

November 2002

Moving from 6bone to IPv6 Internet

[draft-savola-v6ops-6bone-mess-01.txt](#)

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

To view the list Internet-Draft Shadow Directories, see <http://www.ietf.org/shadow.html>.

Abstract

Currently, IPv6 Internet is, globally considered, inseparable from the 6bone network. The 6bone has been built as a tightly meshed tunneled topology. As the number of participants has grown, it has become an untangible mess, hindering the real deployment of IPv6 due to low quality of connections. This memo discusses the nature and the state of 6bone/IPv6 Internet, points out problems and outlines a few ways to start fixing them; also, some rough operational guidelines for different-sized organizations are presented. The most important issues are: not offering transit to everyone and real transit operators being needed to take a more active role.

Internet Draft [draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

Table of Contents

1.	Introduction	2
2.	Current 6bone Routing Practises	3
3.	Identifying the Problem Areas	4
4.	Coping with the Administrative Problems	5
5.	Coping with the Connectivity Problems	5
5.1.	Wait for Transit Providers	6
5.2.	Adjust BGP Path Selection or Route Visibility	6
5.2.1.	Example of Route Visibility Control	8
5.3.	Create Hierarchy and Control Transit	9
5.4.	Continue as Before	9
6.	Guidelines for Proper Routing Policy	9
6.1.	Guidelines for End-sites	9
6.2.	Guidelines for Small or Medium-sized ISPs	10
6.3.	Guidelines for Transit sTLA/pTLA's	11
7.	Example of Controlled Transit in 6NET	12
8.	The Role of 6bone in the Future	13
9.	Acknowledgements	13
10.	Security Considerations	13
11.	References	13
11.1.	Informative References	13
Author's Address	14
A.	Some Vendor-Specific Problems with Communities	14

[1.](#) Introduction

Currently, IPv6 Internet is, globally considered, inseparable from the 6bone network. The 6bone has been built as a tightly meshed tunneled topology. As the number of participants has grown, it has become an untangible mess, hindering the real deployment of IPv6 due to low quality of connections.

Some consider that "6bone" is not (current) "IPv6 Internet". This is wrong, and seems to be just denying the problems. Some operators do manage IPv6 on production-level, yes, even connect to some other IPv6 providers, but that is far from something that can be called "IPv6 Internet": IPv6 Internet is the network which is used to reach the vast majority (at least 90-95%) of sites a user would ever wish to visit. That amount of connectivity (however bad, as described later)

is only available through "6bone network".

This memo tries to outline some ways to gain better connectivity for the future IPv6 Internet. This can be done by increasing the quality of 6bone (in some applicable parts) and moving the 6bone network to

Savola

[Expires May 2003]

[Page 2]

Internet Draft

[draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

the edges of the IPv6 Internet (in some other applicable parts).

6bone [[6BONE](#)] can be considered to consist of two, partially dependent parts: addressing and routing. In this memo, we focus mainly on the routing. Additionally, only the routing at the pTLA/sTLA level is considered in detail: sites are outside of scope except for the guidelines section.

It should be noted that as addressing and routing are quite independent, the same problems also exist, in addition to 6bone-allocated pTLA's, with RIR-allocated "sTLA" addresses too. It can often be assumed, though, that those organizations are at least slightly more knowledgeable and serious than an average pTLA holder, due to the process it takes to get an sTLA if you're not an equivalent of Local Internet Registry (LIR). Nevertheless, these are considered as one category.

First, the structure and a bit of history of 6bone network is mentioned. Then, the current problem areas are described. In sections [4](#) and [5](#), some possible approaches to these problems are outlined. In [section 6](#), some rough guidelines for different kinds of organizations are given. In [section 7](#), an example of a policy in 6NET network is summarized. Last, in [section 8](#), the change in the nature of 6bone these would cause is described.

[2](#). Current 6bone Routing Practises

One of the reasons 6bone still exists today, in roughly the same form it existed years ago, is that many big ISP's and transit providers do not support IPv6 yet.

