BGP-like TE Capabilities for SHIM6

Marcelo Bagnulo, Alberto García-Martínez, Arturo Azcorra
Departamento de Ingeniería Telemática, Universidad Carlos III de Madrid
{marcelo, alberto, azcorra}@it.uc3m.es

Abstract

In this paper we present a comprehensive set of mechanisms that restore to the site administrator the capacity of enforcing Traffic Engineering (TE) policies in a multiaddressed IPv6 scenario. The mechanisms rely on the ability of SHIM6 to securely perform locator changes in a transparent fashion to transport and application layers. Once an outgoing path has been selected for a communication by proper routing configuration in the site, the source prefix of SHIM6 data packets is rewritten by the site routers to avoid packet discarding due to ingress filtering. The SHIM6 locator preferences exchanged in the context establishment phase are modified by the site routers to influence in the path used for receiving traffic. Scalable deployment is ensured by the stateless nature of these mechanisms.

1. Introduction[1]

The growing concern on communications reliability has resulted in a continuous increase of the number of sites that become multihomed, i.e. sites connected to the Internet through multiple providers. This configuration enables an improvement in the fault tolerance, along with the possibility of defining and implementing Traffic Engineering (hereafter TE) policies. In particular, fault tolerance and TE are a must for multimedia applications, which impose tight requirements on reliability and performance.

However, the deployment scope of multihoming in IPv4 networks is limited to somehow large sites, since multihoming depends on the injection in the BGP routing system of the prefix of the multihomed site, and massive prefix injection would lead to the collapse of the interdomain routing system. IPv6, apart from providing a much larger address space to enable a stable end-to-end addressing model, it allows the deployment of multihomed sites without stressing the global routing system. In particular, the large address availability allows small multihomed sites to obtain Provider Aggregatable prefixes from their providers' address blocks. Since the higher level providers only announce their own prefix block into the global routing system, a multihomed host is reachable at a given address only through its corresponding higher level provider. Consequently, in order to be reachable through all the available providers, a host within a multihomed end-site needs to configure as many addresses as prefixes are available in the multihomed site, becoming a *multiaddressed* host.

A multiaddressed host has to include some new mechanisms to properly manage its multiple addresses when establishing new communications. Some basic support is required for preventing the discard of packets due to the ingress filtering performed by the providers [1], since a packet with a source address that does not correspond to the provider to which it has been forwarded will be dropped [2]. Another problem for the multiaddressed hosts is the preservation of an established communication when an outage affects the provider through which the communication is flowing. To solve this, it is required a mechanism to allow diverting the communication to the address of an alternative provider transparently to the transport and application layers. Transparency to upper layers is required because current transport layers identify the endpoints of a communication through the IP addresses of the nodes involved. A multihoming protocol located in a SHIM6 layer within the IP layer is proposed by the IETF [3] to perform a locator change when an outage occurs, in a secure and transparent fashion with respect to transport and application layers.

While fault tolerance support has received much attention, some challenges are raised in the management of TE capabilities. First, it should be noted that, in the multiaddressing model, the local address currently being used for a communication determines the provider through which the multiaddressed host is accessed. Some tools have been proposed to allow the end-host to influence the outgoing and incoming data path [4]: On one hand, DNS record manipulation allows expressing

[1] This work has been partially supported by the OPTINET6 project TIC-2003-09042-C03-01 and by the IMPROVISA project TSI2005-07384-C03-02

Proceedings of the 32nd EUROMICRO Conference on Software Engineering and Advanced Applications (EUROMICRO-SEAA'06)
0-7695-2594-6/06 $20.00 © 2006 IEEE

preferences for the addresses selected by an external host that initiates a communication with the multiaddressed host. On the other hand, the Address Selection procedure [5] standardized for IPv6 hosts can be configured by means of the DAS Policy Table to express preferences for the selection of the source address in a communication initiated by the multiaddressed host.

