TELECOM-SP ISP Network Design Issues

Prepared By:**Paul G. Donner (pdonner@cisco.com)**
Consulting Engineer - Latin America Region
Cisco Systems, Inc.
Miami, FL

Date: **Wednesday, April 30, 1998**

Document Revision: **Version 1.31**

Document: **`ISP Policy Implementation Case Study V_1_31.DOC`**

Table of Contents

1.0 IMPORTANT

This document has been developed based on a real network design project. All references to real, existing Internet Service Providers (ISPs) have been changed to reflect fictitious entities. This document is intended to serve as an illustration of how to implement networking policy and architecture for a small to medium-sized ISP. Under no circumstances should this material be implemented in a network until and unless careful analysis of the material is conducted to determine if it is applicable or requires modification.

1.0 Overview

This document proposes architecture as well as routing policy mechanisms which are intended to satisfy TELECOM-SP's operational requirements for interaction with the Internet and several neighboring autonomous systems (AS). The actual implementation varies to some degree from TELECOM-SP's stated requirements in order to better accommodate the functionality of BGP peering sessions and the inherent nature of inter-AS routing asymmetry. However, the main purpose of this document is to present a possible framework which TELECOM-SP personnel can work with to develop an ultimately more effective and efficient policy and architecture for their network. This document is mainly intended to stimulate a process which leads to a good resolution of the requirements at hand.

2.0 Requirements, Assumptions and Relevant Information

2.1 Requirements

The policy desired by TELECOM-SP which will govern the interaction with the aforementioned autonomous systems is understood as shown below (any requirement not explicitly specified in the following policy discussion is NOT considered). Unless otherwise specified, the "Internet" refers to all ASes other than AS921, AS1399, AS5037, AS7341, AS432, and AS1121. These ASes are hereinafter referred to as the "neighboring" ASes.

1. **AS432** will be permitted to send packets to destinations within and receive packets originated from sources within AS7341 and AS7167. This connectivity will utilize link L8 with AS7341 and link L8 and L7 with AS1121. AS7341's primary link to AS432 is considered to be link L8. Secondary connectivity to AS432 can be provided through AS921 and AS1399.

2. **AS5037** will be permitted to send packets to destinations within and receive packets originated from sources within AS7341 and AS7167. This connectivity will utilize link L9 with AS7341 and link L9 and L7 with AS1121. AS7341's primary link to AS5037 is considered to be link L9. Secondary connectivity to AS5037 can be provided through AS921 and AS1399.

3. **AS1121** will be permitted to send packets to destinations within and receive packets originated from sources within AS7341. Further, AS7341 will serve as a transit path for AS1121 to all destinations on the Internet as well as neighboring ASes. All traffic between AS1121 and AS7341 as well as transit traffic between AS1121 and the Internet and AS7341's neighboring ASes, will utilize link L7. AS1121's primary access to the Internet "and" AS7341's neighboring ASes will be across link L7. Should link L7 fail or access to the Internet or any of the neighboring ASes fail (i.e. failure of the path through AS7341), entirely or partially, AS1121 will rely on link L6 for connectivity through AS854.

4. **AS921** will serve as a primary pathway for egress and ingress traffic to/from TELECOM-SP (AS7341) and XPAC (AS1121). A primary default path and a backup default path will be implemented via CIMR. Traffic to/from CIMR to/from ACSNET, INTEX and OBERON will not be permitted (TELECOM-SP will not be used as a transit AS for these other ASes even in the case of a backup). Partial routing will be obtained from CIMR in an effort to reduce routing overhead due to the large number of routes in the full table, processing resulting from multiple path, possible configuration complexity and hardware overhead. A primary and a backup link with AS921 are required for each of the netblocks.

5. **AS1399** will serve as a primary pathway for egress and ingress traffic to/from TELECOM-SP (AS7341) and XPAC (AS1121). A primary default path and a backup default path will be implemented via OBERON. Traffic to/from OBERON to/from ACSNET, INTEX and CIMR will not be permitted (TELECOM-SP will not be used as a transit AS for these other ASes even in the case of a backup). Partial routing will be obtained from CIMR in an effort to reduce routing overhead due to the large number of routes in the full table, processing resulting from multiple path, possible configuration complexity and hardware overhead. A primary and a backup link with AS1399 are required for each of the netblocks.

2.2 General Assumptions

The following assumptions are made based on the information provided by TELECOM-SP:

1. There is no information which details the termination of the multiple TELECOM-SP circuits inside the CIMR and OBERON networks. Therefore, it is assumed that each circuit terminates at a different peer router and thus eBGP Multihop is not a viable option in this case. Also, special consideration should be taken if implementing eBGP multihop in this environment since the link speeds vary widely. In general, eBGP multihop should only be used on links with the same transmission speeds.

2. For the purposes of this document, it is assumed that PACNET is an integral part of TELECOM-SP AS7341 and not as a separate AS (i.e. AS7167).

3. There is no control over XPAC's routers and no knowledge of XPAC's internal connectivity, router configurations or policy.

4. Traffic levels are completely unknown. Breakdown of prefix announcements has been done in an attempt to distribute traffic loading as much as possible across the peering links with CIMR and OBERON. The announcements have been broken into netblocks based primarily on

traffic loading as shown in **Tables 2.3-A, 4.2-A and 4.2-B**, below. Since the traffic levels generated by each netblock are unknown at this time, this will serve merely as an illustration of how this distribution can be accomplished. This mechanism has been implemented mainly because, bandwidth in South American countries is expensive, provisioning can be lengthy and circuits can be extremely costly forcing extreme efficiency on the use of existing circuits.

Breakdown of prefix announcements across different links to CIMR and OBERON is guesswork at this point and serves only as an illustration. TELECOM-SP must analyze their traffic patterns and decide on the correct announcement breakdown for each link/provider. In an attempt to distribute incoming traffic (traffic destined for TELECOM-SP) as flexibly as possible across the combined CIMR and OBERON links, the entire TELECOM-SP/XPAC address space has been sub-divided into multiple netblocks. These netblocks correspond to those which are currently implemented in TELECOM-SP and XPAC. These netblocks are then selectively announced across the various links utilizing different mechanisms to manage preference on a neighbor basis, including MED and Community attributes and on an end-to-end basis using AS_Path prepending. The breakdown of the netblocks is assumed to be done primarily based on traffic load (which, again, is unknown). These netblocks and the preference mechanisms are discussed in section **4.2 Routing Policy Implemented**. It is also assumed that TELECOM-SP wishes to have all traffic destined for netblocks A through D to primarily traverse the CIMR links and as default, traverse the OBERON links.

5. In a lower tier Internet Service Provider (ISP) Ingress traffic from upstream providers is generally of a much higher magnitude than egress traffic (headed upstream). Ratios range from 2:1 to 4:1 based on statistics collected by various Latin American ISPs. This ratio may in fact actually be higher with respect to ingress versus egress traffic.

