#36

The TEN-34 Mbone Pilot

Jan Novak

Roberto Sabatino

The TEN-34 Mbone Pilot

Jan Novak
Roberto Sabatino

Introduction

Mbone has been defined in several ways, but the definition which has obtained most consensus within the Internet community is *the part of the Internet that supports multicast.* Hosts wanting to send or receive multicast data are connected to a multicast capable router and this router is then connected to a backbone capable of supporting multicast. Although there are a few examples of networks supporting both unicast and multicast on the same infrastructure (or routers), this is not the usual case. Instead, a set of separate routers has often been deployed in the Internet and these are connected via tunnels between themselves to form a virtual physical network over the underlying unicast Internet.

This paper describes the TEN-34 Mbone pilot service, its motivation, achievements and limitations. In addition, a few words will be spent to explain the next steps DANTE intends to carry out with respect to multicast.

1. The Mbone Model

Today's Mbone is built on the co-operation between the participants in this world-wide project: neighbouring partners set up tunnels between their routers and enable the exchange of multicast routes. As a result, traffic can flow in both directions, however, without any policy control. In addition to the lack

Jan Novak and Roberto Sabatino are Network Engineers at DANTE. Their email addresses are Jan.Novak@dante.org.uk and Roberto.Sabatino@dante.org.uk .

of policy control this very co-operative model has a few shortcomings in terms of:

· technology used
· manageability
· bandwidth usage
· scalability

Technology used:

The Mbone was built and at least half of it is still being built on a distance vector routing protocol (DVMRP) which has a number of well-known serious limitations:

- world-wide non-hierarchical routing domain without a central management,

- uncontrolled distribution of very large routing tables,

- a configuration error on one site can cause disruption to the whole Mbone.

The lack of policy routing is viewed as the most serious limitation.

Manageability:

The co-operative model works very well as long as the Mbone network runs without any problems. However, problems do occur and at times it becomes very difficult to identify and resolve the actual problem. The difficulties may vary from not knowing who controls a specific router to an incompatibility in software versions of neighbouring routers. Although these seem trivial problems to solve, they do impact dramatically the service quality the Mbone users receive.

Bandwidth usage:

Some efforts have been made to map the topology of the Mbone to the underlying unicast topology to optimise the backbone usage. The multicast capable routers are usually deployed outside the backbone, leading to sub-optimal use of both backbone and access links. National research networks are more conscious about access link usage, but backbone link usage is very important too. The following two figures illustrate how backbone links and access links may be used several times to transport the same data - which is what multicast is meant to avoid.

Figure 1. highlights how a backbone link is being used twice to carry the same traffic from the ams-unisource to DFN's and CESnet's multicast routers.
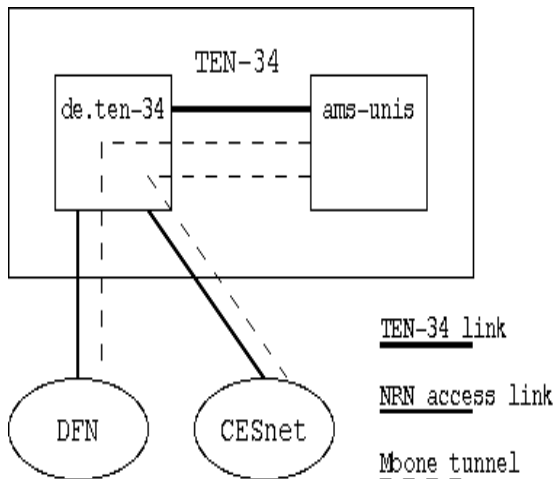

Fig.1: sub-optimal use of backbone links

Figure 2. highlights how the access link for SWITCH can be used three times to send the same data to GARR, ACONET and ams-unisource.
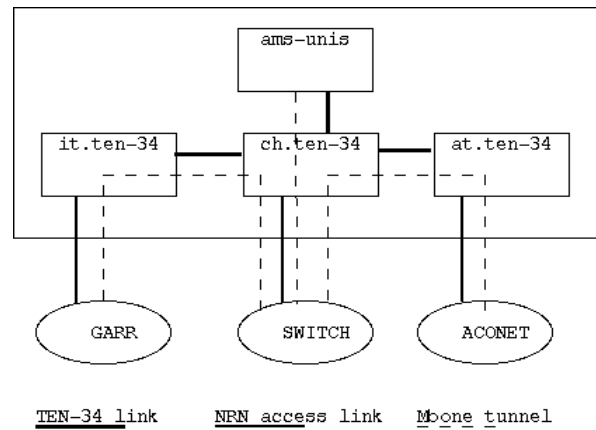

