INTERNET-DRAFT
Danny McPherson
Arbor Networks
Keyur Patel
Cisco Systems
Category
Informational
Expires: March 2005
September 2004

# Experience with the BGP-4 Protocol
<draft-ietf-idr-bgp4-experience-protocol-05.txt>

## Status of this Document



## Copyright Notice

Abstract

The purpose of this memo is to document how the requirements for advancing a routing protocol from Draft Standard to full Standard have been satisfied by Border Gateway Protocol version 4 (BGP-4).

This report satisfies the requirement for "the second report", as described in Section 6.0 of RFC 1264. In order to fulfill the requirement, this report augments RFC 1773 and describes additional knowledge and understanding gained in the time between when the protocol was made a Draft Standard and when it was submitted for Standard.

Table of Contents

## 1. Introduction

The purpose of this memo is to document how the requirements for advancing a routing protocol from Draft Standard to full Standard have been satisfied by Border Gateway Protocol version 4 (BGP-4).

This report satisfies the requirement for "the second report", as described in Section 6.0 of RFC 1264. In order to fulfill the requirement, this report augments RFC 1773 and describes additional knowledge and understanding gained in the time between when the protocol was made a Draft Standard and when it was submitted for Standard.

## 2. BGP-4 Overview

BGP is an inter-autonomous system routing protocol designed for TCP/IP internets. The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. This network reachability information includes information on the list of Autonomous Systems (ASs) that reachability information traverses. This information is sufficient to construct a graph of AS connectivity for this reachability from which routing loops may be pruned and some policy decisions at the AS level may be enforced.

The initial version of the BGP protocol was published in RFC 1105. Since then BGP Versions 2, 3, and 4 have been developed and are specified in [RFC 1163], [RFC 1267], and [RFC 1771], respectively. Changes since BGP-4 went to Draft Standard [RFC 1771] are listed in Appendix N of [BGP4].

### 2.1. A Border Gateway Protocol

The Initial Version of BGP protocol was published in [RFC 1105]. BGP version 2 is defined in [RFC 1163]. BGP version 3 is defined in [RFC 1267]. BGP version 4 is defined in [RFC 1771] and [BGP4]. Appendices A, B, C, and D of [BGP4] provide summaries of the changes between each iteration of the BGP specification.

## 3. Management Information Base (MIB)

The BGP-4 Management Information Base (MIB) has been published [BGP-MIB]. The MIB was updated from previous versions documented in [RFC 1657] and [RFC 1269], respectively.

Apart from a few system variables, the BGP MIB is broken into two tables: the BGP Peer Table and the BGP Received Path Attribute Table.

The Peer Table reflects information about BGP peer connections, such as their state and current activity. The Received Path Attribute Table contains all attributes received from all peers before local routing policy has been applied. The actual attributes used in determining a route are a subset of the received attribute table.

## 4. Implementation Information

There are numerous independent interoperable implementations of BGP currently available. Although the previous version of this report provided an overview of the implementations currently used in the operational Internet, at this time it has been suggested that a separate BGP Implementation Report [BGP-IMPL] be generated.

It should be noted that implementation experience with Cisco's BGP-4 implementation was documented as part of [RFC 1656].

For all additional implementation information please reference [BGP-IMPL].

## 5. Operational Experience

This section discusses operational experience with BGP and BGP-4.

BGP has been used in the production environment since 1989, BGP-4 since 1993. Production use of BGP includes utilization of all significant features of the protocol. The present production environment, where BGP is used as the inter-autonomous system routing protocol, is highly heterogeneous. In terms of the link bandwidth it varies from 56 Kbps to 10 Gbps. In terms of the actual routers that run BGP, it ranges from a relatively slow performance general purpose CPUs to very high performance RISC network processors, and includes

both special purpose routers and the general purpose workstations running various UNIX derivatives and other operating systems.

In terms of the actual topologies it varies from very sparse to quite dense. The requirement for full-mesh IBGP topologies has been largely remedied by BGP Route Reflection, Autonomous System Confederations for BGP, and often some mix of the two. BGP Route Reflection was initially defined in [RFC 1966] and subsequently updated in [RFC 2796]. Autonomous System Confederations for BGP were initially defined in [RFC 1965] and subsequently updated in [RFC 3065].