6. The TELECOM-SP and PACNET are considered to comprise the same Autonomous System in this document.

2.2 Autonomous Systems (ASes) involved

The ASes involved with this design include: (1) CIMR, AS921; (2) OBERON; AS1399; (2) ACSNET, AS5037; (3) TELECOM-SP, AS7341; (4) INTEX, AS432; and (6) XPAC, AS1121.

2.3 Address ranges involved

XPAC	
172.22.0.0/16	obtained from InterNIC
TELECOM-SP	
212.18.22.0/22	obtained from CIMR Address Space
212.18.12.0/21	obtained from CIMR Address Space
201.9.110.0/20	obtained from CIMR Address Space
173.41.220.0/21	obtained from CIMR Address Space
PACNET	
200.5.32.0/19	obtained from InterNIC

200.23.7.0/19 obtained from InterNIC

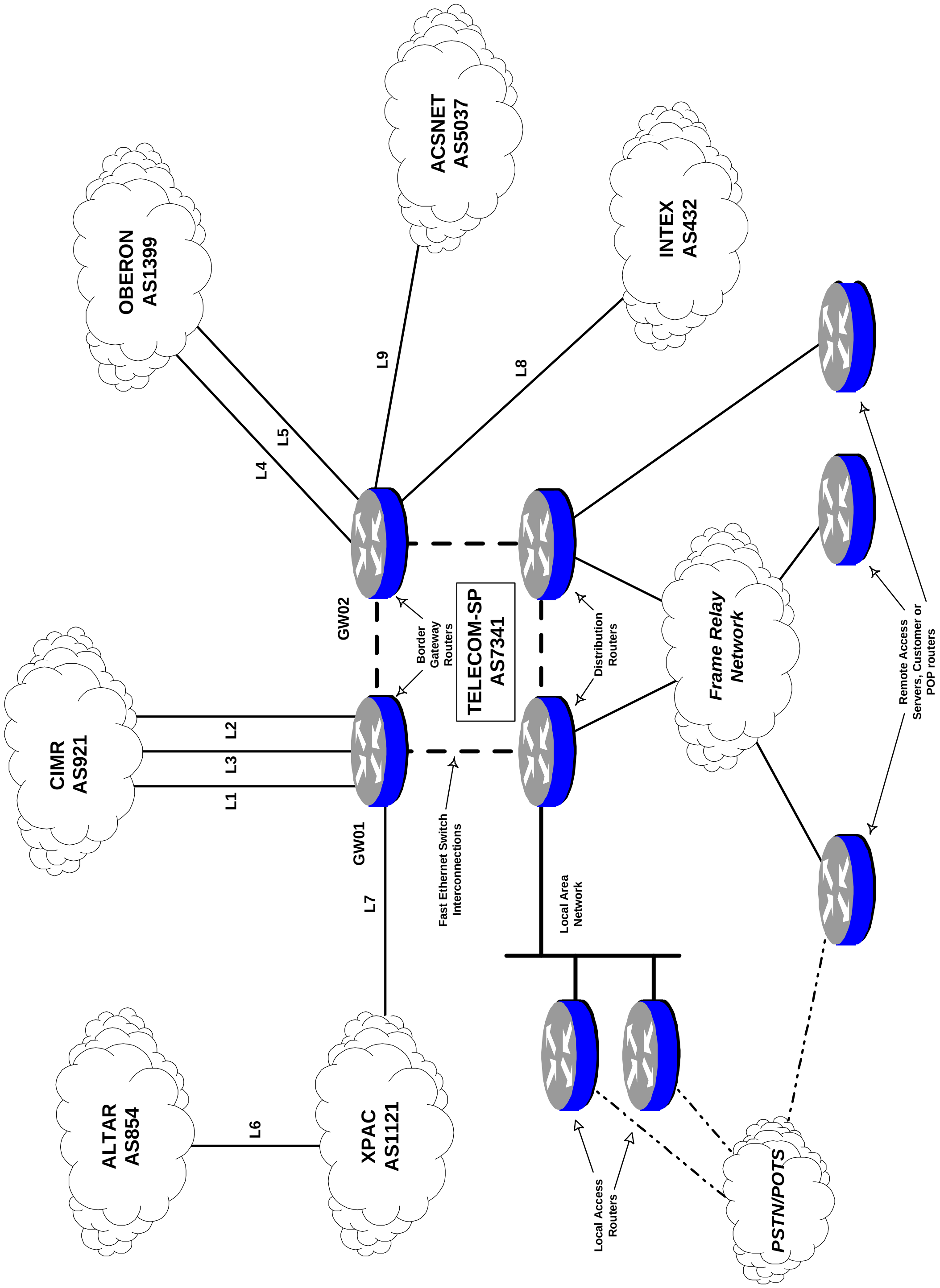
TABLE 4.2-A

3.0 Networking Architecture

3.1 Network Components

The architecture recommended is displayed in **DIAGRAM 3.1-A**. In this case, there are two gateway border routers, GW01 and GW02. These two routers will provide all of the border gateway connectivity between TELECOM-SP AS7341 and other Autonomous Systems. Recommended routers for this task are 7206, 7505 or 7507. The 7206 is a good entry-level platform for this task. The 7507 router would be the most expandable and would suit expanding future requirements such as with support for RSP4 and distributed switching. Two gateway routers are recommended for redundancy. Two 7513s serve as distribution routers providing connectivity between the border gateway routers and the edge routers of the network. The edge routers can consist of local or remote dial-access or private customer routers. This architecture effectively breaks the network out into three layers, gateway, distribution and edge functionality. The gateway and distribution routers can also be located in geographically separate locations depending on the requirements and facilities available.

TELECOM-SP Network Architecture & Connectivity



3.2 Topology

The topology proposed distributes all of the BGP peering links among the two gateway routers. In this case, all CIMR links are connected to GW01 and OBERON links to GW02. This serves to provide redundancy by using CIMR and OBERON as backups to each other. For more complex redundancy, CIMR and OBERON links could be cross-connected to both gateway routers but the complexity must be weighed against the benefits of this option. INTEX and ACSNET are also connected via router GW02 and XPAC is connected to router GW01. Although at this time eBGP multihop or multipath is not considered, the possible future use of eBGP multihop and multipath should be considered and thus it is recommended that links of the same speed to the same neighboring AS be placed on the same local (i.e. TELECOM-SP) router (as has been done in this case).

The gateway and distribution routers are interconnected using a Fast Ethernet switch (as per the TELECOM-SP design information). In this case there is implemented a path (1) between the two gateway routers; (2) between the two distribution routers and (3) between each gateway router and their respective distribution router. The path between the two gateway routers serves to support AS-transit or re-routed AS egress or ingress traffic without affecting the distribution routers. With this configuration there is sufficient physical redundancy to support failures in the FE switch matrix, routers or BGP peering links.

