses look like members of a single, larger network.

Confederation Group of ASs that looks to outside routers as if it were a single AS, with a regular ASN. Described in RFC 1965, this BGP extension allows aggregation of many ASs within a bigger confederation-AS, as well as the possibility to subdivide an AS.

Damping Route flap damping minimizes the instability caused by route flapping by suppressing the propagation of unstable BGP routes. Its goal is to reduce router-processing overhead due to instability without sacrificing convergence time for stable routes. It uses a penalty metric that is incremented every time a route flaps, and that is decremented over time at a given rate. There is a threshold above which a route is suppressed, and a threshold below which a suppressed route is reclaimed.

Exterior BGP (E-BGP) BGP used for communication between router peers from different ASs.

Interior BGP (I-BGP) BGP used for communication between router peers within an AS.

IS-IS An OSI/IP routing protocol, IS-IS stands for Intermediate System to Intermediate System (i.e., router to router).

Multi-Exit Discriminator (MED) Also known as the external metric attribute of a route, it provides information about which path should be selected by the external neighbors accessing an AS with various entry points. Such information is only suggestion, as external neighbors might use other BGP attributes for route selection.

From a backbone-provider perspective, the MED suggests which of their exits to the target AS they should use.

Although MED is not always compared, generally a lower MED value is favored.

Neighbor A pair of BGP-speaking routers that form a TCP connection to exchange routing information between them are called BGP neighbors or peers.

Next Hop Next hop node is the node to send data packets in order to get them closer to the destination. The next hop attribute is used for that purpose.

Open Shortest Path First (OSPF) A link-state routing protocol used by IP routers located within a single Autonomous System (AS) to determine routing paths. MPLS traffic engineering parameters can be distributed with OSPF using extensions to the protocol (OSPF-TE).

Path Vector (PV) Algorithm The PV routing algorithm supplements the advertisement of reachable destinations with information that describes various properties of the paths to these destinations. A path is the recorded sequence of AS numbers through which the reachability information has passed. Each AS is considered equal, independently of its size and internal composition.

A route is defined by the tight coupling of the path to a destination and its attributes, instead of the single distance metric used by Bellman-Ford and other traditional distance-vector algorithms.

Different autonomous domains can have different route optimality notions, as PV only standardizes the results of route selection while allowing heterogeneous criteria across domains. Each AS can have its own policies for route selection.

The Path Vector Algorithm is described in RFC 1322.

Prefix Set of contiguous bits, from 0 to the length of an address, representing all addresses that start with such set of preceding bits. It condenses a (usually large) number of addresses in a compact format.

The prefix attribute of a route determines a section of IP space. For example, IPv4 Class B networks (also known as /16 networks) have a 16-bit network prefix followed by a 16-bit host number. The two highest order bits are set to 1-0 with a 14-bit network number completing the network-prefix (i.e., 128.0.x.x to 191.255.x.x).

Routing Information Protocol (RIP) An Internet routing protocol that uses hop count as a routing metric. RIP is the most common routing protocol among internal routers within a network.

Route Flap Rapid succession of a route advertisement and withdrawal, or withdrawal and re-advertisement.

Route Reflector Concentration router acting as a focal point for I-BGP sessions, adding a hierarchy level to I-BGP. It is primarily recommended for ASs with large internal meshes, and is not recommended for every topology. It is described in RFC 1996 and RFC 2842.

Traffic Engineering Techniques and processes that optimize the routing of network traffic. Traffic engineering mechanisms enable network administrators to manage network traffic's bandwidth, delay, and congestion.

Bibliography

[1] BGP4: Inter-Domain Routing in the Internet, by John W. Stewart III. Publisher: Addison Wesley. (ISBN: 0-20137-951-1)

[2] Interconnections: Bridges and Routers Interconnections: Bridges and Routers, by Radia Perlman (ISBN: 0-20156-332-0)

[3] Internet Routing Architectures, second edition, by Sam Halabi. Publisher: Cisco Press. (ISBN: 157870233X)

[4] Internetworking Technologies Handbook, Fourth Edition. Publisher: Cisco Press, (2003). (ISBN: 1587051192)

[5] TCP/IP Illustrated, Volume 1: The Protocols (Chapter 10. Dynamic Routing Protocols), by W. Richard Stevens. Publisher: Addison Wesley. (ISBN: 0201633469)

[6] RFC 1771: A Border Gateway Protocol 4 (BGP-4)

[7] RFC 1772: Application of the Border Gateway Protocol in the Internet

[8] RFC 1773: Experience with the BGP-4 protocol

[9] RFC 1774: BGP-4 Protocol Analysis

[10] RFC 1322: A Unified Approach to Inter-Domain Routing

[11] RFC 1657: Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2

[12] RFC 2439: BGP Route Flap Damping

[13] RFC 2796: BGP Route Reflection - An Alternative to Full Mesh I-BGP

Acknowledgements

Authors: Dr. Diego Dugatkin and Dennis Crowe.