At the time of this writing BGP-4 is used as an inter-autonomous system routing protocol between all Internet-attached autonomous systems, with nearly 15k active autonomous systems in the global Internet routing table.

BGP is used both for the exchange of routing information between a transit and a stub autonomous system, and for the exchange of routing information between multiple transit autonomous systems. There is no protocol distinction between sites historically considered "backbones" versus "regional" or "edge" networks.

The full set of exterior routes that is carried by BGP is well over 134,000 aggregate entries, representing several times that number of connected networks. The number of active paths in some service provider core routers exceeds 2.5 million. Native AS path lengths are as long as 10 for some routes, and "padded" path lengths of 25 or more autonomous systems exist.

## 6. TCP Awareness

BGP employs TCP [RFC 793] as it's Transport Layer protocol. As such, all characteristics inherent to TCP are inherited by BGP.

For example, due to TCP's behavior, bandwidth capabilities may not be realized due to TCP's slow start algorithms, and slow-start restarts of connections, etc..

## 7. Metrics

This section discusses different metrics used within the BGP

protocol. BGP has a separate metric parameter for IBGP and EBGP. This allows policy based metrics to overwrite the distance based metrics; allowing each autonomous systems to define their independent policies in Intra-AS as well as Inter-AS. BGP Multi Exit Discriminator (MED) is used as a metric by EBGP peers (i.e., inter-domain) while Local Preference (LOCAL_PREF) is used by IBGP peers (i.e., intra-domain).

## 7.1. MULTI_EXIT_DISC (MED)

BGP version 4 re-defined the old INTER-AS metric as a MULTI_EXIT_ DISC (MED). This value may be used in the tie-breaking process when selecting a preferred path to a given address space, and provides BGP speakers with the capability to convey to a peer AS the optimal entry point into the local AS.

Although the MED was meant to only be used when comparing paths received from different external peers in the same AS, many implementations provide the capability to compare MEDs between different autonomous systems as well.

Though this may seem a fine idea for some configurations, care must be taken when comparing MEDs between different autonomous systems. BGP speakers often derive MED values by obtaining the IGP metric associated with reaching a given BGP NEXT_HOP within the local AS. This allows MEDs to reasonably reflect IGP topologies when advertising routes to peers. While this is fine when comparing MEDs between multiple paths learned from a single adjacent AS, it can result in potentially bad decisions when comparing MEDs between different automomous systems. This is most typically the case when the autonomous systems use different mechanisms to derive IGP metrics, BGP MEDs, or perhaps even use different IGP procotols with vastly contrasting metric spaces.

Another MED deployment consideration involves the impact of aggregation of BGP routing information on MEDs. Aggregates are often generated from multiple locations in an AS in order to accommodate stability, redundancy and other network design goals. When MEDs are derived from IGP metrics associated with said aggregates the MED value advertised to peers can result in very suboptimal routing.

The MED was purposely designed to be a "weak" metric that would only be used late in the best-path decision process. The BGP working group was concerned that any metric specified by a remote operator would only affect routing in a local AS if no other preference was specified. A paramount goal of the design of the MED was to ensure

that peers could not "shed" or "absorb" traffic for networks that they advertise.

### 7.1.1. MEDs and Potatoes

In a situation where traffic flows between a pair of destinations, each connected to two transit networks, each of the transit networks has the choice of either sending the traffic to the closest peering to other transit provider or passing traffic to the peering which advertises the least cost through the other provider. The former method is called "hot potato routing" because like a hot potato held in bare hands, whoever has it tries to get rid of it quickly. Hot potato routing is accomplished by not passing the EGBP learned MED into IBGP. This minimizes transit traffic for the provider routing the traffic. Far less common is "cold potato routing" where the transit provider uses their own transit capacity to get the traffic to the point in the adjacent transit provider advertised as being closest to the destination. Cold potato routing is accomplished by passing the EBGP learned MED into IBGP.

If one transit provider uses hot potato routing and another uses cold potato, traffic between the two tends to be symetric. Depending on the business relationships, if one provider has more capacity or a significantly less congested transit network, then that provider may use cold potato routing. An example of widespread use of cold potato routing was the NSF funded NSFNET backbone and NSF funded regional networks in the mid 1990s.