3.3 BGP Peering Sessions

Each inter-AS link will implement an eBGP neighbor peering session with (it is assumed at this point) a separate peer router in a neighboring AS. No eBGP multihop or BGP multipath sessions are demonstrated in the document. Both gateway routers will implement an iBGP peering session between them. This is recommended in order to (1) facilitate the exchange of routing information in order for best-path selection to be coordinated between the two gateway routers, and (2) to provide transit for inter-AS traffic (e.g. XPAC to OBERON, CIMR and beyond) which does not impact the distribution routers.

3.4 Default Routing

Default routing is necessary in this network for several reasons:

1. Partial routing is accepted from all neighbors. As a result, traffic destined for hosts outside of TELECOM-SP AS7341, that are not explicitly listed in the routing table, must be defaulted;

2. BGP-learned routes are not being injected into the Internal Gateway Protocol (IGP) process (in this case EIGRP 7341), therefore, default routing must be implemented within AS7341 in order to provide a default path for traffic destined for hosts external to AS7341. It is sufficient to ensure that packets being forwarded to destinations for which no specific route exists, be

defaulted up to the gateway routers. Once there, more specific routes implemented in the gateway router's routing table (from BGP-learned routes) will enable routing over a specific link to the destination via the best path or the default path.

3.4.1 How Default Routing is Implemented

In this design, default routes to CIMR and OBERON are implemented. CIMR and OBERON, each also implement weighted default routes which are used to provide backup default routes. Default routes are defined by creating static routes pointing to the particular upstream next-hop. A backup default route for the same neighbor AS is implemented by using a distance value of 210. Therefore, for example, should the primary CIMR default link fail, the backup default link will be used. The primary default route will be removed from the local gateway's routing table and the backup default route will be implemented. Should both the primary and backup links fail, the OBERON default link will be used. This will be enabled by GW02's default route being implemented in GW01's routing table. Defaults for destinations external to the TELECOM-SP AS are only implemented for CIMR and OBERON.

These default routes, are injected into the IGP through redistribution. A redistribution filter is implemented to ensure that no other unwanted static routes are injected into the IGP along with the static default route. Once this takes place, the default routes are propagated throughout the TELECOM-SP AS. Left alone, this mechanism implements what is known as "nearest exit defaulting". Packets will follow the default path to the nearest gateway router. Once there, packets destined externally for a host which must be defaulted (its subnet or netblock is not in the routing table), will simply exit via the local gateway's default link. However, a packet which is destined for a host whose subnet (or netblock) is explicitly listed in the routing table of the local gateway will either (1) exit via a link on the local gateway or, (2) be forwarded to the other gateway router and exit there. This is because BGP will select the best path. Because iBGP is implemented between the two gateways, this best path information will be shared/coordinated by the two routers and the routing tables will be adjusted accordingly.

3.5 BGP and IGP Interaction

BGP-learned routes are not injected into the IGP in order to reduce the affect on other routers in the TELECOM-SP AS. The total number of prefix announcements in the Internet is currently about 47,000. Injecting all of these routes into the local IGP could seriously overtax it and cause disruption in routing within the AS. Selective injection of BGP-learned routes into the IGP, through filtering, could be implemented but was not deemed necessary in this case. Although partial routing is being taken from neighbors, this action is still not deemed necessary and is not recommended.

IGP-learned local routes are not being injected into the BGP process. All netblocks for XPAC and TELECOM-SP are defined statically (using BGP *network* statements) at the gateway routers. This eliminates the possibility of TELECOM-SP injecting instability into the Internet by constantly withdrawing and re-announcing "misbehaving" prefixes (e.g. customer networks with flapping access links) within the TELECOM-SP network.

4.0 Policy Architecture

4.1 Assumptions

In section **4.3 Sample Router Configurations** you will find a partial configurations for the gateway routers GW01 and GW02. A number of assumptions have been made in order to develop these configurations. These include:

1. Traffic destined for netblocks A, B, C and D is preferred primarily across CIMR links and then as backup across OBERON links.
2. Each link between TELECOM-SP and it's neighbors is interconnected with a separate neighbor router.
3. CIMR permits the use of the Community attribute to provide customer-originated information regarding the preference of return paths. Please verify with CIMR before implementation of this mechanism. Another mechanism that could be used to achieve this would be to implement the Multi-Exit Discriminator (MED) attribute. This has been done in the case of the OBERON peering sessions and was not used as an example in this case with CIMR.
4. OBERON permits the use of the Multi-Exit Discriminator to provide customer-originated information regarding the preference of return traffic paths. Please verify with OBERON before implementation.
5. Network 212.18.22.0/24 is used as the address space for gateway, distribution & edge routers in the TELECOM-SP AS in the example configurations.
6. Variable Length Subnetting (VLSMs) implementation is assumed for the TELECOM-SP AS.
7. It is assumed that an arrangement has been made between TELECOM-SP and CIMR and OBERON to provide partial routing and that filtering of incoming prefix announcments (to ensure compliance of the receipt of only partial routing) by TELECOM-SP is not necessary. A very limited example of filtering based on AS_Path has been implemented in the configurations provided. This was done to illustrate how filtering can be done should OBERON or CIMR not provide partial routing. Care must be taken in this case because this filtering mechanism restricts acceptable prefixes to OBERON and CIMR local routes only. Any CIMR or OBERON customers who operate their own ASes and peer with OBERON will be filtered unless they are included in the access list upon which the filters are based.
8. It is assumed that filtering of incoming prefix announcements *is* necessary in order to ensure compliance of the receipt of partial routing information from ACSNET, INTEX and XPAC. This was done to ensure that only local routes/prefixes and those from ASes which are customers of INTEX, ACSNET and XPAC are allowed into the TELECOM-SP AS. Care must be taken in this case because this filtering mechanism restricts acceptable prefixes to INTEX, ACSNET and XPAC local routes only. Any INTEX, ACSNET and XPAC customers who operate their own ASes and peer with INTEX, ACSNET and XPAC will be filtered unless they are included in the access list upon which the filters are based. Another benefit of this filtering will be to ensure that ACSNET, INTEX and XPAC do not become transit networks for TELECOM-SP traffic to other than their own customers.
9. In the router configurations the serial interface numeric values correspond to the link numbers. For example, the serial interface on router GW01 for link L1 will be Serial1.