This causes a big problem. Smaller organizations, but still ones big enough to qualify for a pTLA/sTLA, do not have natural IPv6 transit providers: either natively via IPv6 or via tunneling techniques. Other transit providers already supporting IPv6 may be rightfully resistant to taking the burden of terminating tunnels and providing

transit, as there is no customer relationship.

This is the root cause why 6bone network is a mess: as there is no hierarchy, most sites try to gain good connectivity by creating tighter mesh to other pTLA/sTLA's: increase the number of virtual peers by creating tunnels.

The increased number of peers is an alarming fact in itself but it gets worse: most of these pTLA/sTLA's also provide full transit service to every one of their other pTLA/sTLA peers. Often this is done because it "ensures" the others do it as well, one's position in the 6bone network is viewed as significant enough to warrant it, etc.

Savola

[Expires May 2003]

[Page 3]

Internet Draft [draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

[6BROUTING] does not really take any stance on issues like these.

3. Identifying the Problem Areas

One should try to identify which issues are really causing problems.

High amount of BGP peers is not bad in itself, but leads to extreme scaling problems of grade $O(N^2)$ (or at least $O(N)$) with meshed topology. If one only advertised your own routes through multiple BGP peers but provided no transit, there would be no significant problems, except with unnecessary scaling. Some hierarchy would be required.

Using IPv6-in-IPv4 tunneling is also not so bad in itself, if used properly; for example, tunneling over resilient IPv4 connectivity within one AS could be seen as highly advantageous to using dedicated circuits for IPv6 under some circumstances. However, the problem is that tunneling hides the natural, underlying topology: a link may actually consist of many hops, going through many different administrative and technical entities. As such, it becomes very difficult to debug and notice the real problems when using "long" tunnels. Tunneling can also be used to gain connectivity in the absence of native IPv6; as long as one does not provide transit through such "long tunnels", you can only shoot yourself in the foot.

Using old hardware, slow connections and such is also a potential problem, in particular when coupled with transit (as there are more users, and it may be difficult to tell what others will require from the network). For example, using old routers and e.g. T1/Ethernet

-grade connections (even if dedicated) may be ample for some, but totally inadequate for others. Transit should not be provided unless the infrastructure is good enough to cope with it, as this may only degrade others' performance.

Often the pTLA/sTLA's operators are "hobbyists" and providing transit service is not their main business, and monitoring and developing the network is not one of their priorities: for example, the time it takes for new "prefix list filters" to be upgraded throughout the systems should be an evidence of that. It's often noticed that, perhaps in excitement, some want to try "bigger shoes" than they currently have: this is problematic, as most do not have experience or capabilities to deal with the new responsibilities properly.

Full transit, especially in conjunction with points above is the worst problem: there are no longer useful metrics to choose best paths (as AS-path length, the most important global metric in BGP path selection, becomes next to useless), and the network topology gets extremely complicated and unstable. Some metric can naturally be

assigned locally, with usually significant effort, but these are indeed often only locally useful.

4. Coping with the Administrative Problems

Some of the problems are more of an administrative nature. It is clearly visible that under current pTLA assignment policy [[6BROUTING](#)], there are many organizations which are not significant-enough ISP's, transits or such but are still getting pTLA's. This is a double-edged sword: in the absence of IPv6 support by their upstreams, it could be argued it's good that at least someone will do something about IPv6 (note, extending this argument, they should also provide the same services as their ISP would -- ie. not only reap the benefits but also take up the responsibilities); on the other hand, these organizations usually have neither skills, equipment nor customers to understand the operational aspects of being part of a world-wide network, and are the ones worst burdening the 6bone.

Therefore, as one step, the requirements for a pTLA should be made stricter by revising [[6BROUTING](#)]; the revised edition should also take more stance on what kind of organization would be eligible for an allocation (a difficult thing to define precisely, to be sure).

However, as there is an effort to merge pTLA allocation effort with RIR's, this might be just enough to handle this problem for future applicants.