In some cases a provider may use hot potato routing for some destinations for a given peer AS and cold potato routing for others. An example of this is the different treatment of commercial and research traffic in the NSFNET in the mid 1990s. Then again, this might best be described as 'mashed potato routing', a term which reflects the complexity of router configurations in use at the time.

Seemingly more intuitive references that fall outside the vegetable kingdom refer to cold potato routing as "best exit routing", and hot potato routing as "closest exit routing".

### 7.1.2. Sending MEDs to BGP Peers

[BGP4] allows MEDs received from any EBGP peers by a BGP speaker to

be passed to its IBGP peers. Although advertising MEDs to IBGP peers is not a required behavior, it is a common default. MEDs received from EBGP peers by a BGP speaker SHOULD NOT be sent to other EBGP peers.

Note that many implementations provide a mechanism to derive MED values from IGP metrics in order to allow BGP MED information to reflect the IGP topologies and metrics of the network when propagating information to adjacent autonomous systems.

### 7.1.3. MED of Zero Versus No MED

[BGP4] requires that an implementation must provide a mechanism that allows for MED to be removed. Previously, implementations did not consider a missing MED value to be the same as a MED of zero. [BGP4] now requires that no MED value be equal to a value of zero.

Note that many implementations provide a mechanism to explicitly define a missing MED value as "worst" or less preferable than zero or larger values.

### 7.1.4. MEDs and Temporal Route Selection

Some implementations have hooks to apply temporal behavior in MED-based best path selection. That is, all other things being equal up to MED consideration, preference would be applied to the "oldest" path, without preferring the lower MED value. The reasoning for this is that "older" paths are presumably more stable, and thus more preferable. However, temporal behavior in route selection results in non-deterministic behavior, and as such, may often be undesirable.

## 8. Local Preference

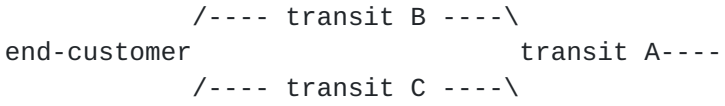
The LOCAL_PREF attribute was added so a network operator could easily configure a policy that overrode the standard best path determination mechanism without independently configuring local preference policy on each router.

One shortcoming in the BGP-4 specification was a suggestion for a

default value of LOCAL_PREF to be assumed if none was provided. Defaults of 0 or the maximum value each have range limitations, so a common default would aid in the interoperation of multi-vendor routers in the same AS (since LOCAL_PREF is a local administration attribute, there is no interoperability drawback across AS boundaries).

[BGP4] requires that LOCAL_PREF be sent to IBGP Peers and must not be sent to EBGP Peers. Although no default value for LOCAL_PREF is defined, the common default value is 100.

Another area where more exploration is required is a method whereby an originating AS may influence the best path selection process. For example, a dual-connected site may select one AS as a primary transit service provider and have one as a backup.




In a topology where the two transit service providers connect to a third provider, the real decision is performed by the third provider and there is no mechanism for indicating a preference should the third provider wish to respect that preference.

A general purpose suggestion that has been brought up is the possibility of carrying an optional vector corresponding to the AS_PATH where each transit AS may indicate a preference value for a given route. Cooperating autonomous systems may then chose traffic based upon comparison of "interesting" portions of this vector according to routing policy.

While protecting a given autonoumous systems routing policy is of paramount concern, avoiding extensive hand configuration of routing policies needs to be examined more carefully in future BGP-like protocols.

## 9. Internal BGP In Large Autonomous Systems

While not strictly a protocol issue, one other concern has been raised by network operators who need to maintain autonomous systems with a large number of peers. Each speaker peering with an external router is responsible for propagating reachability and path

information to all other transit and border routers within that AS. This is typically done by establishing internal BGP connections to all transit and border routers in the local AS.

Note that the number of BGP peers that can be fully meshed depends on a number of factors, to include number of prefixes in the routing system, number of unique path, stability of the system, and perhaps most importantly, implementation efficiency. As a result, although it's difficult to define "a large number of peers", there is always some practical limit.