Fig.2: sub-optimal use of access links

Scalability:

When an organisation wants to join the Mbone, typically it asks if one of its Internet neighbours is willing to provide the feed. Normally there is always someone willing to do so, but the model does not scale for the reasons shown in fig.2: a national research network may end up in a situation where a considerable amount of its access capacity is being used for Mbone distribution. In the example shown, a single transmission from a user in SWITCH (the Swiss national research network) uses the SWITCH access capacity 3 times, which defeats the multicast idea. Adding another national research network to the Mbone, and in this case connecting it to SWITCH, would make the situation even worse.

2. A new Mbone topology

In order to overcome these three major shortcomings of the Mbone, DANTE proposed to distribute the Mbone within the TEN-34 network, rather than from within the national research networks. The proposal relied on the availability of work-stations in the PoPs directly connected to the production routers on which mrouted would be run. This has the immediate advantage of optimising the usage of both backbone and access links, as
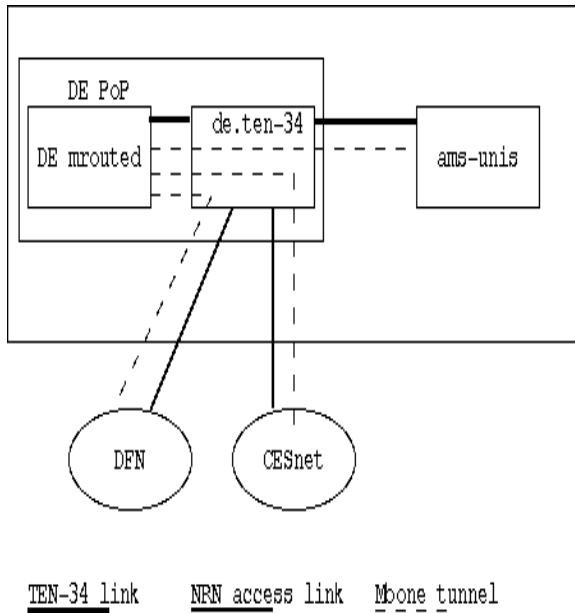
shown in the following two figures:
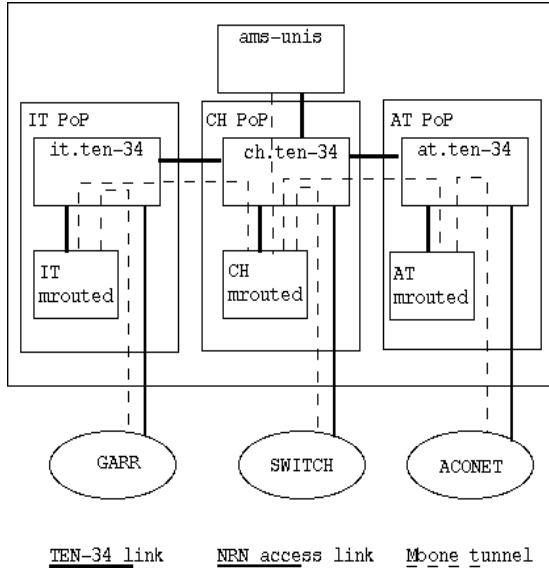

Fig.3: Optimal use of backbone links

Fig.4: Optimal use of access links

The workstations deployed are SUN-ultra-sparc2's running Solaris 2.6 and the public domain software mrouted.

The *scalability* problem has been solved: any national research network wishing to connect

to the Mbone simply has to set up a tunnel to the nearest mrouted, which is in the backbone of TEN-34, thus adding no traffic to neighbouring national research networks. With respect to *manageability*, having one entity (DANTE) managing all the mrouteds to which national research networks connect is a major step forward. National research networks simply connect to the backbone, and in the event of problems with sending or receiving traffic the central managing entity (DANTE in this case) is in the best position to identify and resolve the problem.