However, a relevant concern related with TE is the lack of tools to allow a multihomed site which does not own its addresses, either an end-site or a provider, to influence in the traffic that it receives or sends to its providers. In figure 1 we can see a provider ISP_X that obtains addresses and connectivity to the Internet through ISP_A and ISP_B, with C1 and C2 being client networks. These client networks may also have any other providers apart from ISP_X, as it occurs to C2. The provider ISP_X would require the deployment of some policies for the traffic being exchanged with the rest of the Internet, such as preferring a particular provider to obtain better performance, lower costs, etc.; or balancing traffic among both providers. The mechanism presented so far for TE in multiaddressing contexts [4] requires the configuration of the specific policies that satisfy the requirements of all the providers at each end-host: the administrators of the hosts at C2 should take into account the preferences at site C2, and combine then with the preferences of ISP_X, ISP_Y, along with all the higher level providers that do not inject their prefix into the interdomain routing system. This mechanism results in a very complex configuration, and besides, it is not compatible with dynamic changes in the policies of the sites. Additionally, some pieces of the network can depend on different administrators, requiring a difficult coordination to obtain the required behavior from the network.

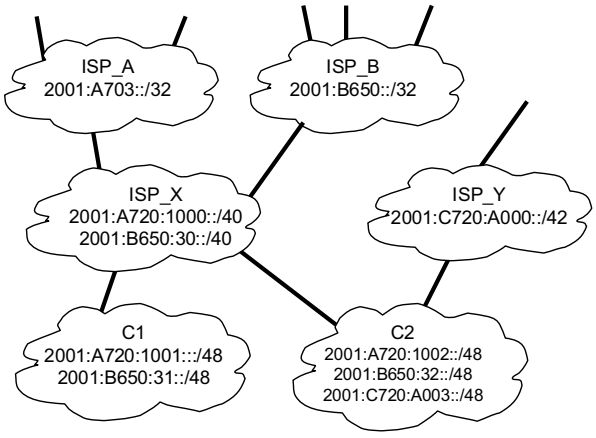

Figure 1. Address assignment in a IPv6 multiaddressed environment

In this paper we propose a comprehensive set of tools to enforce TE policies for SHIM6 data packets without detailed configuration in the end-hosts. Additionally, these tools allow the administrators of end-sites and providers to enforce TE policies in a multiaddressed scenario. Prefix rewriting for source addresses plus proper internal routing configuration allow path selection for egress data packets, while preference rewriting for SHIM6 context establishment packets influences in path selection for ingress data packets.

In the next sections we discuss current BGP TE capabilities, since the aim is to provide similar features to the ones available in the incumbent solution. We also describe the SHIM6 protocol and its security framework. Next, the mechanisms proposed so far for enforcing TE in a multiaddressed environment are presented. We detail the proposed mechanisms for TE with SHIM6, and show an application scenario in which BGP information is available. Finally we end the paper with the conclusions.

2. Traffic Engineering in IPv4 multihomed sites

The most widely deployed multihoming solution for IPv4 networks is based on the announcement of the site prefix through all its providers. In this configuration, the site S obtains a prefix allocation directly from the Regional Internet Registry. Then, the site announces this prefix to its providers using BGP [6]. The providers of the multihomed site announce the prefix to their own providers and so on. This mechanism provides fault tolerance capabilities, which include preserving established connections throughout an outage.

The following TE tools are available to the multihomed site:

- *TE mechanisms for outgoing traffic*: Multihomed sites use BGP attributes to specify preferences for the outgoing link for a prefix received. Essentially, the LOCAL_PREFERENCE attribute is set accordingly to the site TE requirements, so that preferred routes are selected when they are available in order to reach specific destinations. In this scheme, TE is determined and enforced by the site's BGP routers through manual configuration, and hosts are not involved.
- *TE mechanisms for incoming traffic*: Multihomed sites can inject a combination of routes to the interdomain routing system that includes several more or less specific prefixes referring to their own addresses. Less specific prefixes provide fall back routes, in case that the more specific routes are not available. More specific routes express TE policies, so that traffic for these more specific prefixes is routed through the desired path.

  In addition, a multihomed site can somehow influence part of the path through which packets flow

to the site using *AS path prepending*, so that a remote site can perceive one of the paths to the same prefix less attractive than the other ones. However, it must be noted that this configuration can be overridden by the sites that are forwarding the packet.