Another thing to consider would be whether pTLA allocations could be revoked from "irresponsible" operators (even more difficult thing to define as before), possibly as a consequence to not conforming to possible new guidelines for pTLA's. This might eliminate some specific problematic operators, but it might be that then the focus would just shift to the others; in general, removing existing allocations is a very harsh measure and it is improbable that it could be done except for special reasons.

5. Coping with the Connectivity Problems

There are a few main approaches to the connectivity problems described above:

1. Wait until a significant number of transit providers support IPv6 and only try to get anything stable then
2. Try to devise mechanisms to adjust BGP path selection to take the problems into the account

Savola

[Expires May 2003]

[Page 5]

Internet Draft

[draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

3. Try to create some, partial hierarchy and control transit while waiting for transit providers
4. Continue as before, and establish even better connections by getting more peers and participate in the "arms race"

These are discussed in order.

5.1. Wait for Transit Providers

Eventually, transit providers should get interested, especially if enough ISP's start bugging them, asking for (and possibly even paying for) the service.

However, at least latecomers may not understand how the current 6bone

network works (or really: does not work), just connecting to it and being happy: repeating the old mistakes. This seems to be happening regularly. A more active role will be required of full-fledged transit providers if they wish to provide good quality IPv6 connections.

Also, this leads to waiting. As it is, IPv6 connections around the world are, generally speaking, very poor. It is irresponsible to turn on IPv6 by default on e.g. operating systems, as that would only lead to a lower quality for the user: for example, if a website is IPv4/6 dual-stack, trying to use IPv6 by default is often much slower, and should be avoided until the network is in a reasonably good shape.

This is how the network will eventually evolve, but the author believes some solutions must be done in the meantime to make it easier to really start using IPv6: currently, it is often a painful experience.

It should be noted that even though some big transit providers do support IPv6 already, the connections to other IPv6 networks is lacking: "walled gardens" look nice on powerpoint presentations but do not really help in global IPv6 deployment.

5.2. Adjust BGP Path Selection or Route Visibility

Some might even suggest BGP specification modifications (e.g. taking a lower layer end-to-end latency into account), but those are definitely considered out of scope: only operational measures are discussed.

One could try to devise some (more or less) common rules to relay the information about real properties of the paths, for example:

- o Use AS-path prepending (e.g. 1..5 times) every time external tunneling is used, e.g. one for every (roughly) 50 ms or something.
- o Use Multi-Exit-Discriminator (MED) attribute to convey a more fine-grained version of this: MED is not really usable in itself, in the 6bone context, but some pre-defined values could be used to signal the quality to the peers.

- o Use Community tagging to convey even more fine-grained policy, like often used by IPv4 transits today.

If some of these were used in the global context, it would be necessary to have everybody upgrade their policies. Experience has shown that this is difficult, and likely to only shift traffic around to those (even worse) organizations that neglect modifying the policies.

Therefore, it would seem likely that one must not depend on global applicability, but rather, employ it in a some smaller scope (like a region, country, ...).

Intuition would suggest that AS-path prepending would be easiest fix for some of the problems because it requires not so much coordination when used properly. However, if some most responsible parties perform prepending (e.g. for long tunnels they have), it is likely that the most irresponsible just end up advertising the non-prepend paths: this can be avoided by only applying the mechanisms in smaller scopes where everyone can be considered responsible, or applying even longer prepending for parties one wishes to avoid using. However, without coordination, these adjustments are only usable with peers connected to the local AS; the policy could be extended using e.g. communities.

MED is next to useless in itself: with "always-compare-med" -like option, it can be used to influence the path from between two routes of equally long AS-path -- but the path length was deemed a next-to-useless metric in 6bone, so this might only help if AS-paths are very short, e.g. 1 or 2 AS's.

MED could also be used for some signalling, but communities are really better suited for the task. They can be used to convey more complex policies too, but doing so will require coordination.