10. It is assumed that all connections between TELECOM-SP (AS7341) and other ASes are in "good-faith" and that no defaulting of traffic into the TELECOM-SP network is being conducted. Should this be the case several measures can be taken. These options are not discussed in this document version.
11. **`eigrp passive-interface`** statements are implemented on each peering link in order to eliminate unnecessary protocol overhead traffic. EIGRP will still inject the subnet addresses associated with these interfaces, on the TELECOM-SP gateway routers, into the EIGRP process for propagation throughout the EIGRP 7341 process.
12. BGP peering links are numbered as indicated below. The address of the next-hop on the router of the AS neighboring TELECOM-SP is always alpha dotted-decimal address representation for the three Most-Significant-Bytes with a "**`1`**" in the Least Significant Byte (e.g. the L1 link interface on the CIMR router is referred to as **`a.a.a.1`**).
    - **`a.a.a.1`** `- Neighbor next-hop for Link L1 to CIMR`
    - **`b.b.b.1`** `- Neighbor next-hop for Link L2 to CIMR`
    - **`c.c.c.1`** `- Neighbor next-hop for Link L3 to CIMR`
    - **`e.e.e.1`** `- Neighbor next-hop for Link L4 to OBERON`
    - **`f.f.f.1`** `- Neighbor next-hop for Link L5 to OBERON`
    - **`g.g.g.1`** `- Neighbor next-hop for Link L9 to ACSNET`
    - **`h.h.h.1`** `- Neighbor next-hop for Link L8 to INTEX`
    - **`j.j.j.1`** `- Neighbor next-hop for Link L7 to XPAC`
    - **`y.y.y.y`** `- Subnet for internal Fast Ethernet connection between GW01 and GW02. y.y.y.1 is the address for the interface on GW01 and y.y.y.2 is the address for the interface on GW02.`

4.2 Routing Policy Implemented

The purpose of this section is to present the specific policy mechanisms that have been implemented in the router configurations in order to support the general policy discussed in section **2.1 Requirements**.

Inter-Autonomous System routing is largely asymmetric. As a result, traffic flowing out of an ISP from a specific subnet or net-block over one link will not necessarily use the same link for traffic which is flowing back into the AS destined for that same subnet or netblock. The end-to-end paths over which the traffic travels are largely asymmetric and independent. Attempting to attain a high degree of traffic flow symmetry for a specific subnet or net-block requires (1) a lot of knowledge of the particular AS's operational conditions (i.e. traffic flow patterns and levels), (2) periodic updating of the configurations as well as the the policy implemented to map to the dynamically changing Internet environment, and (3) an understanding of the operational conditions of the immediate upstream neighbors.

1. **Policy for neighbor CIMR (AS921):**
    - BGP Community attribute is used to indicate to CIMR the **local path preference** for each netblock of address space, as shown in the table below (Table 4.2-A). The community attribute is set for each prefix (as defined in the table) by the TELECOM-SP gateway router, GW01, and is passed to the CIMR neighbor routers. The neighbor routers then use this community attribute value to set the Local_Preference value for

the specific prefixes. A higher value of the Community attribute, and thus Local_Pref, indicates a higher preference for a particular route. A value of 100 for the Local_Pref is the default and need not be set by setting the Community attribute value. CIMR policy (implemented in the CIMR neighbor routers) is configured to filter prefix announcements, coming from the TELECOM-SP gateway routers, by Community attribute, and based on the value, to set the Local_Preference attribute accordingly (see the references, RFC-1998 for a specific example). This mechanism could be implemented with TELECOM-SP's other neighbors and customer ASes in order to implement an effective preference mechanism.

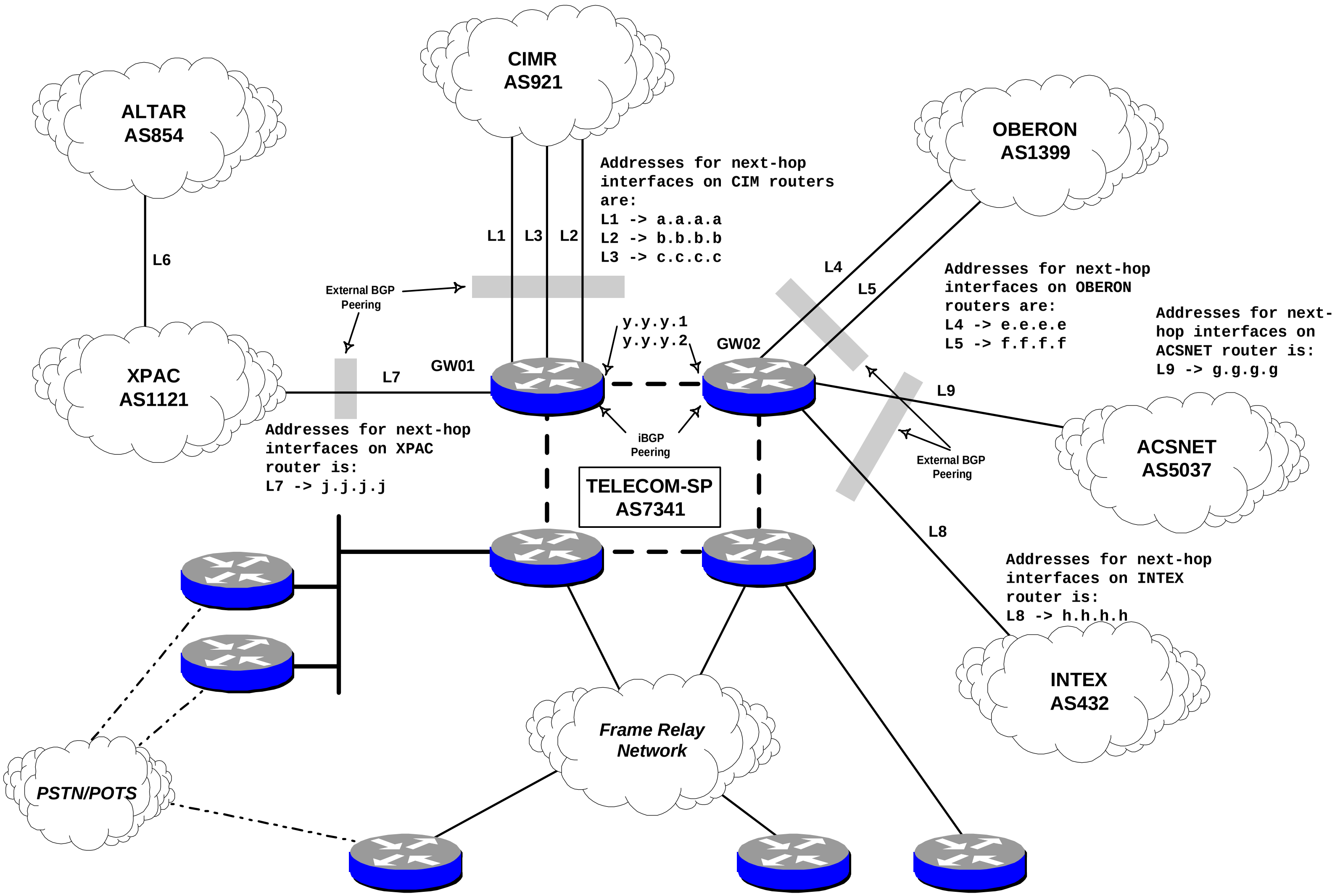
Network Block	Net-Block Group	Link L1/ Preference	Link L2/ Preference	Link L3/ Preference
212.18.22.0/22	A	*default*	90	N/A
212.18.12.0/21	B	*default*	90	N/A
201.9.110.0/20	C	*default*	90	N/A
173.41.220.0/21	D	*default*	90	N/A
200.5.32.0/19	E	*default*	90	N/A
200.23.7.0/19	F	N/A	90	*default*
172.22.0.0/16	G	N/A	90	*default*

TABLE 4.2-A

Note 1: N/A - Not Announced out this interface/link.