Finally, other BGP attributes can be used to express preference for incoming traffic. The COMMUNITY attribute allows tagging some announces in order to inform that some policy previously accorded between the sites should be applied. The MED attribute is also used to express policies between neighbor sites.

In any case, the TE capabilities reside in the routers, and hosts cannot influence the path used.

While the presented IPv4 multihomed solution provides fairly good features regarding to fault tolerance and TE, it presents very limited scalability with respect to the interdomain routing system, since each multihomed site using this solution contributes with new routes to the already oversized routing table. For this reason, more scalable multihoming solutions are being explored for IPv6.

3. IPv6 multihoming architecture

As it has been presented in the introduction, Provider Aggregatable addressing is used to guarantee the scalability of the multihoming solution. Multihomed sites obtain one prefix per each one of their providers. Consequently, as each provider only announces its own prefix to the rest of the Internet, a given provider is used to reach the multihomed site only when the destination addresses belong to the prefix associated with the provider. So, in order to be reachable through all the providers of the site, each host within the multihomed site has to configure multiple addresses, one per provider.

To provide fault tolerance to established communications, the SHIM6 architecture defines a protocol [3] and a security framework based on addresses with cryptographic properties. The SHIM6 architecture allows diverting a packet of a communication to an address of the host delegated by an alternative ISP. This change has to be performed in a transparent fashion with respect to transport and application layers, in order to preserve the established communication, since current transport layers identify the endpoints of a communication through the IP addresses of the nodes involved. The multihoming mechanism located in the SHIM6 layer translates the address used for exchanging packets (namely *locator*) according to the available providers, while always presenting a constant address (*identifier)* to the upper layers of the stack. As a result, the SHIM6 layer performs a mapping between the identifier presented to the upper layers and the locator actually used to exchange packets on the wire. Note that a given address can be simultaneously a locator and an identifier, since it can be used for packet forwarding and also be presented to the higher layers. Both ends exchange the information about alternative locators using a multihoming protocol between the SHIM6 layers.

3.1. SHIM6 security architecture

The security architecture proposed for the multihoming protocol is based in the use of cryptographic addresses such as CGA (Cryptographically Generated Addresses, [7]). CGA incorporate into the 64-bit interface identifier (II) a cryptographic one-way hash of a public key and P_{CGA}, a prefix owned by the node, creating a binding between this public key and the resulting address.

$II_{CGA}=hash_{|64}\ (K_{public\_key} \mid P_{CGA})$

The CGA is built appending the resulting CGA interface identifier to the CGA network prefix: $P_{CGA}::II_{CGA}$.

The private key corresponding to the K_{public_key} used to generate the address can sign the alternative locators that are conveyed in the SHIM6 protocol exchange described later. The trust chain is as follows: the identifier used for the communication is securely bound to the key pair, because it contains the hash of the public key, and the alternative address is bound to the public key through the signature. This chain provides hijacking protection, requiring at least $O(2^{59})$ operations to impersonate a given address if the key pair is strong enough.

3.2. SHIM6 protocol

The SHIM6 protocol [3] is used to create and manage the SHIM6 context associated with the communication between two end-points, and then to be able to exchange data packets using different locators while preserving the established communication. Additional protocols, as defined in [8], are used to detect failures affecting the currently used path, and to explore alternative paths and select among them the most appropriate one to divert the communication to.

We next describe the context establishment and data packet exchange functions of the SHIM6 protocol.

3.2.1. SHIM6 context establishment. Consider the case where one of the parties involved in a communication decides to create a SHIM6 context in order to benefit from the enhanced fault tolerance capabilities of multihoming. We refer to the party that decides to initiate the SHIM6 context creation process as the *initiator*, and the other party involved in the communication as the *receiver*. We assume that at least one of the parties involved in the communication is multihomed. The multihomed host(s) has generated a CGA and has signed with the private key a set of its

locators that belong to some of the multiple prefixes available in the multihomed site.

