

A VPRN Solution for Fully Secure and Efficient Group Communications

Lina ALCHAAL

Netcelo S.A., Echirolles, FRANCE

INRIA Rhône-Alpes, Planète project, FRANCE

lina.alchaal@inrialpes.fr

Vincent ROCA*

INRIA Rhône-Alpes, Planète project, FRANCE

vincent.roca@inrialpes.fr

Ayman EL-SAYED

INRIA Rhône-Alpes, Planète project, FRANCE

ayman.elsayed@inrialpes.fr

Michel HABERT

Netcelo S.A., Echirolles, FRANCE

michel.habert@netcelo.fr

Abstract

In this paper we show how to build a fully secure and efficient group communication service between several sites. This service is built on top of a VPN environment where IPSec tunnels are created, on-demand, between the various sites that need to communicate. This paper is a follow-up of previous work on group communications in a VPN environment and on application-level multicast. We show that these proposals naturally fit with one-another and lead to the concept of Virtual Private Routed Network, or VPRN. This concept enables us to largely improve the data distribution efficiency, and in particular reduces the physical link stress. We are convinced that security is critical in many situations and must be the primary concern of a group communication service.

Keywords: Security, VPN, VPRN, IPSec, Group Communications, Multicast

*corresponding author: Vincent ROCA; INRIA Rhone-Alpes; Zirst; 655 av. de l'Europe; Montbonnot; 38334 ST ISMIER cedex - FRANCE; phone: (+33) 4.76.61.52.16; fax: (+33) 4.76.61.52.52; email: vincent.roca@inrialpes.fr

1 Introduction

Many applications like collaborative work applications and bulk data distribution require an efficient group communication service. It is the only viable solution when network resources must be preserved, either because of their scarcity, of the large amount of data transmitted, or the high number of receivers. If intra-domain multicast (within a LAN or a site) is widely available, this is different for inter-domain multicast. Today many ISPs are still reluctant to provide a wide-area multicast routing service [3]. The important activity around application-level multicasting [4], that most of the time try to offer a pragmatic alternative group communication service when there is no native multicast routing, proves there is an important need. The idea is to build an application-level overlay topology, made of point-to-point tunnels between the group members, over which data is distributed. This is the reason why this approach is also called Overlay Multicast. In the remaining of this paper both names are used indifferently.

But an aspect that lacks in the overlay multicast proposals is security. In a companion paper [1] we have shown how to build a group communication service on top of a fully secure IP VPN (Virtual Private Network) environment. In this paper we show how to further improve the distribution efficiency of this solution and how to reduce the stress laid on the physical infrastructure thanks to the help of application-level multicast techniques.

The rest of the paper is organized as follows: section 2 introduces the IP VPN concepts and our proposal to create a group communication service on top of it; section 3 details both the concept of VPRN and the HBM overlay multicast proposal; section 4 discusses the merge of the IVGMP and HBM protocols in order to create a multicast-enabled VPRN; section 5 introduces a performance evaluation that highlights the benefits of the concept; section 6 introduces related works, and finally we conclude this paper.

2 Offering a VPN Based Secure Group Communication Service

Before describing our solution, we first introduce the VPN specificities, how to build a group communication service on top of it, and how this approach departs from the work carried out in related IETF groups.

2.1 Definition of an IP VPN

An IP VPN [7][11] is an extension of a private network that encompasses links across a shared or public network like the Internet. A secure VPN uses a combination of tunneling and data encryption to securely connect remote users and remote offices. Thus VPNs can replace troublesome remote-access systems and costly leased lines. There are currently three major tunneling protocols for VPNs [11]: Point-to-Point Tunneling Protocol (PPTP), Layer 2 Tunneling Protocol (L2TP), and Internet Protocol Security (IPSec). IPSec [9] has the advantage of offering advanced cryptographic services and proves to be the best security protocol for LAN-to-LAN VPNs (which is what we need), while other security protocols work better for Host-to-Host connections. Besides IPSec is now well known and integrated in many operating systems (e.g. FreeS/Wan for the Linux OS [6]). Therefore our work relies on IPSec.

2.2 A Centralized Approach that Meets Secure Group Communication Needs

We assume that a VPN service provider (or VPN SP) is responsible of the IPSec/VPN deployment and management between the various sites, and that this service provider controls a VPN Edge Device (or ED), a small router with IPSec support, in each site. The Virtual Network Operation Center (or VNOC) is the central point of the service provider that collects all the configuration and policy information and that remotely configures the ED of each site during IPSec tunnel establishment.