It seems clear that using MED is generally useless when networks are operated as they should be ("production" scenario of connecting to a neighbor AS with more than one peering is naturally excluded). That leaves us with AS-path prepending and communities.

In addition to the implemented example of 6NET communities, see below, one could sketch additional ones, like often used in IPv4. Some possibilities are show below.

Communities consist of <ASN>:<VALUE>; these could be taken to have different applications.

<ASN>:

- o The AS number of organization receiving an advertised route
- o The AS number of organization where a route is advertised

<VALUE>:

- o Do not announce to any external peers
- o Do not announce to any external peers except your direct customers which do not perform transit
- o Announce to external peers only using no-export community
- o Announce to external peers and prepend N (3, 6, ..) times
- o Do not announce this route at all

(Note that there are several vendor-specific problems with no-export communities, see [Appendix A](#) for more information.) There are many possibilities; these are indeed just examples. It is not the purpose of this memo to try to standardize (requiring IANA action) "well-known" communities; but either to give ideas what could be implemented in a non-global scope, e.g. regionally or locally, or come up with simple "if you don't know what to do, you can do this" -rules.

If using the second interpretation of <AS>, there is one particularly interesting case which deserves additional discussion; that is coupled with the last bullet above, this could be used as "If you have a peer in <AS>, do not advertise this route to him at all". This attribute could be transitive (to the extent where communities are sent anyway) or non-transitive (cleared when <AS> is first "blocked"). This could be extremely effective in creating "blockades" around badly behaving AS's.

It should also be noted that this kind of fine-grained policing could prove to be very useful in traditional 6bone routing context -- but that is what we try to escape: simplicity is very rarely gained by adding more complexity.

[5.3.](#) Create Hierarchy and Control Transit

One could try to create at least some hierarchy and remove some peerings; to remove transit as much as possible, or offer it to stub (or at least non-transiting) organizations only.

For some this may be difficult, as people may not want to give up their "toys". Unfortunately, for some others, the "toys" are really guns, some of them loaded.

On commercial front, it seems unlikely that this would happen: competing interests are probably too much. However, IPv6 is still in a start-up phase, so it might be still considered experimentation/research, and this kind of activity would be possible.

But academic side could have some promise. One example policy of 6NET is discussed in a section below.

[5.4.](#) Continue as Before

One should be able to get reasonably good connections by having a significant number of peers (e.g. so that every destination can be reached through 1-3 AS's) and adjust policies to prefer these as appropriate.

But this does not help others. If one wants connections, should one really be required to participate in this kind of "arms race"? This appears to be an extremely harmful and morally questionable requirement.

[6.](#) Guidelines for Proper Routing Policy

There are many ways to properly operate networks: this memo does not try to produce a definite recipe that must be followed. Rather, the subsections below try to bring up issues that should be seriously considered when designing and operating IPv6 networks.

[6.1.](#) Guidelines for End-sites

End-sites, usually obtaining a /48 assignment, do not really need to do much, as their possibilities in affecting 6bone routing are, by design, limited (common prefix filters do not allow them to advertise their specific routes).

The most important things for the end-sites are to:

1. Ask their local ISP(s), both marketing and technical people, for IPv6 if they do not yet provide service: ISP(s) will not usually react unless there is an observed demand, and this will help in making production networks closer to reality.
2. When obtaining IPv6 addresses/routing, if not possible from the local ISP, get them from some provider hopefully as close topologically as possible. For example, trans-Atlantic or trans-European tunnels will just degrade performance.
3. Due to designed restrictions regarding advertising more specific routes, usually obtaining two connections doesn't really help much and may cause problems if the ISP(s) implement ingress filtering. It's better to just use one for simplicity unless there are special reasons to do otherwise.

[6.2](#). Guidelines for Small or Medium-sized ISPs

The definition of "small or medium-sized ISP" is purposely left vague; but often one could characterize them as Internet service or connectivity providers whose main business is not providing transit to other sTLA-level organisations (in the IPv4 world). Vast majority of pTLA's and a majority of sTLA's can be considered to belong to this loose category. However, the organization is always assumed to have an sTLA/pTLA.