Migration from the previous Mbone topology to the topology proposed by DANTE was discussed at RIPE 29 in Amsterdam in January 1998 and was acknowledged as a very positive and practical answer to the well known shortcomings of the Mbone. A TEN-34 document describing the actual migration is available from http://www.dante.net/staff/roberto/docs/RS-98-03.html.

The following figure outlines the topology of the core of the European Mbone and how the national research networks are connected to it:
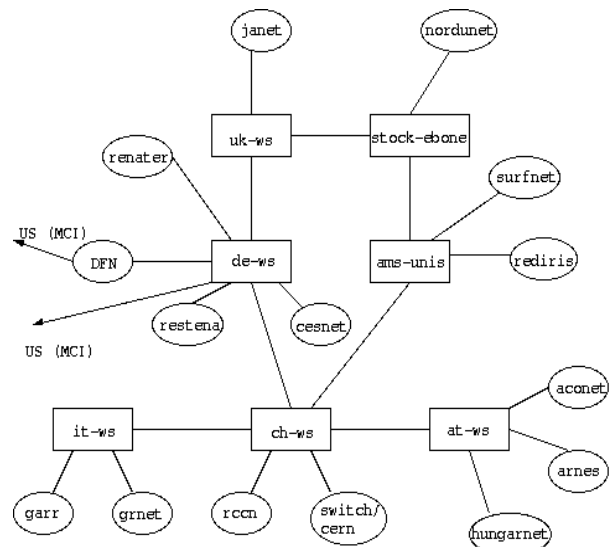

Fig.5: TEN-34 Mbone topology

Effectively the proposal led to DANTE providing an Mbone pilot service to the NRNs,

which has been positively accepted by the community. In order to improve this service (which officially is still in the pilot phase), DANTE decided to add further value by:

· making the topology resilient by adding backup tunnels for national research networks to be used in case the main feed fails,

· providing monitoring tools based on SNMP and by making these publicly accessible.

3. Monitoring with SNMP

During May 1998, Merit's mrouted 3.8.2 with SNMP support was deployed on all the TEN-34 workstations in Fig.5. Effectively this resulted in a downgrade from a higher version of mrouted, but the possibility to collect data from mrouteds via remote SNMP poll gave more overall benefits than the 3.9beta3 version that was previously being used. A recent version of 3.9beta3 with SNMP support has recently been deployed and its functionality is being monitored.

The SNMP agent enables data collection mainly for the following variables: input/output data rates and status of corresponding interfaces, number of routing entries in DVMRP table, the routing table itself, the multicast forwarding table and active multicast groups, information about the neighbours (like version and status) etc.

A debugging tool based on the public domain software Looking glass and Merit's SNMP agent software was written. The user interface of this tool is available at http://www.dante.net/mbone together with the classical mrinfo information, the TEN-34 Mbone topology information and basic statistics collection graphs (DVMRP routes and data rates on each tunnel in TEN-34 infrastructure) using MRTG. With the web interface the user may:

· perform relatively easy debugging (as

compared to ordinary mrouted tools)

· identify the tunnels that are down

· identify multicast groups forwarded or pruned off on different interfaces

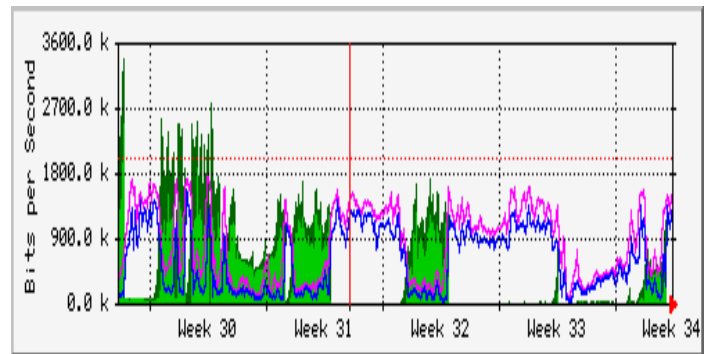Figures 6 and 7 are examples of graphs collected over the same time period.

