Internet Engineering Task Force
INTERNET DRAFT

T. Przygienda
Bell Labs, Lucent Technologies
19 November 1997

BGP-4 over ATM and Proxy PAR
<draft-przygienda-bgp4-atm-00.txt>

Status of This Memo

Abstract

This draft specifes for BGP-4 implementors and users mechanisms describing how the protocol operates in ATM networks over PVC and SVC meshes with the presence of Proxy PAR. These recommendations do not require any protocol changes and allow for simpler, more efficient and cost-effective network designs. Proxy PAR can help to distribute changes of peer relationships when BGP-4 capable routers are reconfigured on the ATM cloud.

1. Introduction

1.1. Introduction to Proxy PAR

Proxy PAR [CPS96, PD97a] is an extension allowing for different ATM attached devices to interact with PAR capable switches and obtain information about non-ATM services without executing PAR [Ca96] which is an extension of PNNI [AF96b] themselves. The client side is much simpler in terms of implementation complexity and memory requirements than a complete PAR protocol stack and should allow for easy implementation in e.g. existing IP routers. Additionally, clients can use Proxy PAR to register different non-ATM services and protocols they support. Proxy PAR has consciously not been included as part of ILMI due to the complexity of PAR information passed in the protocol and the fact that it is intended for integration of non-ATM protocols and services only. A device executing Proxy PAR does not necessarily need to execute ILMI or UNI signaling although this normally will be the case. The context or reference model is aligned with the one included in ILMI [AF96a].

The protocol in itself does not specify how the distributed service registration and data delivered to the client is supposed to be driving other protocols so e.g. OSPF routers finding themselves through proxy PAR could use this information in RFC1577 [Lau94] fashion, forming a full mesh of point-to-point connections to interact with each other to simulate broadcast interfaces. For the same purpose LANE [AF95] or MARS [Arm96] could be used.

As a by-product, Proxy PAR could provide the ATM address resolution for IP attached devices but such resolution can be achieved by other protocols under specification in IETF as well, e.g. [CH97a, CH97b]. And last but not least, it should be mentioned here that the protocol coexists with and complements the ongoing work in IETF on server detection via ILMI extensions [Dav97] and opaque LSAs [CH97a, CH97b].

1.1.1. Proxy PAR Scopes

Any Proxy PAR registration is carried only within a defined scope that is set during registration and is equivalent to the PNNI routing level. Since no assumptions except scope values can be made about the information distributed (e.g. IP addresses bound to NSAPs are not assumed to be aligned with them in any respect such as

encapsulation or functional mapping), registration information cannot be summarized. This makes a careful handling of scopes necessary to preserve the scalability. More detailed comments on these issues and optimizations possible when logical IP topology aligns with aspects of ATM topology can be found in [PD97b].

1.2. Introduction to BGP

Border Gateway Protocol (BGP) is an Exterior Gateway Protocol (EGP) and described in [RL95, RL97] from which most of the following paragraphs have been taken almost literally.

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 runs over a reliable transport protocol. This eliminates the need to implement explicit update fragmentation, retransmission, acknowledgment, and sequencing. Any authentication scheme used by the transport protocol may be used in addition to BGP's own authentication mechanisms. The error notification mechanism used in BGP assumes that the transport protocol supports a "graceful" close, i.e., that all outstanding data will be delivered before the connection is closed.

BGP deployments are normally configured such that that all BGP speakers within a single AS must be fully meshed so that any external routing information must be re-distributed to all other routers within that AS. This represents a serious scaling problem that has been well documented with several alternatives proposed. The alternative supported in Proxy PAR are route reflectors [Bat96] due to their simplicity, easy migration and compatibility with existing BGP configurations.

2. BGP over ATM

2.1. Model

The model used for BGP operation over ATM in connection with Proxy PAR assumes that not only pre-configured peers exist but neighbor relationships can be formed dynamically based on discovery mechanisms. Such a discovery must be provided by an underlying layer since BGP does not include peer auto-detection that would be comparable with e.g. OSPF's hellos used to find all OSPF routers on a specific subnet. To fulfill this purpose, Proxy PAR allows BGP to register and query the following data with the server:

- ATM address
- IP instance
- IP address
- IP mask
- BGP Identifier
- route reflector type as one of:
    * reflector of a certain cluster or
    * client of a certain cluster or
    * non-client

The motivation of such a model is to allow for a simpler maintainance of BGP router configuration when some router interfaces are connected over ATM. As an example, full mesh connectivity on a specific subnet does not require the configuration of peer relationsships in routersa priori but a router can register as providing BGP services on an interface and his possible peers discover it through Proxy PAR queries. Figure 1 illustrates a possible BGP scenario with several cases of relationsships based on the following Proxy PAR registrations:

- Router R1 is configured to be BGP capable and has the interface
    * 1.1.1.1 reaching into DMZ subnet 1.1.1/255.255.255

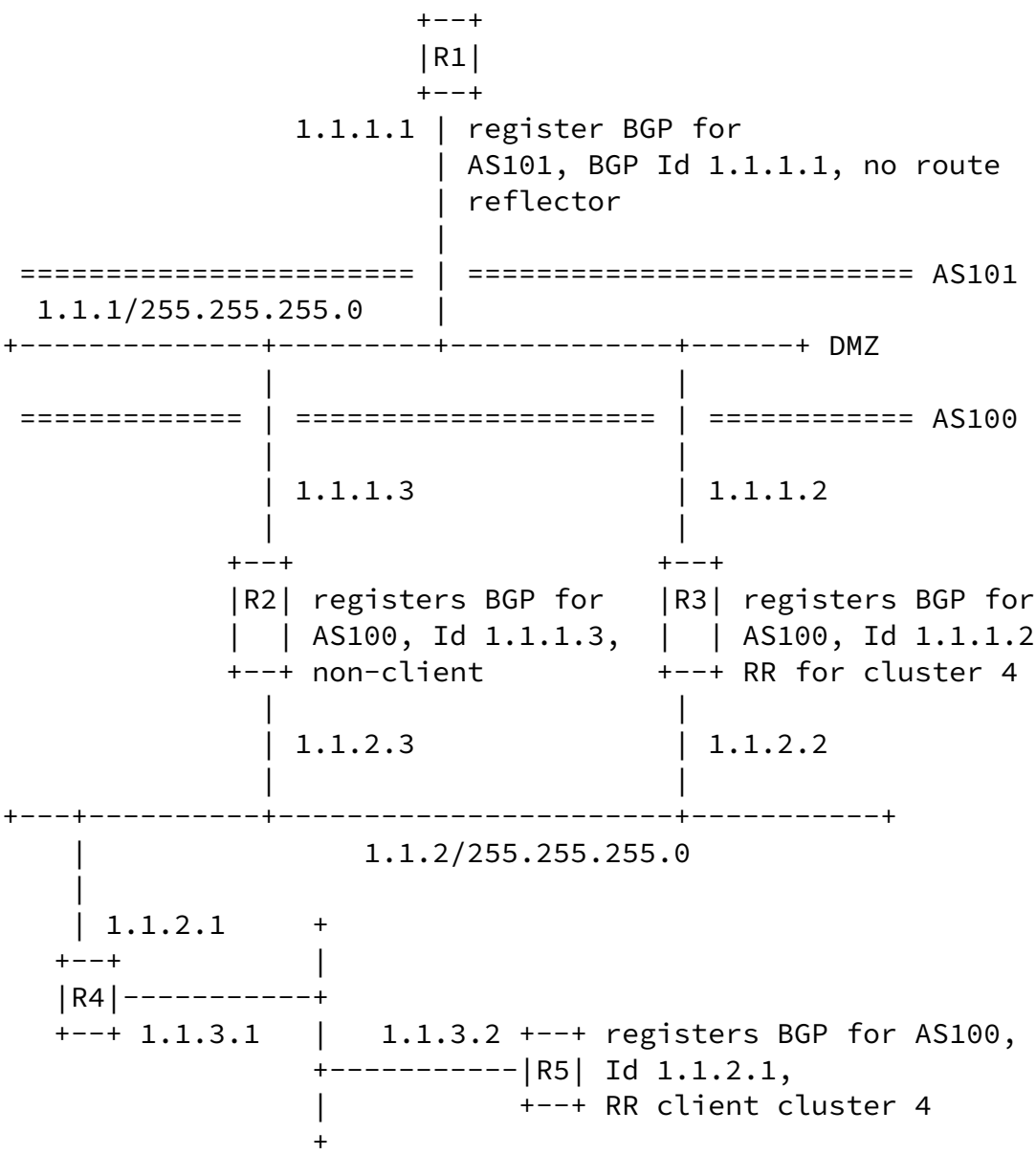

Figure 1: Logical IP Topology with Proxy PAR Registrations (Single ATM network)

- Router R3 is configured to be BGP capable, is route reflector for cluster id 4, and has the interfaces