In a large AS, this leads to a full mesh of TCP connections (n * (n-1)) and some method of configuring and maintaining those connections. BGP does not specify how this information is to be propagated, so alternatives, such as injecting BGP routing information into the local IGP have been attempted, though it turned out to be a non-practical alternative (to say the least).

Several alternatives to a full mesh IBGP have been defined, to include BGP Route Reflection [RFC 2796] and AS Confederations for BGP [RFC 3065], in order to alleviate the the need for "full mesh" IBGP.

## 10. Internet Dynamics

As discussed in [BGP4-ANALYSIS], the driving force in CPU and bandwidth utilization is the dynamic nature of routing in the Internet. As the Internet has grown, the frequency of route changes per second has increased.

We automatically get some level of damping when more specific NLRI is aggregated into larger blocks, however, this isn't sufficient. In Appendix F of [BGP4] are descriptions of damping techniques that should be applied to advertisements. In future specifications of BGP-like protocols, damping methods should be considered for mandatory inclusion in compliant implementations.

BGP Route Flap Damping is defined in [RFC 2439]. BGP Route Flap Damping defines a mechanism to help reduce the amount of routing information passed between BGP peers, and subsequently, the load on these peers, without adversely affecting route convergence time for relatively stable routes.

None of the current implementations of BGP Route Flap Damping store route history by unique NRLI and AS Path although it is listed as mandatory in RFC 2439. A potential result of failure to consider

each AS Path separately is an overly aggressive suppression of destinations in a densely meshed network, with the most severe consequence being suppression of a destination after a single failure. Because the top tier autonomous systems in the Internet are densely meshed, these adverse consequences are observed.

Route changes are announced using BGP UPDATE messages. The greatest overhead in advertising UPDATE messages happens whenever route changes to be announced are inefficiently packed. As discussed in a later section, announcing routing changes sharing common attributes in a single BGP UPDATE message helps save considerable bandwidth and lower processing overhead.

Persistent BGP errors may cause BGP peers to flap persistently if peer dampening is not implemented. This would result in significant CPU utilization. Implementors may find it useful to implement peer dampening to avoid such persistent peer flapping [BGP4].

## 11. BGP Routing Information Bases (RIBs)

[BGP4] states "Any local policy which results in routes being added to an Adj-RIB-Out without also being added to the local BGP speaker's forwarding table, is outside the scope of this document".

However, several well-known implementations do not confirm that Loc-RIB entries were used to populate the forwarding table before installing them in the Adj-RIB-Out. The most common occurrence of this is when routes for a given prefix are presented by more than one protocol and the preferences for the BGP learned route is lower than that of another protocol. As such, the route learned via the other protocol is used to populate the forwarding table.

It may be desirable for an implementation to provide a knob that permits advertisement of "inactive" BGP routes.

It may be also desirable for an implementation to provide a knob that allows a BGP speaker to advertise BGP routes that were not selected by decision process.

## 12. Update Packing

Multiple unfeasible routes can be advertised in a single BGP Update

message. In addition, one or more feasible routes can be advertised in a single Update message so long as all prefixes share a common attribute set.

The BGP4 protocol permits advertisement of multiple prefixes with a common set of path attributes to be advertised in a single update message, this is commonly referred to as "update packing". When possible, update packing is recommended as it provides a mechanism for more efficient behavior in a number of areas, to include:

- Reduction in system overhead due to generation or receipt of fewer Update messages.
- Reduction in network overhead as a result of less packets and lower bandwidth consumption.
- Allows you to process path attributes and look for matching sets in your AS_PATH database (if you have one) less frequently. Consistent ordering of the path attributes allows for ease of matching in the database as you don't have different representations of the same data.

The BGP protocol suggests that withdrawal information should be packed in the begining of Update message, followed by information about more or less specific reachable routes in a single UPDATE message. This helps alleviate excessive route flapping in BGP.

## 13. Limit Rate Updates

The BGP protocol defines different mechanisms to rate limit Update advertisement. The BGP protocol defines MinRouteAdvertisementInterval parameter that determines the minimum time that must be elapse between the advertisement of routes to a particular destination from a single BGP speaker. This value is set on a per BGP peer basis.

Due to the fact that BGP relies on TCP as the Transport protocol, TCP can prevent transmission of data due to empty windows. As a result, multiple Updates may be spaced closer together than orginally queued. Although this is not a common occurrence, implementations should be aware of this.