Note 2: Local_Pref default value is 100.

TELECOM-SP Network Architecture & Connectivity



Addresses for next-hop interfaces on CIM routers are:
L1 -> a.a.a.a
L2 -> b.b.b.b
L3 -> c.c.c.c

Addresses for next-hop interfaces on OBERON routers are:
L4 -> e.e.e.e
L5 -> f.f.f.f

Addresses for next-hop interfaces on ACSNET router is:
L9 -> g.g.g.g

Addresses for next-hop interfaces on XPAC router is:
L7 -> j.j.j.j

Addresses for next-hop interfaces on INTEX router is:
L8 -> h.h.h.h

- Partial routing is taken from CIMR (AS921). In this case, (1) an arrangement can be made with CIMR so that it only announces its local prefixes and its customer prefixes or (2) filtering can be implemented by TELECOM-SP which will restrict the acceptable prefixes to those which originate from CIMR's AS (i.e by filtering incoming announcements based on AS origination). One possible problem with option (2) is that unless direct communication takes place with CIMR to obtain a list of the provider's customer ASes (if CIMR wishes to make this information available) this information is difficult to obtain. If customer prefixes are not obtained and implemented in the filter the list of "visible" prefixes will be restricted to prefixes local to CIMR.
- **Private address space prefix filtering**, as detailed in RFC1918, are filtered from ingress routing announcements.
- **End-to-end path preference** is implemented through the use of AS-Path prepend. Refer below to the section under **2. Policy for neighbor OBERON (AS1399)** regarding end-to-end path preference for more detail.
- Filtering of prefix announcements based on AS_Path is implemented to prevent CIMR from receiving routing announcements from any ASes other than XPAC AS1121 (and its customers) and TELECOM-SP AS7341 (and its customers). This is accomplished by implementing an AS_Path filter which filters out all prefix announcements which have *passed through* ASes 432, 5037, 1399, 921 and 854. The main purpose for this is to prevent AS7341 TELECOM-SP, from becoming a transit AS for CIMR. This is implemented on the two TELECOM-SP gateway routers and affects prefixes being announced to the respective neighbor ASes.

2. **Policy for neighbor OBERON (AS1399):**

- BGP MED (Multi-Exit Discriminator) attribute is used to indicate to OBERON the **local path preference** for each netblock of address space, as shown in the table below. In the case of the MED, the lower value is the most preferred path. For example, a MED value of 0 has a higher preference than a value of 50, therefore traffic destined for net-block A (within TELECOM-SP) will prefer a path over a link with the lowest MED value, in this case 0, on link L4. For netblock A, the preference value of 50 for link L5 is actually set in router GW02 and passed to the OBERON neighbor router. A value of 0 does not have to be set by GW02 since it is the default value set by the BGP process in the OBERON neighbor router.

Network Block	Net-Block Group	Link L4/ Preference	Link L5/ Preference
212.18.22.0/22	A	*default*	50
212.18.12.0/21	B	*default*	50
201.9.110.0/20	C	*default*	50
173.41.220.0/21	D	*default*	50
200.5.32.0/19	E	*default*	50
200.23.7.0/19	F	50	*default*

172.22.0.0/16	G	50	*default*

TABLE 4.2-B

Note 1: N/A - Not Announced out this interface/link.
Note 2: MED default value is 0.

- **Partial routing** is taken from OBERON (AS1399). In this case, (1) an arrangement can be made with OBERON so that it only announces its local prefixes and its customer prefixes or (2) filtering can be implemented by TELECOM-SP which will restrict the acceptable prefixes to those which originate from OBERON's AS (i.e by filtering incoming announcements based on AS origination). One possible problem with option (2) is that unless direct communication takes place with OBERON to obtain a list of the provider's customer ASes (if OBERON wishes to make this information available) this information is difficult to obtain. If customer prefixes are not obtained and implemented in the filter the list of "visible" prefixes will be restricted to prefixes local to OBERON.
- **Private address space prefix filtering**, as detailed in RFC1918, are filtered from ingress routing announcements.
- **End-to-end path preference** is implemented through the use of AS-Path prepending. In this case, traffic destined for netblocks A through D is preferred across the CIMR links, utilizing the OBERON links as a backup for this traffic. All other traffic (that destined for netblocks E through G) is governed by policy set for local path preference through the use of Community and MED attributes (in this case for CIMR and OBERON, respectively). In this case, AS_Path prepend is applied to all prefix announcements contained within netblocks A through D that are propagated across the OBERON links (links L4 and L5) from GW02. This should lengthen the AS_Path of these announcements and thus "suggest" acceptance by upstream ASes of the pathway through CIMR for traffic to destinations within these netblocks. This is, of course, under ideal conditions. Interactions between OBERON, CIMR and their neighbors (e.g. filtering, aggregation, modification of AS-Path, etc…) and of ASes even beyond that, will decide the ultimate validity of this policy implementation. It is possible that if OBERON decides not to advertise the CIMR netblocks (i.e netblocks A through D), that the preferred path for these netblocks will be through CIMR anyhow, since this would be the only source of announcements.
- Filtering of prefix announcements based on AS_Path is implemented to prevent OBERON from receiving routing announcements from any ASes other than XPAC AS1121 (and its customers) and TELECOM-SP AS7341 (and its customers). This is accomplished by implementing an AS_Path filter which filters out all prefix announcements which have *passed through* ASes 432, 5037, 1399, 921 and 854. The main purpose for this is to prevent AS7341 TELECOM-SP, from becoming a transit AS for OBERON. This is implemented on the two TELECOM-SP gateway routers and affects prefixes being announced to the respective neighbor ASes.

3. **Policy for neighbor ACSNET (AS5037):**
    - **Private address space prefix filtering**, as detailed in RFC1918, are filtered from ingress routing announcements.

- **Partial routing** is forcefully implemented for ACSNET destinations. In this case, ACSNET is not providing partial routing, but rather, TELECOM-SP is filtering routing announcements so as to permit only those which originate within the ACSNET network. This can be implemented by filtering particular address prefixes or by filtering on AS-Path information (specifically on the originating AS). In the case of this recommendation, this filtering is based on the originating AS number (i.e. the first number in the list of AS numbers in the AS_Path attribute). In order for this to be accurate, the AS numbers for all customers of ACSNET must be obtained and implemented in the access-list/filter. In the case of the router configurations below, only the ACSNET AS number 5037 is implemented. In this case only prefixes which originate in AS5037 will be accepted.
- Filtering of prefix announcements based on AS_Path is implemented to prevent ACSNET from receiving routing announcements from any ASes other than XPAC AS1121 (and its customers) and TELECOM-SP AS7341 (and its customers). This is accomplished by implementing an AS_Path filter which filters out all prefix announcements which have *passed through* ASes 432, 5037, 1399, 921 and 854. The main purpose for this is to prevent AS7341 TELECOM-SP, from becoming a transit AS for ACSNET. This is implemented on the two TELECOM-SP gateway routers and affects prefixes being announced to the respective neighbor ASes.