The initiator requests the creation of a SHIM6 context associated with a pair of identifiers, at least one of them being the CGA. The initiator issues an *I1* message to provide some form of Denial-of-Service attack protection by allowing the receiver to refuse the creation of any context-related state until the initiator has proven its location through this preliminary packet exchange. This message just informs the receiver about the initiator's will to establish a SHIM6 context. Upon the reception of this message, the receiver does not create any state, but it simply replies with a *R1* message.

Once that the initiator has received the *R1* message, it sends an *I2* message that contains the pair of identifiers, the locator set available at the initiator, and the *context tag* that will be used to identify data packets sent with alternative locators, as it is detailed in next section. The *I2* message includes the parameters associated to the initiator's CGA, if available, i.e. the public key of the initiator and its prefix, and the alternative locators signed with the private key. It can also convey a preference specification for each one of the locators exchanged in the form of a *Locator Preferences* option.

Upon the reception of the *I2* message, the receiver verifies the initiator's identity. If the initiator is using a CGA, the interface identifier of the CGA must be the result of the hash of the public key received and the prefix carried in the *I2* message, and this *I2* prefix must also be the same as the prefix of the CGA address used to exchange the SHIM6 protocol information. If these verifications are successful, the receiver creates the SHIM6 context using the received information, and it replies with a *R2* message, in which the receiver includes, if it is also multihomed, its own locator set, its own context tag, the corresponding validation information for its identifier, and optionally the Locator Preferences parameter. The initiator verifies then the receiver's identity with the parameters received in the *R2* message. If the verification is successful, the initiator associates the received alternative locators to the SHIM6 context state and the SHIM6 context establishment process is finished.

3.2.2. Exchanging data packets in SHIM6. As it has been commented above, the SHIM6 layer performs the translation between the identifiers and the locators used for exchanging packets. For that purpose, it has to properly identify the packets that need to be translated. While a locator change is not required, the address included in the data packet is used as both identifier and locator, as it occurs in normal IP operation. However, if the locators are changed for an established communication, because of an outage or a TE policy, the initial identifiers have to be preserved when interfacing to upper layers. Considering that all the addresses available in a multihomed host can be used both as locators and as identifiers, and that it is possible that the same address is being simultaneously used as a locator in one communication and as an identifier in another one, the addresses contained in the received packets are not enough to identify a particular communication. In order to overcome this difficulty, additional information needs to be carried in the packets themselves. The communication to which incoming packets belong is indicated by the context tag included in the data packets into a SHIM6 Payload Extension Header when the locators are different from the identifiers. An *Update* message can be sent at any time to inform about new locators or to modify the preferences.

3.2.3. Failure detection and recovery. The failure detection procedure gathers information from several sources to determine if an outage has occurred. In particular, it considers feedback from upper layer protocols, ICMP error messages, and keepalive SHIM6 specific messages used to probe the reachability of the destination through a given pair of locators.

If a failure is detected, probes are sent using different pairs of the available locators as destination and source addresses to find new valid paths. The host then selects for the next data packets to be sent one of the locator pairs that have been acknowledged by the correspondent host. It is not required to use a single locator pair for both directions of the communication.

4. Traffic Engineering in multiaddressed sites

With respect to TE, the multiaddressing configuration greatly modifies the situation currently available in IPv4. We first present the elements that determine the ingress and egress path from a multiaddressed site, and we then discuss the solutions for TE enforcement for a basic multiaddressed scenario in which a SHIM6 mechanism is not available, showing the limitations of this scenario.

The comparison between path selection capabilities for the current IPv4 solution based on BGP and the IPv6 one based on multiaddresing is the following

- *Outgoing traffic selection.* In the current IPv4 solution, any of the outgoing paths can forward packets that contain a source address with the prefix assigned to the multihomed site. In the multiaddressing scenario, packets have to flow through the ISP associated with the prefix of the source address in order to avoid being discarded by ingress filtering.
- *Incoming traffic selection.* In the currently deployed IPv4 multihoming solution, each multihomed host usually is assigned a single IP address, and there are multiple paths available in the interdomain routing system to that particular address. TE is then performed

by proper configuration (more or less specific routes, AS prepending, etc.) of the multiple routes available in the routing system for that address. When multiaddressing is adopted, the multihomed site is reachable through a given route/ISP only through the proper prefix, so in order to reach the multihomed site through a given ISP, the correspondent prefix/address has to be used in the communication. This implies that the ingress path used is determined by which locator is used among the multiple addresses available for a host.

