

BGP-4 Protocol Document Roadmap and Implementation Experience

Status of this Memo

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Introduction

Border Gateway Protocol v4 (BGP-4) [1] is an inter-Autonomous System routing protocol. It is built on experience gained with BGP as defined in RFC-1267 [2] and BGP usage in the connected Internet as described in RFC-1268 [3].

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 from which routing loops may be pruned and some policy decisions at the AS level may be enforced.

BGP-4 provides a new set of mechanisms for supporting classless inter-domain routing. These mechanisms include support for advertising an IP prefix and eliminates the concept of network "class" within BGP. BGP-4 also introduces mechanisms which allow aggregation of routes, including aggregation of AS paths. These changes provide support for the proposed supernetting scheme [4].

The management information base has been defined [5] and security considerations are discussed in the protocol definition document [1].

Applicability Statement for BGP-4

BGP-4 is explicitly designed for carrying reachability information between Autonomous Systems. BGP-4 is not intended to replace interior gateway protocols such as OSPF [7] or RIP [6].

Implementations

Four vendors have developed independent implementations at the time of this memo:

ANS (gated)
Europanet
3COM
cisco

The complete interoperability matrix between all known implementations of various versions of BGP is available under separate cover [9].

Implementation Testing

One implementation has been extensively tested in a network designed to mirror the complex connectivity present at many major Internet borders. This network consists of multiple BGP-3 and BGP-4 speakers carrying full routing information injected from Alternet, EBone, Sprint, CERFnet, and cisco. In many cases additional AS adjacencies are simulated via the use of IP over IP tunnels to increase the complexity of the routing topology.

The primary feature of BGP-4 is the ability to carry network reachability information without regard to classfull routing. In addition to canonical routing information, CIDR prefixes (both supernets and subnets) are being injected from IGP information and aggregated using the methods described in BGP-4. AS set aggregation and policy decisions based upon AS sets have been tested.

Secondary extensions incorporated as part of version 4 of this protocol include enhancements to use of the INTER_AS_METRIC (now called MULTI_EXIT_DISC), the addition of a LOCAL_PREF parameter to influence route selection within an AS, and a specified method of damping route fluctuations. All of these features have been tested in at least one implementation.

Observations

All implementations, are able to carry and exchange network reachability information.

Not all implementations are capable of generating aggregate information based upon the existence of more specific routes.

No implementation supports automatic deaggregation (enumeration of all networks in an aggregate block for backwards compatibility with routing protocols that do not carry mask information (e.g. BGP-3)). However, most implementations do allow for statically configured controlled deaggregation for minimal backwards compatibility with non-CIDR capable routers.

At least one implementation capable of running earlier versions of BGP deliberately does not automatically negotiate to earlier versions. Connections to BGP-4 peers must be explicitly configured as such.

Conclusions

The ability to carry and inject natural networks and CIDR supernets is the immediate requirement for BGP-4. The ability to carry subnet information (useful when reassigning parts of class A networks to organizations with different routing policies) is of secondary concern.

The ability to conditionally aggregate routing information may be worked around by injecting static or IGP network information into BGP, or aggregation may be performed by an upstream router that is capable.

Deaggregation is dangerous. It leads to information loss and unless tightly controlled by a manual mechanism, will create a routing information explosion.

Automatic version negotiation is dangerous due to the state-less nature. Given packet losses or spontaneous restarts, it is possible for two BGP peers capable of BGP-4 to negotiate a BGP-3 or BGP-2 connection, which is incapable of carrying super/subnet reachability information and AS set information.

Acknowledgments

The author would like to acknowledge Yakov Rekhter (IBM) and Tony Li (cisco) for their advice, encouragement and insightful comments.

References

- [1] Rekhter, Y., and T. Li, "A Border Gateway Protocol 4 (BGP-4), RFC 1654, cisco Systems, T.J. Watson Research Center, IBM Corp., July 1994.
- [2] Loughheed K., and Y. Rekhter, "A Border Gateway Protocol 3 (BGP-3)", RFC 1267, cisco Systems, T.J. Watson Research Center, IBM Corp., October 1991.
- [3] Gross P., and Y. Rekhter, "Application of the Border Gateway Protocol in the Internet", RFC 1268, T.J. Watson Research Center, IBM Corp., ANS, October 1991.

- [4] Fuller V., Li. T, Yu J., and K. Varadhan, "Supernetting: an Address Assignment and Aggregation Strategy", Work in Progress. [Note: This is an expired draft, and is also referred to in BGP4.6.]
- [5] Willis S., Burruss J., and J. Chu, "Definitions of Managed Objects for the Border Gateway Protocol (Version 4) using SMIV2", RFC 1657, Wellfleet Communications Inc., IBM Corp., July 1994.
- [6] Hedrick, C., "Routing Information Protocol", RFC 1058, Rutgers University, June 1988.
- [7] Moy J., "Open Shortest Path First Routing Protocol (Version 2)", RFC 1583, Proteon, March 1994.
- [8] Varadhan, K., Hares S., and Y. Rekhter, "BGP4/IDRP for IP---OSPF Interaction", Work in Progress, September 1993.
- [9] Li, T., and P. Traina, "BGP Interoperabilty Matrix", Work in Progress, November 1993.

Security Considerations

Security issues are not discussed in this memo.

Author's Address

Paul Traina
cisco Systems, Inc.
1525 O'Brien Drive
Menlo Park, CA 94025

EMail: pst@cisco.com