4. **Policy for neighbor INTEX (AS432):**
   - **Private address space prefix filtering**, as detailed in RFC1918, are filtered from ingress routing announcements.
   - **Partial routing** is forcefully implemented for INTEX destinations. In this case, INTEX is not providing partial routing, but rather, TELECOM-SP is filtering routing announcements so as to permit only those which originate within the INTEX network. This can be implemented by filtering particular address prefixes or by filtering on AS-Path information (specifically on the originating AS). In the case of this recommendation, this filtering is based on the originating AS number (i.e. the first number in the list of AS numbers in the AS_Path attribute). In order for this to be accurate, the AS numbers for all customers of INTEX must be obtained and implemented in the access-list/filter. In the case of the router configurations below, only the INTEX AS number 432 is implemented. In this case only prefixes which originate in AS432 will be accepted.
   - Filtering of prefix announcements based on AS_Path is implemented to prevent INTEX from receiving routing announcements from any ASes other than XPAC AS1121 (and its customers) and TELECOM-SP AS7341 (and its customers). This is accomplished by implementing an AS_Path filter which filters out all prefix announcements which have *passed through* ASes 432, 5037, 1399, 921 and 854. The main purpose for this is to prevent AS7341 TELECOM-SP, from becoming a transit AS for INTEX. This is implemented on the two TELECOM-SP gateway routers and affects prefixes being announced to the respective neighbor ASes.

5. **Policy for neighbor XPAC (AS 1121):**
   - **Private address space prefix filtering**, as detailed in RFC1918, are filtered from ingress routing announcements.

- **Partial routing** is forcefully implemented for XPAC destinations. In this case, XPAC is not providing partial routing, but rather, TELECOM-SP is filtering routing announcements so as to permit only those which originate within the XPAC network. This can be implemented by filtering particular address prefixes or by filtering on AS-Path information (specifically on the originating AS). In the case of this recommendation, this filtering is based on the originating AS number (i.e. the first number in the list of AS numbers in the AS_Path attribute). In order for this to be accurate, the AS numbers for all customers of XPAC must be obtained and implemented in the access-list/filter. In the case of the router configurations below, only the XPAC AS number 1121 is implemented. In this case only prefixes which originate in AS1121 will be accepted.

6. **Defaulting Policy**

- Primary Default via CIMR (and gateway router GW01) points to neighbor **`b.b.b.1`** (CIMR across link L2). Backup default will point to **`a.a.a.1`** with a weight (distance) of 210. Because of the behavior of static routes in Cisco routers, if there is a failure of link L2, the default route to **`a.a.a.1`** will be placed into the routing table. In the case of a failure of the primary default link to CIMR, the backup link will become the new default. A failure of both the primary and backup default links will leave a default path through OBERON as the sole remaining default for all egress traffic from TELECOM-SP to destinations other than INTEX, ACSNET and XPAC. Static routes are configured to the default 0/0 and these routes are, in turn, injected into the EIGRP 7341 process through redistribution and thus propagated throughout the TELECOM-SP network. Based on this mechanism, traffic exiting from the network will take the “nearest exit” out of the AS.
- Primary default via OBERON points to neighbor **`e.e.e.1`** (OBERON across link L4). Backup default to OBERON will point to **`f.f.f.1`** using a weight of 210.
- Careful consideration must be taken in the design of egress and ingress traffic loading in the case of neighbor link failures as these failures can cause other links to become severely loaded or even saturated.

4.3 Sample Router Configurations

ROUTER TELECOM-SP GATEWAY 01 (GW01)

```
hostname GW01

router eigrp 7341
network 212.18.22.0
redistribute static route-map ONLY_DEFAULT
passive-interface serial 1
passive-interface serial 2
passive-interface serial 3
passive-interface serial 7

router bgp 7341
no synchronization
no auto-summary
network 212.18.22.0 mask 255.255.252.0
network 212.18.12.0 mask 255.255.248.0
network 201.9.110.0 mask 255.255.240.0
network 173.41.220.0 mask 255.255.248.0
network 200.5.32.0 mask 255.255.224.0
network 200.23.7.0 mask 255.255.224.0
network 172.22.0.0 mask 255.255.0.0
neighbor PEER_MAP peer-group
neighbor PEER_MAP remote-as 921
neighbor PEER_MAP filter-list 20 out
neighbor PEER_MAP distribute-list 2 in
neighbor PEER_MAP distribute-list 2 out
neighbor a.a.a.1 peer-group PEER_MAP
neighbor b.b.b.1 peer-group PEER_MAP
neighbor c.c.c.1 peer-group PEER_MAP
neighbor a.a.a.1 route-map SET_COMMUNITY_1 out
neighbor b.b.b.1 route-map SET_COMMUNITY_2 out
neighbor c.c.c.1 route-map SET_COMMUNITY_3 out
neighbor a.a.a.1 send-community
neighbor b.b.b.1 send-community
neighbor c.c.c.1 send-community
neighbor j.j.j.1 remote-as 1121
neighbor j.j.j.1 filter-list 21 in
neighbor j.j.j.1 distribute-list 2 in
neighbor j.j.j.1 distribute-list 2 out
neighbor y.y.y.2 remote-as 7341

! Access list summary:
! #2 Used for Ingress to filter private space prefixes as well
!    as any other prefixes desired.
! #3 Used for redistribution into EIGRP process to permit only
!    static route to default 0/0 to be redistributed.
! #10 Used to set community attribute for CIMR sessions and to
!     implicitly filter prefixes out to CIMR peers other than
!     those specified.
! #11 Used to set community attribute for CIMR sessions and to
!     implicitly filter prefixes out to CIMR peers other than
!     those specified.
! #12 Used to set community attribute for CIMR sessions and to
!     implicitly filter prefixes out to CIMR peers other than
!     those specified.
! #20 AS Path filter to ensure that AS7341 serve only as transit
!     to AS1121 and not to any other neighboring AS.
! #21 AS Path filter to ensure that only prefixes originating
!     from AS1121 are allowed to be propagated throughout AS7341.

! Ingress filtering to prevent 1918 private address space from
! being injected into AS7341.  This access-list can be used to
! add other filters which TELECOM-SP wishes to impose at the ingress
! to their AS.
access-list 2 deny 10.0.0.0 0.255.255.255
access-list 2 deny 172.6.0.0 0.15.255.255
access-list 2 deny 192.168.0.0 0.0.255.255
access-list 2 permit any

! Access list to permit default only being injected into EIGRP
! process.
access-list 3 permit 0.0.0.0 0.0.0.0
access-list 3 deny any

! Prefixes A, B, C, D and E.
access-list 10 permit 212.18.22.0 255.255.252.0
access-list 10 permit 212.18.12.0 255.255.248.0
access-list 10 permit 201.9.110.0 255.255.240.0
access-list 10 permit 200.5.32.0 255.255.224.0
access-list 10 permit 173.41.220.0 255.255.248.0
access-list 10 deny any

! Prefixes F and G.
access-list 11 permit 200.23.7.0 255.255.224.0
access-list 11 permit 172.22.0.0 255.255.0.0
access-list 11 deny any

! AS Path filter list for outgoing prefixes.  This filter can be
! used on all peering sessions with CIMR, OBERON, INTEX and
! ACSNET.  In each case there may be overlap but it will still
! work.  It is intended to prevent AS7341 from acting as a transit
! network for everyone except XPAC (AS1121).
ip as-path access-list 20 deny ^432_
ip as-path access-list 20 deny ^5037_
ip as-path access-list 20 deny ^1399_
ip as-path access-list 20 deny _854_
ip as-path access-list 20 permit .*

! AS Path filter list for incoming prefixes.  This filter is
! used on the XPAC peering session.
ip as-path access-list 21 permit _1121$
ip as-path access-list 21 deny any

! Permit A, B, C, D and E to be propagated - they will default
! to Local_Pref of 100 in the upstream neighbor.
route-map SET_COMMUNITY_1 permit 10
match ip address 10

! Explicitly deny all other prefixes from being
! propagated to the associated peer router.
route-map SET_COMMUNITY_1 deny 20

route-map SET_COMMUNITY_2 permit 10
match ip address 10
match ip address 11
set community 0x0DE9005A

route-map SET_COMMUNITY_2 deny 20

route-map SET_COMMUNITY_3 permit 10
match ip address 11

route-map SET_COMMUNITY_3 deny 20

! This route-map is necessary to permit only the redistribution
! of the default route into the EIGRP process 7341.  There is
! no need to introduce the other “place-holder” routes to null0
! into the EIGRP process.
route-map ONLY_DEFAULT permit 10
match ip address 3

route-map ONLY_DEFAULT deny 20

! Static route definitions.
ip route 0.0.0.0 0.0.0.0 b.b.b.1
ip route 0.0.0.0 0.0.0.0 a.a.a.1 210
ip route 212.18.22.0 255.255.252.0 null0
ip route 212.18.12.0 255.255.248.0 null0
ip route 201.9.110.0 255.255.240.0 null0
ip route 173.41.220.0 255.255.248.0 null0
ip route 200.5.32.0 255.255.224.0 null0
ip route 200.23.7.0 255.255.224.0 null0
ip route 172.22.0.0 255.255.0.0 null0
```