### 13.1. Consideration of TCP Characteristics

If a TCP receiver is processing input more slowly than the sender or if the TCP connection rate is the limiting factor, a form of backpressure is observed by the TCP sending application. When the TCP buffer fills, the sending application will either block on the write or receive an error on the write. Common errors in either early implementations or an occasional naive new implementation are to either set options to block on the write or set options for non-blocking writes and then treat the errors due to a full buffer as fatal.

Having recognized that full write buffers are to be expected additional implementation pitfalls exist. The application should not attempt to store the TCP stream within the application itself. If the receiver or the TCP connection is persistently slow, then the buffer can grow until memory is exhausted. A BGP implementation is required to send changes to all peers for which the TCP connection is not blocked and is required to remember to send those changes to the remaining peers when the connection becomes unblocked.

If the preferred route for a given NLRI changes multiple times while writes to one or more peers is blocked, only the most recent best route needs to be sent. In this way BGP is work conserving. In times of extremely high route change, a higher volume of route change is sent to those peers which are able to process it more quickly and a lower volume of route change is sent to those peers not able to process the changes as quickly.

For implentations which handle differing peer capacity to absorb route change well, if the majority of route change is contributed by a subset of unstable NRLI, the only impact on relatively stable NRLI which make an isolated route change is a slower convergence for which convergence time remains bounded regardless of the amount of instability.

## 14. Ordering of Path Attributes

The BGP protocol suggests that BGP speakers sending multiple prefixes per an UPDATE message should sort and order path attributes according to Type Codes. This would help their peers to quickly identify sets of attributes from different update messages which are semantically different.

Implementers may find it useful to order path attributes according to Type Code so that sets of attributes with identical semantics can be more quickly identified.

## 15. AS_SET Sorting

AS_SETs are commonly used in BGP route aggregation. They reduce the size of AS_PATH information by listing AS numbers only once regardless of any number of times it might appear in process of aggregation. AS_SETs are usually sorted in increasing order to facilitate efficient lookups of AS numbers within them. This optimization is entirely optional.

## 16. Control over Version Negotiation

Because pre-BGP-4 route aggregation can't be supported by earlier version of BGP, an implementation that supports versions in addition to BGP-4 should provide the version support on a per-peer basis. At the time of this writing all BGP speakers on the Internet are thought to be running BGP version 4.

## 17. Security Considerations

BGP a provides flexible and extendable mechanism for authentication and security. The mechanism allows to support schemes with various degree of complexity. BGP sessions are authenticated based on the IP address of a peer. In addition, all BGP sessions are authenticated based on the autonomous system number advertised by a peer.

Since BGP runs over TCP and IP, BGP's authentication scheme may be augmented by any authentication or security mechanism provided by either TCP or IP.

### 17.1. TCP MD5 Signature Option

[RFC 2385] defines a way in which the TCP MD5 signature option can be used to validate information transmitted between two peers. This method prevents any third party from injecting information (e.g., a TCP Reset) into the datastream, or modifying the routing information carried between two BGP peers.

TCP MD5 is not ubiquitously deployed at the moment, especially in inter- domain scenarios, largely because of key distribution issues. Most key distribution mechanisms are considered to be too "heavy" at this point.

It was naively assumed by many for some time that in order to inject a data segement or reset a TCP transport connection between two BGP peers an attacker must correctly guess the exact TCP sequence number (of course, in addition to source and destination ports and IP addresses). However, it has recently been observed and openly discussed that the malicous data only needs to fall within the TCP receive window, which may be quite large, thereby significantly lowering the complexity of such an attack.

As such, it is recommended that the MD5 TCP Signature Option be employed to protect BGP from session resets and malicious data injection.

### 17.2. BGP Over IPSEC

BGP can run over IPSEC, either in a tunnel, or in transport mode, where the TCP portion of the IP packet is encrypted. This not only prevents random insertion of information into the data stream between two BGP peers, it also prevents an attacker from learning the data which is being exchanged between the peers.

IPSEC does, however, offer several options for exchanging session keys, which may be useful on inter-domain configurations. These options are being explored in many deployments, although no definitive solution has been reached on the issue of key exchange for BGP in IPSEC.