This centralized but dynamic approach, which was designed for unicast transmissions, is well suited to our needs. The VPN SP can easily take in charge the group security management aspects (authentication and access control of the sites that want to join a VPN) on behalf of the communication group. The dynamic aspect of the VPN topology (since a site can join or leave a VPN at any time) fits well with the dynamic nature of a multicast group. Therefore taking advantage of the VPN infrastructure to offer a fully secure group communication service seems reasonable. It must be noticed that *security is managed on a per-sites basis, not on a per-node basis, which is reasonable when threats arise from the Internet.*

2.3 Security Versus Scalability

We believe that this approach meets many needs, in particular for the deployment of services in commercial and competitive environments requiring a high level of security. A typical example is a headquarter that needs to distribute a large confidential database to its remote offices. In this case the number of sites concerned

is limited (a few tens), but communications must be fully secure (i.e. the source must be authenticated, the content encrypted and the integrity verified).

In this example, typical of the problem we address, security is the primary concern, not scalability. The number of sites that take part in the VPN is limited, at most a few hundreds and usually only a few tens. Therefore having a centralized approach for VPN management (and also for overlay multicast management as we will see later on) is by no means an issue. Besides, within each site, the number of nodes, senders or receivers, is not limited, which largely increase the effective scalability (in terms of nodes). Finally the scalability in terms of the number of VPNs (rather than the number of sites for each VPN) is a different issue that can easily be addressed by having several VNOC.

2.4 The IVGMP architecture

We now give an overview of the group communication service introduced in [1]. Each VPN ED must implement the Internet VPN Group Management Protocol (or IVGMP) that we propose. The first goal of IVGMP is to discover group members and sources local to this site. To that goal it relies on the Query/Report mechanism of IGMP. When a local host needs to join a new group not already received by this site, the ED informs the VNOC which performs some policy checking. If the site is authorized to join the group, the VNOC then sends back the new IPSec/VPN configuration to this ED and to all other EDs concerned by the VPN. Thus only authorized sites can join a group. Interested readers are invited to refer to [1] for further details.

A limitation though is that traffic replication is performed by the ED attached to the source. Therefore when the number of remote sites increases, the performances quickly degrade. This is the reason why this paper introduces the VPRN concept along with an overlay multicast solution to improve this efficiency.

2.5 An Non-Conventional Approach

Our approach departs from the work carried out in the MSEC IETF working group where the problem is to add security to an already existing multicast routing infrastructure. Our goal is to add a group communication service in a fully secure environment based on IPSec point-to-point tunnels. Scalability aspects, addressed by the MSEC group, is not the primary goal. For instance we do not plan to use our proposal for large-scale video distribution whereas this could be possible with protocols considered at the MSEC working group.

Our work also departs from that carried out in the PPVPN IETF working group where the VPN Service

Provider, in addition to providing a VPN solution, also masters the core network and is an Internet Service Provider (ISP). Group communication solutions developed by PPVPN providers can easily take advantage of their own provider equipments (e.g. IP routers or MPLS-enabled infrastructure) to offer multicast-capable VPNs [12]. In our case the two entities, the VPN SP and the ISP, are different entities. It avoids ISP dependencies and enables to set up a VPN across sites connected to the Internet via different ISPs, without requiring preliminary ISP agreements and mutual confidence.

The priorities and assumptions made in the MSEC and PPVPN working groups are in both cases largely different from ours. Non-surprisingly, the approaches considered differ and in fact complement each other, by addressing different needs.

3 The Benefits of the VPRN and Overlay Multicast Solutions

3.1 The Traditional VPRN Concept Versus our View of a VPRN

Many aspects introduced by VPNs have a direct analogue with those of physical networks. One of them is the way in which VPN sites are connected together and traffic is forwarded. If a fully meshed topology between the various sites is feasible, it is not the only possibility and creating a non-fully connected topology can be highly beneficial in some situations. It naturally leads to the concept of Virtual Private Routed Network, or VPRN, which emulates a multi-site wide area routed network using IP facilities.