ROUTER TELECOM-SP GATEWAY 02 (GW02)

```
hostname GW02

router eigrp 7341
network 212.18.22.0
redistribute static route-map ONLY_DEFAULT
passive-interface serial 4
passive-interface serial 5
passive-interface serial 8
passive-interface serial 9

router bgp 7341
no synchronization
no auto-summary
network 212.18.22.0 mask 255.255.252.0
network 212.18.12.0 mask 255.255.248.0
network 201.9.110.0 mask 255.255.240.0
network 173.41.220.0 mask 255.255.248.0
network 200.5.32.0 mask 255.255.224.0
network 200.23.7.0 mask 255.255.224.0
network 172.22.0.0 mask 255.255.0.0

neighbor PEER_MAP peer-group
neighbor PEER_MAP filter-list 20 out
neighbor PEER_MAP distribute-list 2 in
neighbor PEER_MAP distribute-list 2 out
neighbor e.e.e.1 peer-group PEER_MAP
neighbor f.f.f.1 peer-group PEER_MAP
neighbor g.g.g.1 peer-group PEER_MAP
neighbor h.h.h.1 peer-group PEER_MAP
neighbor e.e.e.1 remote-as 1399
neighbor e.e.e.1 route-map SET_MED_L4 out
neighbor f.f.f.1 remote-as 1399
neighbor f.f.f.1 route-map SET_MED_L5 out
neighbor g.g.g.1 remote-as 5037
neighbor g.g.g.1 filter-list 21 in
neighbor h.h.h.1 remote-as 432
neighbor h.h.h.1 filter-list 22 in
neighbor y.y.y.1 remote-as 7341

! Access list summary:
! #2 Used for Ingress to filter private space prefixes as well
!    as any other prefixes desired.
! #3 Used for redistribution into EIGRP process to permit only
!    static route to default 0/0 to be redistributed.
! #20 AS Path filter to ensure that AS7341 serve only as transit
!     to AS1121 and not to any other neighboring AS.
! #21 AS Path filter to ensure that only prefixes originating
!     from AS5037 are allowed to be propagated throughout AS7341.
! #22 AS Path filter to ensure that only prefixes originating
!     from AS432 are allowed to be propagated throughout AS7341.

! Ingress filtering to prevent 1918 private address space from
! being injected into AS7341.  This access-list can be used to
! add other filters which TELECOM-SP wishes to impose at the ingress
! to their AS.
access-list 2 deny 10.0.0.0 0.255.255.255
access-list 2 deny 172.6.0.0 0.15.255.255
access-list 2 deny 192.168.0.0 0.0.255.255
access-list 2 permit any

! Access list to permit default only being injected into EIGRP
! process.
access-list 3 permit 0.0.0.0 0.0.0.0
access-list 3 deny any

! AS Path filter list for outgoing prefixes.  This filter can be
! used on all peering sessions with CIMR, OBERON, INTEX and
! ACSNET.  In each case there may be overlap but it will still
! work.  It is intended to prevent AS7341 from acting as a transit
! network for everyone except XPAC (AS1121).
ip as-path access-list 20 deny ^432_
ip as-path access-list 20 deny ^5037_
ip as-path access-list 20 deny ^1399_
ip as-path access-list 20 deny ^921_
ip as-path access-list 20 deny _1121_854_
ip as-path access-list 20 permit .*

! AS Path filter list for incoming prefixes.  This filter is
! used on the ACSNET peering session.
ip as-path access-list 21 permit _5037$
ip as-path access-list 21 deny any

! AS Path filter list for incoming prefixes.  This filter is
! used on the INTEX peering session.
ip as-path access-list 22 permit _432$
ip as-path access-list 22 deny any

! Prefixes A, B, C, and D.
access-list 10 permit 212.18.22.0 255.255.252.0
access-list 10 permit 212.18.12.0 255.255.248.0
access-list 10 permit 201.9.110.0 255.255.240.0
access-list 10 permit 173.41.220.0 255.255.248.0
access-list 10 deny any

! Prefix E.
access-list 12 permit 200.5.32.0 255.255.224.0
access-list 12 deny any

! Prefixes F and G.
access-list 11 permit 200.23.7.0 255.255.224.0
access-list 11 permit 172.22.0.0 255.255.0.0
access-list 11 deny any
```

```
! The following two route maps are used to modify the
! length of the AS_Path attribute for various prefixes
! to make them more desirable to OBERON on particular
! links.  For example, prefix groups A, B, C, D and E
! will have their AS_Paths prepended with the local AS
! number on link L4 but not on link L5.  All other metrics
! being equal, this should therefore induce the OBERON network
! to decide that link L4 is less desirable than link L5
! for traffic to be routed to AS7341.  This will be manifested
! in the BGP path selection process which will select the route
! with the shortest AS-Path.

route-map SET_MED_L4 permit 10
match ip address 11
set metric 50

route-map SET_MED_L4 permit 20
match ip address 10
set as-path prepend 7341 7341

route-map SET_MED_L4 permit 30
match ip address 12

route-map SET_MED_L4 deny 40

route-map SET_MED_L5 permit 10
match ip address 10
set metric 50

route-map SET_MED_L5 permit 20
match ip address 12
set metric 50

route-map SET_MED_L5 permit 30
match ip address 11

route-map SET_MED_L5 deny 40

! This route-map is necessary to permit only the redistribution
! of the default route into the EIGRP process 7341.  There is
! no need to introduce the other “place-holder” routes to null0
! into the EIGRP process.
route-map ONLY_DEFAULT permit 10
match ip address 3

route-map ONLY_DEFAULT deny 20

! Static route definitions.
ip route 0.0.0.0 0.0.0.0 e.e.e.1
ip route 0.0.0.0 0.0.0.0 f.f.f.1 210
ip route 212.18.22.0 255.255.252.0 null0
ip route 212.18.12.0 255.255.248.0 null0
ip route 201.9.110.0 255.255.240.0 null0
ip route 173.41.220.0 255.255.248.0 null0
ip route 200.5.32.0 255.255.224.0 null0
ip route 200.23.7.0 255.255.224.0 null0
ip route 172.22.0.0 255.255.0.0 null0
```