  * 1.1.1.2 reaching into DMZ subnet 1.1.1/255.255.255 and

  * 1.1.2.2 to the subnet 1.1.2/255.255.255 inside its AS

- Router R2 is configured to be BGP capable and has the interfaces
  - 1.1.1.3 reaching into DMZ subnet 1.1.1/255.255.255 and
  - 1.1.2.3 to the subnet 1.1.2/255.255.255 inside its AS
- Router R5 is configured to be BGP capable, is a client of route reflector cluster id 4, and has the interface
  - 1.1.2.1 to a subnet 1.1.2/255.255.255 inside its AS

It has to be stated here that the model assumes that E-BGP-multihop will not be supported through auto-configuration. Based on such an assumption, the following queries are generated by the routers and conclusions drawn concerning the BGP sessions to be formed:

Q1> Router R1 queries for all BGP capable routers on the DMZ subnet (1) and discovers R2 and R3 supporting interfaces 1.1.1.3 and 1.1.1.2 and being in a different AS. Router R1 concludes to

- build a E-BGP connection to router R3 (shown with &'s in Figure 2)
- build a EBPG connection to router R2 (shown with #'s in Figure 2)

Q2> Router R2 queries for all BGP capable routers on the DMZ subnet and discovers R3 and R1 on the same subnet and concludes to

- build a E-BGP connection to router R1 since it is in a different AS
- not to build a E-BGP connection to router R3 since it is in the same AS

Q3> Router R3 issues a symmetric query to Q2 and comes to conclusions analogous to Q2>

Q4> Router R2 queries for all routers supporting BGP inside of the same AS, detects R3 and R5 and concludes to