Therefore, the ISP used by packets to egress or ingress for a given site is determined by the local address selected as identifier for the communication. This address remains unchanged for the lifetime of the communication if SHIM6 is not used. For a non-SHIM6 communication, TE can only be enforced by influencing in the address selection mechanism at the beginning of the communication. However, different considerations have to be made to externally and to internally initiated communications.

When a host outside the multihomed site attempts to initiate a communication with a host within the multihomed site, it obtains the set of destination addresses, and it selects one according to RFC 3484. It seems then that the only place where the multihomed site can express TE considerations is through the DNS server replies. The DNS server can be configured to modify the order of the addresses returned to express some form of TE. When the host receives the list of addresses, it processes them according to the rules specified in RFC 3484 to express its local preferences. If none of those rules applies, the list is unchanged and the first address received from the DNS is tried.

For internally initiated communications, the exit ISP for both incoming and outgoing traffic is determined by the source address included in the initiating packet. This means that the source address selection mechanism defined in RFC 3484 determines the exit ISP. RFC 3484 defines a DAS Policy Table that can be configured to express TE considerations.

These mechanisms are limited in several ways. The most relevant limitation is that sites are no longer capable of defining their own TE policies, but these policies are fully determined by the hosts. While end-sites could deploy tools for uploading TE policies in the form of DAS table to the hosts, or as DNS configurations to the corresponding DNS server, this is not always feasible due to the internal organization, and it is not an option for a higher level provider of the multiaddressed end-site whose connectivity with the Internet depends also on multiaddressing. Therefore, a comprehensive set of mechanisms enabling sites to gain control on TE enforcement in a multiaddressing scenario are required.

5. Traffic Engineering with SHIM6

In this section we present a set of tools for enabling the enforcement of TE for SHIM6 data traffic by site routers. The main objective pursued is to enable TE enforcement at the routing system, avoiding complex configuration in the hosts, and so that updates in TE policies could be easily applied.

The main issues to consider for routing-based TE enforcement are the following:

- For outgoing packets, the source address set by the internal host determines the possible outgoing providers, since ingress filtering in external ISPs restricts the source addresses acceptable to the prefixes that have been delegated by those ISPs.
- For incoming packets, the destination address set by the remote host determines the incoming path.

A solution restoring TE enforcement to the site administrator must

- Return to the routers the capacity they have in the IPv4 multihoming model for selecting the path that egress packets will follow, overcoming the restrictions imposed by ingress filtering within an appropriate security framework, and
- Provide the capacity of influencing in the destination address that determines the path for incoming packets.

We will propose tools for both problems in the next sections.

5.1 TE mechanisms for outgoing traffic

We propose a model in which internal routers are configured to select the egress provider taking into account the destination address for SHIM6 data flows. Prefix rewriting for source address is used to avoid the discard of packets due to ingress filtering. The mechanism does not depend on any particular host (either the internal or the remote) initiating the communication. We first consider which packets can be policed in this way, and we next describe the solution.

In a simple multiaddressing scenario without SHIM6 support, it has to be enforced that the provider used for a data packet is the one that delegated the prefix of the source address carried in the IPv6 header. Otherwise, the packet is discarded. To ensure that packets are forwarded to the proper provider, Source Address Dependant (SAD) routing at the site can be deployed [1]. When SAD is used, the site maintains in each internal router as many different additional routing table instances as prefixes has been assigned. The prefix of the source address of the packet is used to determine the routing table instance, and

proper configuration of each table ensures that the provider corresponding to the source address is selected.

When SHIM6 is used, most outgoing data packets do not need to follow the path that the source prefix is determining. In this case, the internal routing system can use TE policies to determine the egress provider, avoiding the discarding of packets due to ingress filtering by rewriting the source address of the packets at the internal egress routers.