APPENDIX A - NSP Filter Policy (some examples)

```
OBERON'S FILTERING POLICY
=========================

Return-Path: owner-nanog@merit.edu
X-Authentication-Warning: iscserv.res.SPRINTlink.net: vgoel owned process doing -bs
Date: Tue, 15 Apr 1997 23:50:24 -0400 (EDT)
From: Vab Goel <vgoel@SPRINT.net>
X-Sender: vgoel@iscserv.res.SPRINTlink.net
To: Gregory Hersh <ghersh@bbnplanet.com>
cc: nanog@merit.edu
Subject: Re: SPRINT Filters - what are they?
Sender: owner-nanog@merit.edu



SPRINT currently filters announcements from its
non-customers as follows:

        RFC 1597 reserved space: accept nothing                    (**)
        In the classical "A" space: accept nothing longer than /8
        In the classical "B" space: accept nothing longer than /16
                in 24/8 space: accept nothing longer than /19       (*)
                in 195/8: accept nothing longer than /19
                in 206/8 - 223/8: accept nothing longer than /19
                in 192/8 - 205/8: accept nothing longer than /24

(*)  in-line with the IP registries allocation
(**) RFC 1597>  Private Address Space

RFC 1597>   The Internet Assigned Numbers Authority (IANA) has reserved
RFC 1597>   the following three blocks of the IP address space for private
RFC 1597>   networks:
RFC 1597>        10.0.0.0        -   10.255.255.255
RFC 1597>        172.16.0.0      -   172.31.255.255
RFC 1597>        192.168.0.0     -   192.168.255.255

Vab..

On Tue, 15 Apr 1997, Gregory Hersh wrote:

> A while ago we had a customer with some routing problems which were caused
> by SPRINT filters (127/8 - 191/8, deny anything longer than /16). When
> customer called SPRINT, he's been told by someone from tech support that
> 'SPRINT doesn't filter anything'. When I called support, I've been told
> that on 127/8 - 191/8 range SPRINT denies anything longer than /16. This
> morning the same customer called SPRINT again, to confirm this information,
> and been told that for the above range SPRINT denies anything longer than /19.
> It seems like each time you call SPRINT support, you hear a different story.
>
> Will someone from SPRINT _who knows_ confirm their filtering policy which
> I have as follows:
>
> 0/8 - 126/8, deny subnets of historical A's
> 127/8 - 191/8, deny anything longer than /16
> 192/8 - 205/8, deny anything longer than /24
> 206/8 - 223/8, dney anything longer than /19
> 192/8 [RIPE], deny anything longer than /19
>
>
> It would be really nice to have it posted at NANOG web page as well
> (alas, only AGIS done so).
```

```
>
> Thanks,
>
> - Greg -
>


AGIS'S FILTERING POLICY
=======================

As per Peter Kline's NANOG 8 presentation:

- Filter TWD at /24
- Filter 206+ at /19
- Filter everything else at /16
- Use of Next-Hop-Self at all exchange point peering sessions.
  - Expect peers to do likewise and will route map if necessary.
```

References

[RFC-1918] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot & E. Lear, **“Address Allocation for Private Internets”**, RFC-1918, February 1996

[RFC-1930] J. Hawkinson, T. Bates, **“Guidelines for creation, selection, and registration of an Autonomous System”**, RFC-1930, March 1996

[RFC-1998] E. Chen & T. Bates, **“An Application of the BGP Community Attribute in Multi-home Routing”**, RFC-1998, August 1996

Document Updates

1. **[Date: April 5, 1998]** - Modification to router configurations beginning on page 17 to implement *"send-community"* statements which were left out of the original document. In this case, Communities are only being sent to CIMR. If communities must be sent to CIMR via third-party ASes (e.g. ALTAR and/or OBERON), the **`neighbor send-community`** command will need to be implemented for each neighbor in the third-party AS. This has not been done in this document. With this configuration, it is implied that third-party ASes will not forward community information to CIMR. This modification affects only the configuration for router **GW01**. ***Version changed to 1.10***
2. **[Date: April 28, 1998]** - Modifications to the router configurations beginning on page 17 to implement the **`no auto-summary`** command for BGP and the filter direction identifiers (i.e. in/out) in the GW01 configuration for the **`neighbor-route-map SET_COMMUNITY_1, "_2, and "_3`** statements as well as for the **`neighbor-route-map SET_MED_L4 and "_L5`** in the GW02 configuration. ***Version changed to 1.20***
3. **[Date: April 30, 1998]** - Modifications to the router configurations beginning on page 17 to implement outgoing prefix filters for RFC1918 blocks. ***Version changed to 1.30***
4. **[Date: May 2, 1998]** - (1) Minor corrections needed to be made to correct the AS1800 to make it AS854 as per the diagrams. (2) Corrected some of the AS_Path filters in the router configurations to more accurately reflect the policy to be implemented. (3) Corrected and updated the diagrams. ***Version changed to 1.31.***