It should be noted that since BGP runs over TCP and IP, BGP is vulnerable to the same denial of service or authentication attacks that are present in any other TCP based protocol.

### 17.3. Miscellaneous

Another issue any routing protocol faces is providing evidence of the validity and authority of the routing information carried within the routing system. This is currently the focus of several efforts at the moment, including efforts to define the threats which can be used against this routing information in BGP [draft-murphy, attack tree], and efforts at developing a means to provide validation and authority for routing information carried within BGP [SBGP] [soBGP].

In addition, the Routing Protocol Security Requirements (RPSEC) working group has been chartered within the Routing Area of the IETF in order to discuss and assist in addressing issues surrounding routing protocol security. It is the intent that this work within RPSEC will result in feedback to BGPv4 and future enhancements to the protocol where appropriate.

## 18. PTOMAINE and GROW

The Prefix Taxonomy (PTOMAINE) working group, recently replaced by the Global Routing Operations (GROW) working group, is chartered to consider and measure the problem of routing table growth, the effects of the interactions between interior and exterior routing protocols, and the effect of address allocation policies and practices on the global routing system. Finally, where appropriate, GROW will also document the operational aspects of measurement, policy, security and VPN infrastructures.

One such item GROW is currently studying is the effects of route aggregation and the inability to aggregate over multiple provider boundaries due to inadequate provider coordination.

It is the intent that this work within GROW will result in feedback to BGPv4 and future enhancements to the protocol as necessary.

## 19. Internet Routing Registries (IRRs)

Many organizations register their routing policy and prefix origination in the various distributed databases of the Internet Routing Registry. These databases provide access to the information using the RPSL language as defined in [RFC 2622]. While registered

information may be maintained and correct for certain providers, the lack of timely or correct data in the various IRR databases has prevented wide-spread use of this resource.

## 20. Regional Internet Registries (RIRs) and IRRs, A Bit of History

The NSFNET program used EGP and then BGP to provide external routing information. It was the NSF policy of offering differing pricing and providing a different level of support to the Research and Education (RE) networks and the Commercial (CO) networks that led to BGP's initial policy requirements. CO networks were not able to use the NSFNET backbone to reach other CO networks, in addition to being charged more. The rationale was that commercial users of the NSFNET with business with research entities should subsidize the RE community. Recognition that the Internet was evolving away from a hierarchical network to a mesh of peers led to changes from EGP and BGP-1 that eliminated any assumptions of hierarchy.

Enforcement of NSF policy was accomplished through maintenance of the NSF Policy Routing Database (PRDB). The PRDB not only contained each networks designation as CO or RE, but also contained a list of the preferred exit points to the NSFNET to reach each network. This was the basis for setting what would later be called BGP LOCAL_PREF on the NSFNET. Tools provided with the PRDB generated complete router configurations for the NSFNET.

Use of the PRDB had the fortunate consequence of greatly improving reliability of the NSFNET relative to peer networks of the time and offering more optimal routing for those networks sufficiently knowledgeable and willing to keep their entries current.

With the decommission of the NSFNET Backbone Network Service in 1995, it was recognized that the PRDB should be made less single provider centric and its legacy contents plus any further updates made available to any provider willing to make use of it. The European networking community had long seen the PRDB as too US centric. Through Reseaux IP Europeens (RIPE) the Europeans had created an open format in RIPE-181 and had been maintaining an open database used for address and AS registry more than policy. The initial conversion of the PRDB was to RIPE-181 format and tools were converted to make use of this format. The collection of databases was termed the Internet Routing Registry, with the RIPE database and US NSF funded Routing Arbitrator (RA) being the inital components of the IRR.

A need to extend RIPE-181 was recognized and RIPE agreed to allow the

extensions to be defined within the IETF in the RPS WG. The result was the RPSL language. Other work products of the RPS WG provided an authentication framework and means to widely distribute the database in a controlled manner and synchronize the many repositories. Freely available tools were provided primarily by RIPE, Merit, and ISI, the most comprehensive set from ISI. The efforts of the IRR participants has been severely hampered by providers unwilling to keep information in the IRR up to date. The larger of these providers have been vocal, claiming that the database entry, simple as it may be, are an administrative burden and some acknowledge that doing so provides a advantage to competitors that use the IRR. The result has been an erosion of the usefulness of the IRR and an increase in vulnerability of the Internet to routing based attack or accidental injection of faulty routing information.