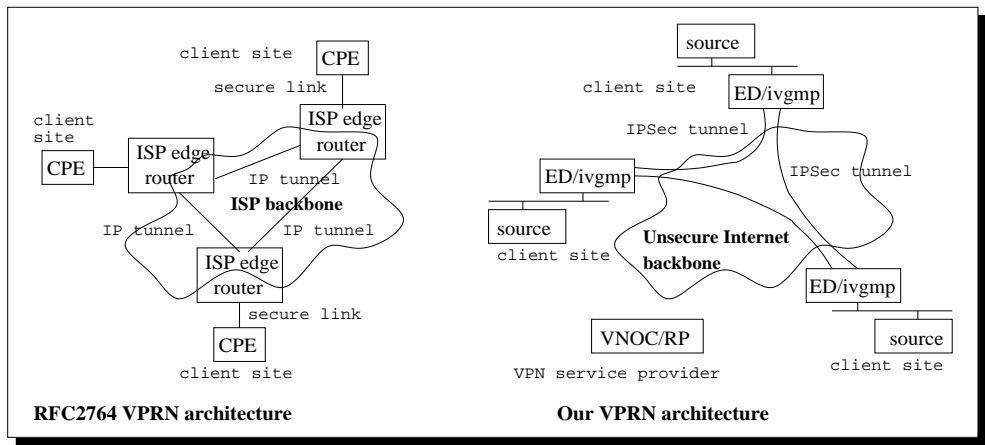


Figure 1: RFC 2764 versus ours VPRN architecture for group communications.

The traditional VPRN model described in RFC 2764 [7] (figure 1-left) considers a provider network as an opaque IP cloud where only nodes on cloud border are part of VPN description; nodes within the cloud are

transparent. Users access the network via a Customer Premises Equipment (CPE) router, which is a router connecting the customer internal network to the provider's edge router, using a non-shared secure link. The provider is responsible for establishing a mesh of tunnels between the provider's edge routers that have at least one attached CPE belonging to a given VPN. This mesh represents a new dedicated network that virtualizes the physical one. Conceptually, there is a dedicated mesh per VPN, and the mesh topology is arbitrary (partially or fully meshed, depending on customer needs). The main benefit of this approach is that it moves the complexity and the configuration tasks from the CPE router to the provider's edge router. Besides the mechanisms proposed intrinsically rely on the features provided by the underlying physical infrastructure (most of the time an MPLS network).

The VPRN model discussed in this paper differs quite a lot from the previous model. In our case (figure 1-right) the Edge Device (ED) located in each customer's site behaves as a VPRN node. The complexity and configuration tasks remain hidden to the customer since these ED are remotely managed by the VNOC. Another difference is that the ED/VPRN nodes have a more dynamic nature (compared to an ISP edge router) and are concerned by group management (e.g. by discovering local sources and receivers with IVGMP). Besides, our VPRN architecture is built on top of a generic IP network, without making any assumption on the underlying physical infrastructure. Likewise it does not need any ISP agreement when sites are connected through different ISPs, which is a big asset.

3.2 The Overlay Multicast Concept and the HBM Protocol

The usual Overlay Multicast goal is to offer an alternative to the lack of deployment of inter-domain multicast routing. Many protocols, largely different in their approach, have been defined, but they all share some similarities that distinguish them from traditional multicast routing [4]:

- A forwarding node in the overlay topology can be either a end-host (i.e. running the application), a dedicated server within a site, or a border router.
- With an overlay topology, the underlying physical topology is completely hidden. A directed (or often undirected) virtual graph is created between all the nodes, and metric measurements taken between these nodes (or a subset).
- In traditional multicast, the membership knowledge is distributed in the multicast routers. With an

overlay multicast, group members are known either by a Rendez-vous Point (or RP), by the source, by everybody, or is distributed among members (e.g. for increased scalability).

- The overlay topology is potentially under complete control. In particular the topology creation process is often optimized using the distance metrics collected between the nodes.

In [13] we have defined the Host-Based Multicast (HBM) protocol. This protocol automatically creates a virtual overlay topology between the various group members, using point-to-point UDP tunnels between them. Everything is under the control of a Rendez-vous Point, or RP. This RP knows the members, their features, and the communication costs between them. He is responsible of the topology calculation and its dissemination among group members.

The data distribution efficiency highly depends on the quality of the distribution tree. This is addressed by the periodic node-to-node (in our case site-to-site) measurements performed by HBM nodes and that are communicated to the RP. This latter then create the distribution topology, using the available metrics. If existing solvers can easily create an optimal topology, in practice the metric database only gives a partial, more or less outdated, view of the networking conditions. Yet we assume that the resulting topology is reasonably good. By default, a shared shortest-path tree is created, but other topologies are possible, for instance a per-source tree when the application is known to be single-source.