---

1. since one of its interfaces is on this subnet

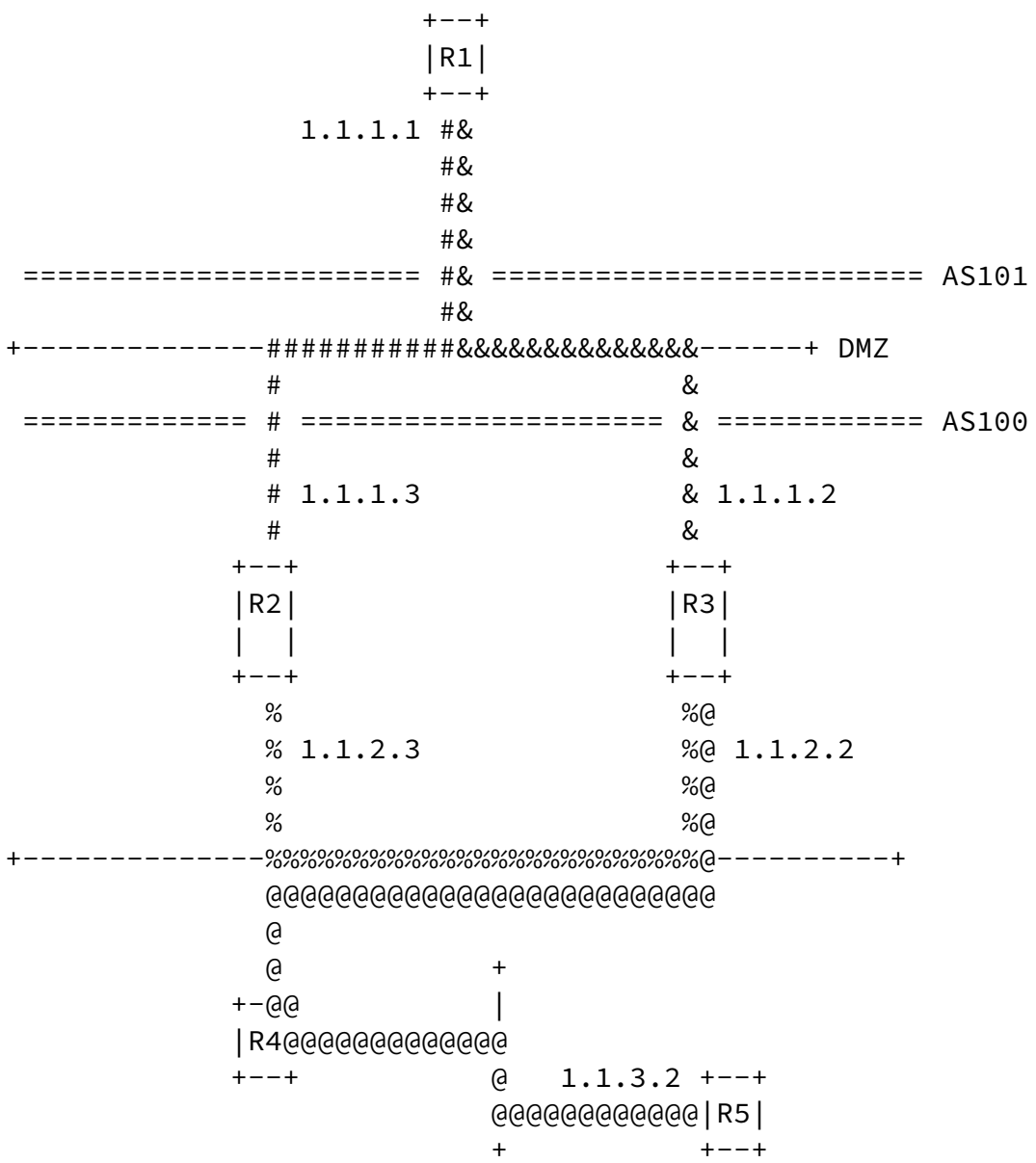

Figure 2: Active BGP Connections after Auto-Discovery in Figure 1.

* build an I-BGP connection to R3 since R3 is a reflector and R2 is a non-client (shown with %'s in Figure 2)

* not build an I-BGP connection to R5 since R5 is a client of a route reflector

Q5> Router R3 queries for all routers supporting BGP inside of the same AS, detects R2 and R5 and concludes to

* build an I-BGP connection to R2 since R3 is a reflector and R2 is a non-client

* build an I-BGP connection to R5 since R5 is a client of the route reflector for the same cluster (shown with @'s in Figure 2)

Q6> Router R5 queries for all routers supporting BGP inside of the same AS, detects R2 and R3 and concludes to

* not build an I-BGP connection to R2 since R5 is a reflector client and R2 is a non-client

* to build an I-BGP connection to R3 since R3 is the reflector for the same cluster R5 is client of

The resulting peerings are visualized in Figure 2. Based on the configuration of BGP properties the network automatically set up valid and necessary connections between routers. It should be obvious that especially for I-BGP such a mechanism faciliates the maintainance of many routers inside of an AS. The necessary route reflector and mesh connections for BGP are built correctly. A carefull reader observes as well that the automatically formed full set of E-BGP connections between AS border routers is not always a good thing. This problem will be given some special consideration.

The intended auto-configuration behavior when registering and retrieving information can be split across the internal and external BGP functionality boundary. Since I-BGP requires a full mesh configuration (2) Proxy PAR information proves very beneficial to meet this necessary constraint in an automatic manner. For E-BGP, as mentioned above, a full mesh between all peers on the same subnet is not always a good solution and therefore Proxy PAR information has to be treated more carefully or not used at all.

---

2. with exceptions in presence of route reflectors, of course

2.2. BGP Configuration Interaction with Proxy PAR

To resolve problems with multiple IP subnets operating on top of a single ATM NSAP, multiple BGP instances, and possibly even multiple ATM clouds the router attaches to, router configuration has to define what information is feasible to be registered. As default, any new upcoming IP interface running on top of an ATM link should be registered with the server on one of the ATM links interfacing with the same ATM cloud. The necessary IP instance is determined by the BGP instance and the NSAP is equivalent to the NSAP of the ATM interface through which the registration is performed.

2.2.1. Registration of Information for Autoconfiguration of External BGP Peerings

An implicit assumption when using Proxy PAR for autoconfiguration of BGP external peerings is that multihop peers are not supported. A BGP router with an IP over ATM interface that attaches to a subnet between different AS'es registers the interface for the according IP instance with one of the proxy PAR servers on the same cloud. It is possible, although not necessary, to omit multiple registrations in the case of a BGP router having multiple interfaces to the same IP subnet with broadcast capabilities.

2.2.2. Registration of Information for Autoconfiguration of Internal BGP Peerings

For the IP over ATM interfaces on subnets being entirely inside of the router's AS, BGP instances should register with proxy PAR server. This allows for necessary sessions to be formed and consecutively provides full mesh connectivity between non-clients, and star connectivity inside route reflector clusters. Same optimizations as described in section 2.2.1 are possible.

2.3. Proxy PAR Interaction with BGP Configuration

2.3.1. Autoconfiguration of Internal BGP Peerings

Proxy PAR presence in a BGP network 'on the internal side' is helping to meet the requirement that all I-BGP peers have to be connected as full-mesh or connect to their route reflectors. To make sure that

all route reflector clients and non-clients are configured correctly, Proxy PAR queries will present enough information to let the routers configure a minimal valid connectivity graph. After being provided with the information about all BGP peers running in the same AS, a BGP router determines which peers it must initiate connections to based on the following criteria:

- looking at the other router's BGP identifier no session has been formed yet and
- the other router is in the same AS and
  * one router is route reflector with the same cluster ID and the other router is a client of this cluster or
  * one of the routers is a non-client

The example in section 2.1 encompasses the different cases that can trigger initiation of connections.