It should be noted that even with SHIM6, some outgoing packets must be forwarded through the providers specified by the source address selected by the host. For example, when a failure occurs through the outgoing path in use, the SHIM6 layer could detect it and initiate the exploration of the paths available through the alternative locators. In this case, SHIM6 Probe Messages are issued to check reachability through different locator pairs. These Probe Messages should be routed through the providers corresponding to the source prefixes selected by the host. Moreover, if a valid locator pair has been found in the address pair exploration, subsequent data packets should be routed through the paths specified by the source prefixes set by the host. Additionally, applications could also require the enforcement of a given path that has been identified in a locator pair exploration as more convenient in terms of packet drop rate or delay. For this communications, site TE policies should be overridden, as it is discussed below.

To be able to rewrite the source address for outgoing data packets, the SHIM6 Payload Extension Header, which contains the context tag used for SHIM6 flow identification, is required to be included in all data packets belonging to a communication for which a SHIM6 context has been established, instead of only requiring it for the packets for which a locator has changed. A *Source Rewriting Enabled* bit is included in the SHIM6 Payload Extension Header to express that source address rewriting can be performed, because the communication has not suffered from failures and the host does not require any specific path for this communication. This bit cannot be modified in transit.

The packets that carry a SHIM6 Payload Extension Header with a Source Rewriting Enabled bit set are forwarded using a new routing table instance, additional to the ones defined for SAD routing. In this internal routing instance, internal egress routers inject routes to propagate reachability to some external prefixes. By proper configuration of the preference of these announces, an administrator can control the amount of traffic that is directed to a given exit link, therefore defining the TE policy of the site for egress packets. This model has been successfully applied for enforcing TE in IPv4 sites with BGP capabilities.

To ensure that source address rewriting can be performed properly, the following conditions must also be honored by the hosts:

- The hosts must have been assigned one CGA per provider prefix. Per each of this CGA a related set of addresses is constructed with each one of the rest of the prefixes assigned to the host and the Interface Identifier of the CGA considered. Then the hosts have the following addresses, being II_{CGA_i} the Interface Identifier generated for prefix P_i:
  $P_1::II_{CGA\_1}$, $P_2::II_{CGA\_1}$, …, $P_N::II_{CGA\_1}$
  …
  $P_i::II_{CGA\_i}$, $P_1::II_{CGA\_i}$, …, $P_N::II_{CGA\_i}$
  ...
  Therefore, if N prefixes are available at a given host, N^2 addresses are configured.
- Communications are initiated using only the CGA addresses ($P_1::II_{CGA\_1}$,…, $P_i::II_{CGA\_i}$,…). These addresses should be the only ones available in the DNS for allowing externally initiated communications, and should be preferred in the source address selection process defined in RFC 3484 for internally initiated communications.
- When an internal host establishes a SHIM6 context for a given CGA, it exchanges the set of locators that share the Interface Identifier with the CGA.

Address assignment rules are also required to guarantee that prefix rewriting can be performed properly. Consider a multiaddressed site for which several prefixes from its providers have been delegated, with different lengths for the significant bits (suppose ISP_J delegated a /40 to the site, while an ISP_K delegated a /42 – see figure 2). For ease prefix rewriting, the site must use the same number of bits for its own address assignment, regardless of the particular prefix used. The number of bits to use is then determined by the most specific prefix delegated (in the previous example, only a /42 would be used from the ISP_J delegation, and prefix rewriting would always be performed for the 42 most significant bits of each address). The site and its clients must also assign the same network bits for all the prefixes assigned for each segment, as it is shown in the internal segment of figure 2, so rewriting of the /42 bits of the prefix generates a valid prefix that corresponds to the same network segment.