Fig.6: Input (green/full dark areas) and output (blue line/white areas) on the tunnel from ch-ws.ten-34 to ams-unisource
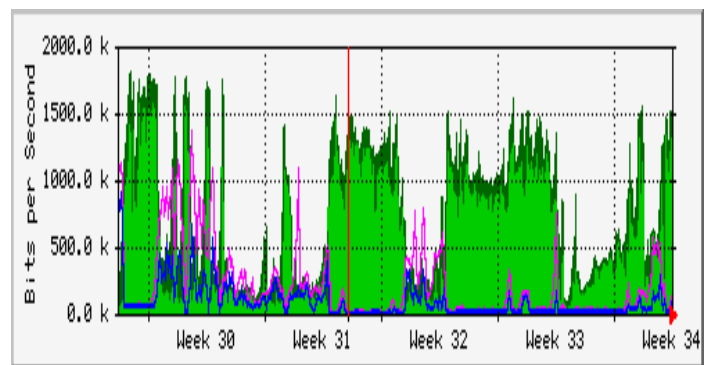

Fig.7: Input (full dark areas) and output (small peaks at bottom) data rates on the tunnel from de-ws.ten-34 to DFN (provides backup access to the US)

The situation with high input data rates on Fig.6 (middle of the weeks 31 and 32) corresponds to the ordinary routing set up in the TEN-34 network whereby the US feed is received from the stock-ebone box in Fig.5.

The situation between week 31 and 32 and after the green/dark peak in week 32 corresponds to the failure of the US feed via Stock-

holm (stock-ebone), with the consequence that the US feed was received via DFN. All these graphs were collected during the MBGP experiments in NORDUnet in August 1998 and before the new DVMRP feed for the TEN-34 Mbone was connected.

4. Limitations of the current model and future work

Despite DANTE's efforts to improve the service quality of the Mbone, the current model of the Mbone has a number of inherent limitations which can only be resolved by reviewing the multicast technology.

The current Mbone infrastructure does not allow reliable and high quality multicast data distribution mainly due to the following reasons:

· the Mbone uses links in direct competition with normal unicast traffic; multimedia/multicast data are encapsulated into unicast IP packets and no preference is given to these packets; the encapsulation/tunnelling adds packet and processing overhead to the transmission which results in a significant decrease of the transmission efficiency.

· There is no policy routing, i.e. it is not possible to control the traffic flows on the Mbone. Once connected, anyone can send or receive traffic to or from anywhere which has direct implications on the amount of bandwidth that can be consumed.

· There is a large number of users connected to the world-wide Mbone that do not implement *pruning* correctly. If a user wants to receive data from a multicast group, it has to join that group. If no users have joined a group by sending the appropriate message to their multicast capable router (mrouter), the mrouter should generate a prune message to its upstream mrouter. If this does not happen, for whatever reason (configuration error, wrong software version, etc.), then that router will continue to receive traffic for a group hence adding unnecessary traffic on the Internet.

The combination of rate limitations on tunnels and the existence of routers that do not implement pruning correctly leads to a large amount of packet-loss which in turn leads to poor (or non-optimal) performance at user level. One consequence is that it often happens that audio sessions are very difficult to understand.

Future work, to be carried out on TEN-155, will include the evaluation of native multicast, together with the evaluation and deployment of MBGP, which allows policy routing with multicast.

The technology is believed to not yet be mature enough (at least for MBGP) to be deployed on production routers for normal unicast traffic, but the intention is to evaluate this on the international testbed provided by the Quantum Test Programme. Once there is confidence in the maturity of the technology, it will be transferred to the production routers.

5. Conclusions

The TEN-34 MBONE pilot service run by DANTE was accepted with enthusiasm by the NRNs as it was the first time that the scalability, manageability and bandwidth usage problems were resolved on a European-wide scale. In addition, the availability of a backbone Mbone pilot helped to identify and in some case resolve a number of problems associated with the technology.

DANTE's plans for MBGP within the Quantum Test Programme should provide another step forward in the European multicasting and contribute significantly in conversion of the worldwide Mbone from the out-dated DVMRP routing protocol to modern

multicast technologies.

6. Acknowledgements

Many thanks go to the staff at DANTE who participated in different ways to the success of the Mbone pilot service, and to the maintainers of the multicast routers in all the national research networks who helped during the migration from the previous topology to the current topology, providing advice and feedback on the performance and troubleshooting expertise.

7. References

http://www.dante.net/ten-34/

http://www.dante.net/staff/roberto/docs/ RS-98-03.html: TEN-98-04, Proposed Mbone topology for European NRNs over TEN-34