There have been numerous cases of accidental disruption of Internet routing which were avoided by providers using the IRR but highly detrimental to non-users. As filters have had to be relaxed due to the erosion of the IRR to less complete coverage, these types of disruptions have continued to occur very infrequently, but have had increasingly widespread impact.

## 21. Acknowledgements

We would like to thank Paul Traina and Yakov Rekhter for authoring previous versions of this document and providing valuable input on this update as well. We would also like to explicitly acknowledge Curtis Villamizar for providing both text and thorough reviews. Thanks to Russ White, Jeffrey Haas, Sean Mentzer, Mitchell Erblich and Jude Ballard for supplying their usual keen eye.

Finally, we'd like to think the IDR WG for general and specific input that contributed to this document.

## 22. References

### 22.1. Normative References

[RFC 1519] Fuller, V., Li. T., Yu J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", RFC 1519, September 1993.

[RFC 1966] Bates, T., Chandra, R., "BGP Route Reflection: An alternative to full mesh IBGP", RFC 1966, June 1996.

[RFC 2385] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, August 1998.

[RFC 2439] Villamizar, C. and Chandra, R., "BGP Route Flap Damping", RFC 2439, November 1998.

[RFC 2796] Bates, T., Chandra, R., and Chen, E, "Route Reflection - An Alternative to Full Mesh IBGP", RFC 2796, April 2000.

[RFC 3065] Traina, P., McPherson, D., and Scudder, J, "Autonomous System Confederations for BGP", RFC 3065, Febuary 2001.

[RFC 3345] McPherson, D., Gill, V., Walton, D., and Retana, A, "BGP Persistent Route Oscillation Condition", RFC 3345, August 2002.

[BGP4-ANALYSIS] "BGP-4 Protocol Analysis", Internet-Draft, Work in Progress.

[BGP4-IMPL] "BGP 4 Implementation Report ", Internet-Draft, Work in Progress.

[BGP4] Rekhter, Y., T. Li., and Hares. S, Editors, "A Border Gateway Protocol 4 (BGP-4)", BGP Draft, Work in Progress.

[RFC 1657] Willis, S., Burruss, J., Chu, J., " Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2", RFC 1657, July 1994.

[SBGP] "Secure BGP", Internet-Draft, Work in Progress.

[soBGP] "Secure Origin BGP", Internet-Draft, Work in Progress.

[RFC 793] Postel, J., "Transmission Control Protocol", RFC 793, September 1981.

## 22.2. Informative References

[RFC 1105] Lougheed, K., and Rekhter, Y, "Border Gateway Protocol BGP", RFC 1105, June 1989.

[RFC 1163] Lougheed, K., and Rekhter, Y, "Border Gateway Protocol BGP", RFC 1105, June 1990.

[RFC 1264] Hinden, R., "Internet Routing Protocol Standardization Criteria", RFC 1264, October 1991.

[RFC 1267] Lougheed, K., and Rekhter, Y, "Border Gateway Protocol 3 (BGP-3)", RFC 1105, October 1991.

[RFC 1269] Willis, S., and Burruss, J., "Definitions of Managed Objects for the Border Gateway Protocol (Version 3)", RFC 1269, October 1991.

[RFC 1656] Traina, P., "BGP-4 Protocol Document Roadmap and Implementation Experience", RFC 1656, July 1994.

[RFC 1771] Rekhter, Y., and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995.

[RFC 1772] Rekhter, Y., and P. Gross, Editors, "Application of the Border Gateway Protocol in the Internet", RFC 1772, March 1995.

[RFC 1773] Traina, P., "Experience with the BGP-4 protocol", RFC 1773, March 1995.

[RFC 2622] C. Alaettinoglu et al., "Routing Policy Specification Language", RFC 2622, June 1999.

## 23. Authors' Addresses

Danny McPherson
Arbor Networks
Email: danny@arbor.net

Keyur Patel
Cisco Systems
Email: keyupate@cisco.com

## Intellectual Property Statement



## Disclaimer of Validity



## Copyright Statement

## Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.