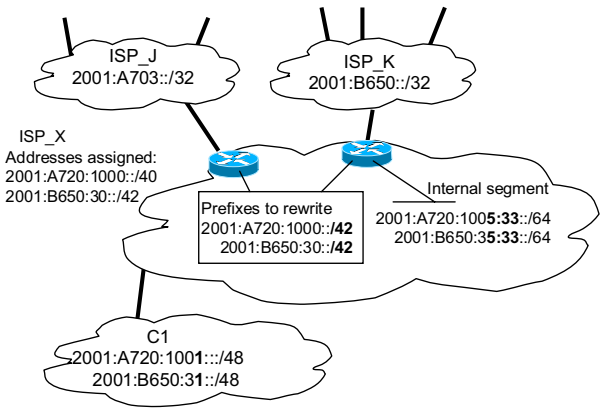

Figure 2. Address delegation for prefix rewriting

If all the conditions stated held, it can be assured that the remote host is aware of all the alternative prefixes that can be used for this communication. Then, the router can change the source prefix to the one corresponding to the provider through which the packet should to flow. This operation is performed without per-communication state in the router.

Note that the SHIM6 security checks assure that the rewriting process can only be performed among the valid prefixes protected by the signature of the internal host, because otherwise the remote host will discard the packet. It can be highlighted that the prefix rewriting operation does not affect to end-to-end identities, thanks to the SHIM6 layer, so address rewriting does not require specific application level packet manipulation at the routers, as it occurs in NATs.

In this section we have presented a prefix rewriting mechanism for SHIM6-enabled hosts using CGA as identifiers. Another cryptographic address type has been proposed for SHIM6 usage, namely Host Based Addresses (HBA, [9]), that allows cheaper identity verification than CGA in terms of computing power [10]. The main obstacle for applying the rewriting process proposed above to HBA is that the locators associated to HBA are currently defined to generate a different Interface Identifier per prefix to provide some level of privacy. Per communication state should be required for prefix rewriting with the current HBA specification, which is highly undesirable. However, a new HBA type could be defined in which the artificial reordering of the prefixes specified to generate different Interface Identifiers per prefix were not performed.

5.2 TE mechanisms for incoming traffic

As it has been said before, in a non-SHIM6 multiaddressed configuration the incoming path is determined by the prefix of the local site address used for the first packet. This can be influenced by proper configuration of the DNS for externally initiated communications, and by proper configuration of the Policy Table used in the Source Address Selection procedure for internally initiated communications.

In a SHIM6 communication the path followed by incoming packets is determined by the correspondent host. After a SHIM6 context has been established, the correspondent host can decide at any moment to select a different locator for the packets it sends to the internal host, either because of a failure or due to any preference. Remember that the locator pair used in SHIM6 is not required to be the same for both directions of the communication, so the change in the locator of the internal host for incoming traffic does not have to be agreed with the host at the local site in any way. Fortunately, SHIM6 provides a mechanism to influence in the choices of the correspondent host, through the optional exchange of a Locator Preferences parameter. This parameter, carried in an *I2* or *R2* messages, can be used to express preference for each locator exchanged through a 16-bit number. The remote host will perform the change suggested by the preference, since in general it is not going to interfere with the enforcement of TE in its surroundings, which is determined by the remote locator used for the communication.

The administrator of a site can influence on the selection of the locators at the remote site by intercepting *I2* and *R2* data packets at the site egress routers, and by adding or modifying the Locator Preferences parameter to express the preferred policies for the site. *Update* messages including the Locator Preferences parameter can also be sent at any time for a communication for which a SHIM6 context has been established to enable dynamic changes in the policies. These operations are stateless, and the criteria for determining the preferences to establish can be based again in the prefix matching for the remote addresses. Note that the Locator Preferences parameter is not signed by the end-sites, so its rewriting does not break the context establishment procedure. The rewriting of the preference can be performed by several providers to express the TE policies of each of the higher level providers.

6. Applying site-based TE enforcement to a site receiving BGP announcements

We present a specific configuration that can provide additional benefits to a SHIM6-based multiaddressing deployment. In this case, the site receives BGP route announcements from its providers, although it is not allowed to inject its own prefixes because it has not been assigned a Provider Independent address range. The information received through the BGP route announcements provide detailed information to the site for providing better fault tolerance and TE support. Prefix

rewriting is used to assure ingress filtering compatibility, and the transmission of preference information can be used for influencing in the path for the incoming traffic.