This kind of organization is typically big enough to be significant in a regional scope (e.g. country), but not in itself in the global scope.

Therefore the benefit (for the whole Internet) of having such an organization establish peering relations with others all over the world, just for itself and the customers, is small.

These kinds of organizations must seek transit providers to create hierarchy: when there is enough hierarchy, the branch can be considered globally significant. The most natural action would be to find these from the IPv4 upstreams, but that may not always be possible. If not, as above, it's always good to try to underline

that IPv6 transit service is something that's needed. There are some organizations in the community which are often willing to perform transit free of charge, though, too.

sTLA/pTLA's must seek these local transit providers and try to get connectivity through them; after that they can first change their other peers to non-transit mode, and then even remove the non-local peerings completely.

Savola

[Expires May 2003]

[Page 10]

Internet Draft [draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

In this way we achieve:

1. Some hierarchy is created, and these "transits" are less in number and can be interconnected more easily.
2. There is less need to provide "fully meshed" transit to all of your peers as you have your hierarchical transits which handle it except for locally connected peers.
3. A huge number of completely useless, long tunnels can be removed from these "globally insignificant" sTLA/pTLA holders.

At the very least, the following should be carefully considered:

1. Changing most (or some) current peerings from transit to pure peering: advertising own routes with no-export community, and accepting only the peer's AS-number (and possibly its customers) by a simple AS-path access list (or something similar).
2. Disbanding especially "long tunnels", in particular if these include providing transit. If they cannot be removed, some extra measures like AS-path prepending (outlined in sections above) should be implemented.

[6.3.](#) Guidelines for Transit sTLA/pTLA's

"Real" transit-providing organizations should consider whether they could provide limited transit to other organizations too, at least initially. For example, research organizations could be able to provide research-to-commercial and commercial-to-research transit for others too, but fully commercial might be unacceptable due to higher-

level policies.

Organizations whose main business is not providing transit service, but having adequate infrastructure to do it at some scope, could also step up, providing transit until enough "real" transits are available.

Transit providers should be especially careful that they do not make the problem worse: it is important to control the routes sent by the smaller organizations they provide transit for using some mechanisms, like AS-path access lists, prefix lists, or having the smaller organizations add pre-agreed community tagging to mark routes.

Transit providers should also be careful when interconnecting with other transit providers: this should be done natively, or over short tunnels if necessary. Long tunnels, especially if no AS-path

Savola

[Expires May 2003]

[Page 11]

Internet Draft [draft-savola-v6ops-6bone-mess-01.txt](#)

November 2002

prepending is done, should be avoided at all costs.

To re-iterate, especially when providing transit for others, care must be observed so that it is done in a controlled fashion. Transit networks are in the key position when applying and forcing proper policies on 6bone: working together, they might be able to force some of the more "irresponsible" pTLA/sTA's to act in a more sensible manner: together they might have much more power in the arm-wrestling match against the 6bone routing mess.

[7.](#) Example of Controlled Transit in 6NET

6NET [[6NET](#)] is an EU project with about 30 participants, mostly national research networks; there exists native IPv6 infrastructure (initially STM-1) between the partners. In addition, partners naturally have their own connections to other organizations, ISP's, etc.

6NET has implemented a community-based tagging scheme [[6NET-D13](#)] and is in the process of expanding it to accomplish some goals outlined above for transit operators and small-to-medium sized ISP's (national research networks).

The idea is as follows. Most partners probably have a few peers locally (in the country or the region), either natively or through

tunnels. The partner can advertise all 6NET routes (from other partners) to such peers with no-export community. Also, to not to make routes more unidirectional as one has to, the partner also advertises the AS of the peer and possibly the peer's local customers (but not transit routes) to 6NET, for the enjoyment of other partners. Community-based tagging is used to be able to have any partner decide whether it wants to take part in this or not (ie. to discard specific routes or deny his prefixes from being advertised to foreign ISP's), and to ensure traffic from one ISP to another does not leak through the research network.