2.3.2. Autoconfiguration of external BGP peerings

Proxy PAR registration information made available can be used to determine which BGP routers are present to form sessions with. Normally, all routers on a specific DMZ subnet are interested in forming relationships with routers in different ASes to exchange route information. However, to prevent unnecessary or insecure external sessions, each of the IP interfaces on a subnet reaching into other AS'es can filter information from query results based on any of the fields or combinations thereof. The filter would prevent BGP from autodetecting the registration and effectively the possible neighbor. Since the connection could be initiated from either side, the filters should be symmetrical in both BGP peers that try to prevent that session from forming. If this is unenforcable, a peer accepting an E-BGP connection for which Proxy PAR information is filtered, could explicitly close it after providing appropriate notification.

2.4. IP to ATM Address Resolution

Given the nature of Proxy PAR registrations that contain not only BGP specific information but always carry IP interface address and

the attached NSAP, when running BGP over IP interfaces on top of ATM with Proxy PAR capabilities, the information obtained in queries can be used to provide address resolution for the lower layers. When BGP chooses to initiate a connection to a peer, lower layers of the TCP/IP protocol stack could use the available Proxy PAR information to resolve the IP address into the necessary NSAP of the registration point. Such a solution however necessitates an appropriate stack architecture.

3. Acknowledgments

Comments and contributions from several sources, especially Rob Coltun are included in this work.

4. Security Consideration

Security issues in the context of BGP autoconfiguration in presence of Proxy PAR can be split into parts specific to either of the protocols. BGP protocol addresses the issues in existing RFCs and ongoing work. PNNI protocol in version 2 contains peer authentication mechanisms and Proxy PAR in itself could be extended to encompass the same security features in the future. To address the problem of security of Proxy PAR client/server interactions, especially registrations that could be used for denial-of-service attacks is an issue not addressed so far. Its scope is similar to the problem of a secure ILMI [AF96a].

5. Conclusions

This RFC specifes for BGP implementors and users mechanisms describing how the protocol operates in ATM networks over PVC and SVC meshes with the presence of Proxy PAR. These recommendations do not require any protocol changes and allow for simpler, more efficient and cost- effective network designs. Proxy PAR can help to distribute configuration changes when BGP capable routers are reconfigured on the ATM cloud and greatly facilitates consistence of I-BGP meshes and can be used for E-BGP auto-configuration as well.

References

[AF95] ATM-Forum. LAN Emulation over ATM 1.0. ATM Forum af-lane-0021.000, January 1995.

[AF96a] ATM-Forum. Interim Local Management Interface (ILMI) Specification 4.0. ATM Forum 95-0417R8, June 1996.

[AF96b] ATM-Forum. Private Network-Network Interface Specification Version 1.0. ATM Forum af-pnni-0055.000, March 1996.

[Arm96] G. Armitage. Support for Multicast over UNI 3.0/3.1 based ATM Networks, RFC 2022. Internet Engineering Task Force, November 1996.

[Bat96] T. Bates. BGP Route Reflection, RFC 1966. Internet Engineering Task Force, June 1996.

[Ca96] R. Callon and al. An Overview of PNNI Augmented Routing. ATM Forum 96-0354, April 1996.

[CH97a] R. Coltun and J. Heinanen. Opaque LSA in OSPF. Internet Draft, 1997.

[CH97b] R. Coltun and J. Heinanen. The OSPF Address Resolution Advertisement Option. Internet Draft, 1997.

[CPS96] R. Coltun, T. Przygienda, and S. Shew. MIPAR: Minimal PNNI Augmented Routing. ATM Forum 96-0838, June 1996.

[Dav97] M. Davison. Simple ILMI-Based Server Discovery. Internet Draft, 1997.

[Lau94] M. Laubach. Classical IP and ARP over ATM, RFC 1577. Internet Engineering Task Force, January 1994.

[PD97a] T. Przygienda and P. Droz. Proxy PAR. ATM Forum 97-0495, 97-0705, 97-0882, July 1997.

[PD97b] T. Przygienda and P. Droz. Proxy PAR. Internet Draft, 1997.

[RL95] Y. Rekhter and T. Li. A Border Gateway Protocol 4 (BGP-4), RFC 1771. Internet Engineering Task Force, March 1995.

[RL97] Y. Rekhter and T. Li. A Border Gateway Protocol 4 (BGP-4). Internet Draft, 1997.

Authors' Addresses

Tony Przygienda
Bell Labs, Lucent Technologies
101 Crawfords Corner Road
Holmdel, NJ 07733-3030
prz@dnrc.bell-labs.com