When an outage occurs in the Internet, and some prefixes are withdrawn, BGP information allows the routers in the site to determine the egress routers that are still valid for reaching the destination prefix. The internal routing then diverts the traffic to the new egress router selected for the communication. In this case, an *Update* message can be sent to the remote host to inform about the new preferred path for receiving traffic. This process can be performed transparently to the hosts, without triggering the costly SHIM6 recovery procedure.

Note that some network failures could not be fixed only with BGP information. Consider for example a failure in the link between a remote multiaddressed site with its provider. Although the provider is announcing through BGP the prefix aggregate that includes the address of the remote site, the site is not reachable through this address. SHIM6 can recover this failure by selecting an alternative operational locator pair. However, BGP would recover from a wide range of failures, although the SHIM6 failure detection mechanism and recovery procedures should not be disabled in general. As a consequence, a reduced rate for the sending of SHIM6 probe packets when BGP information is available could diminish the control traffic generated by SHIM6 while resulting in a robust enough solution.

7. Conclusions

In this paper we have described a comprehensive set of mechanisms for restoring the capacity of the site administrator to enforce TE policies in a multiaddressed IPv6 scenario. SHIM6 data traffic is directed to a given egress router based on the destination address of the traffic. The egress router rewrites the prefix of the source address to prevent from discarding the packet due to ingress filtering at its provider. Some messages belonging to the SHIM6 context establishment process are modified by the routers to transmit to the remote end-host the preferences of the site for incoming traffic. These operations are applied to SHIM6 communications without requiring explicit collaboration of the internal end-host. Some conditions are required for the SHIM6 host, such as the inclusion of the SHIM6 Payload Extension Header in all data packets, and also for the site's addressing assignment. Stateless operation leads to a low impact in router processing.

The applicability of these mechanisms is expected to be broad. First, the scalability concern for the IPv6 routing system is leading to an scenario in which small and even medium sites will be multiaddressed. The end-hosts located in these multiaddressed networks will require a mechanism such as SHIM6 to be able to benefit from the multihomed nature of their site or their providers. While not all the communications will incur in the cost (in terms of local state, computational requirements and messages exchanged) of establishing a SHIM6 context, long-lasting and data intensive communications will do, mainly to obtain fault tolerance. This kind of applications are expected to generate a relevant part of the traffic generated by the site, specially if we consider the high rates required for the transference of multimedia contents or real-time communications. Additionally, a pricing policy rewarding with lower costs the usage by the hosts of SHIM6 rewritable packets could increase even more the amount of traffic that could be conveniently engineered by the site.

8. References

[1] M. Bagnulo, A. García-Martínez, J. Rodríguez, A. Azcorra. "End-site routing support for IPv6 multihoming". *Computer Communications Journal*. Elsevier Eds. Vol. 29, Issue 7, April 2006, pp. 893-899.

[2] C. Huitema, R. Draves and M. Bagnulo, "Host-Centric IPv6 Multihoming", draft-huitema-multi6-hosts-03, Internet Draft, Work in progress, Feb. 2004.

[3] E. Nordmark, M. Bagnulo, "Level 3 multihoming shim protocol", draft-ietf-shim6-proto-03.txt, Internet Draft, Work in progress, Mar. 2006.

[4] M. Bagnulo, A. García-Martínez, C. J. Bernardos, A. Azcorra. "Traffic Engineering in Multihomed Sites". *10th IEEE Symposium on Computers and Communications (ISCC 2005)*, La Manga, Spain, Jun 2005, pp 495-500.

[5] R. Draves, "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, Feb. 2003.

[6] I. Van Beijnum. "BGP". O'Reilly, 2002.

[7] T. Aura, "Cryptographically Generated Addresses (CGA)", RFC 3972, Mar. 2005.

[8] J. Arkko, I. Beijnum, "Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming", draft-ietf-shim6-failure-detection-03, Internet Draft, Work in progress, Dec. 2005.

[9] M. Bagnulo, A. García-Martínez, A. Azcorra, "Efficient Security for IPv6 Multihoming". *ACM Computer Communications Review*, Vol. 35, nº 2, Apr. 2005, pp 61-68.

[10] M. Bagnulo, "Hash Based Addresses (HBA)", draft-ietf-shim6-hba-01. Internet Draft, Work in progress, Oct. 2005.