The advantages should be self-evident: every partner can take advantage of others' good connectivity to others' peers, and there is much less need for tunnels and a large number of peers.

The scheme can be extended to other networks too: to add proper certain communities and preferences to routes learned through other networks (like commercial transits, other research networks, etc.) and that way the benefits of good, native connections can be more widely realized. Perhaps then commercial pTLA's/sTLA's would need a lot smaller number of tunnels and transit services (at least to research organizations).

Something similar could be done in other networks too. The key is to build something that prefers local (hopefully native, but short tunnels are also okay) connections and disallows any uncontrolled transit.

[8.](#) The Role of 6bone in the Future

In the operational community it has often been argued (by the author, too) that 6bone should be disbanded. However, there may be some use for it yet, but mainly for connecting stub or non-transiting sites or pTLA's -- to enable the interested people to get IPv6 address space for testing easily, before getting "production" addresses.

So, the role should change from addressing and routing to only (or mainly) addressing.

A good question is where these stub/non-transiting sites/pTLA's would get transit from -- who would offer them that. There must be

solution for this if this path is to be adopted.

One solution would be getting that from one of their ISP's (or their ISPs' upstreams), or secondarily, from some willing third parties. In some cases this may not be possible though.

Ideas for the ways to proceed would be appreciated.

9. Acknowledgements

The ideas presented have formed during the years at 6bone and discussing the issues with numerous people: many forgotten and too lengthy to list here.

10. Security Considerations

Operational routing practises like these have few, if any, new security considerations. Current practises can be seen to cause a high amount of availability problems, though.

11. References

11.1. Informative References

[6BONE] "6Bone -- testbed for deployment of IPv6",
<http://www.6bone.net/>

[6BROUTING] Rockell, R., Fink, R., "6Bone Backbone Routing Guidelines", [RFC2772](#), February 2000.

[6NET] EU IST-2001-32603, "The 6NET Project -- Large-Scale International IPv6 Pilot Network", <http://www.6net.org>.

[6NET-D13] Mohacsi, J. (Ed.), "Operational procedures to be followed by organizations connected to 6NET", 6NET deliverable 1.3, <http://www.6net.org/publications/deliverables/D1.3.pdf>, June 2002.

[BGPCOMM] Chandra, R., et al, "BGP Communities Attribute", [RFC1997](#), August 1996.

Author's Address

Pekka Savola
CSC/FUNET
Espoo, Finland
EMail: psavola@funet.fi

[A.](#) Some Vendor-Specific Problems with Communities

There are a few vendor-specific issues the author has noticed with the use of community tagging, and they're listed here. If others are aware of any others which might severely impair the use of operational procedures listed in this memo, it is encouraged to bring them to attention.

There are two relatively serious long-standing feature requests in Cisco IOS BGP implementation that will cause some setbacks when using inter-AS communities:

- o When a route with a well-known community (like no-export) is received from an eBGP peer, the route is propagated to iBGP without the community unless "send-community" feature has been configured for all iBGP peers. Note that this behaviour seems to be in violation of [\[BGPCOMM\]](#) definition of e.g. NO_EXPORT_SUBCONFED: "All routes received carrying a communities attribute containing this value MUST NOT be advertised to external BGP peers".

One fix is making send-community the default for iBGP sessions (CSCdk38549).

- o When a route with a well-known community (like no-export) is received, it may be inadvertently deleted or overwritten like any other community (e.g. in the process of scrubbing the routes clean of communities).

One fix is requiring some explicit deletion for these communities (CSCdt21886)

A partial workaround for the first problem (which is assumed to be

much more serious and widespread than the second) would be to use "no-advertise" rather than no-export.

An operative workaround, on the other hand, would be always ensuring that the peers' AS has send-community enabled for all iBGP peers if Cisco equipment is being used.