4 The IVGMP/HBM Architecture

We have described so far the various concepts and protocols. In this section we describe how they nicely fit with one another.

4.1 General Architecture

The VPN approach considered so far is centralized around the VNOC. Thanks to the IVGMP protocol running on each ED, the VNOC is also responsible of collecting and distributing configuration policies and membership information (in terms of sites) for each multicast group. The HBM protocol also assumes the presence of a central RP which collects membership and distance information, and performs topology creation. Therefore *it is natural to merge the various features and add a RP functionality to the VNOC* (figure 2).

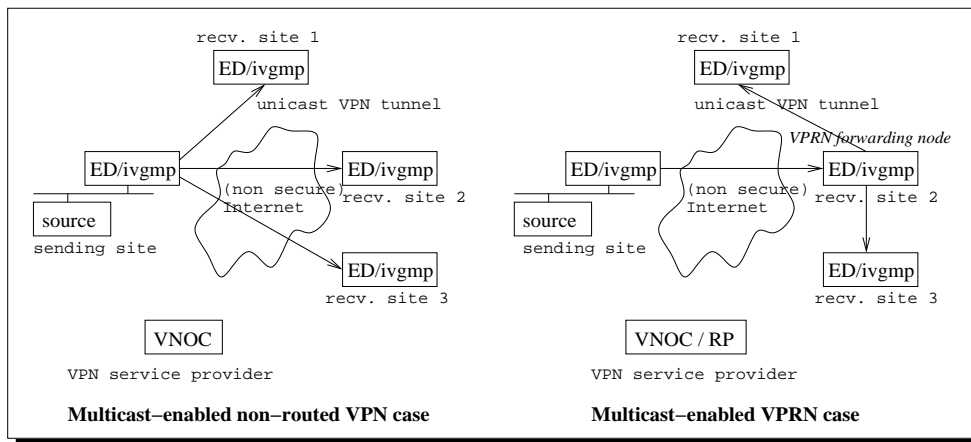


Figure 2: The non-routed versus VPRN approaches for group communications in a VPN environment.

Improved Scalability

Each VPN site can now act as a VPRN node and can forward traffic to its neighbors in the topology. Doing so reduces the fan-out of the site where the source settles, and because it removes a hot spot in the network, the scalability (in terms of number of sites) is significantly improved.

Dynamic Aspects

The group membership dynamic triggers both VPN updates (e.g. by removing tunnels set-up to/from sites that no longer participate in the group) and VPRN/distribution topology updates (e.g. to avoid forwarding traffic to the site that left the group). There is a risk of topology partitioning (and of packet losses) when a forwarding site leaves the group, until the topology is updated. Yet, and this is a major difference with the Overlay Multicast general case, nodes considered here are well administered routers (the ED), instead of end hosts that are far less stable. Therefore the ED departures are almost always negotiated, and appropriate measures can easily be taken.

ED-to-VNOC/RP Security

Our architecture gives security the priority over other services. If site-to-site security is addressed by IPSec, site-to-VNOC (or RP since the VNOC acts as a RP too) communications must also be secure. To that goal, each ED establishes a secure communication channel with the VNOC based on SSL and certificates. The ED and the VNOC first authenticate one another, and then establish a secure SSL connection. Remote configuration and other control operations can then take place, using the SOAP approach [2].

The VNOC must be able to deploy all the IPsec features and ensure that key exchange can be handled properly on the VPN EDs. In order to enable secure communications between two EDs, the VNOC supports a framework for automatic key management, IKE [8]. The IPsec Security Associations (or SAs) generated dynamically by the VNOC, are created between two VPN sites to exchange keys as well as any details on the cryptographic algorithms that will be used during a session.

4.2 Detailed Description

A more detailed analysis of the VPRN/IVGMP/HBM integration exhibits several slight differences with the initial HBM proposal. In this section we detail the operation of HBM in this environment and highlight the specificities related to its integration in a VPRN environment.

4.2.1 ED Functionalities

In addition to its VPN and IVGMP functionalities, an ED now needs to participate in the metric evaluation with the other EDs of the group. The list of such EDs is necessarily present on each ED since IPsec tunnels are created between them. This is in line with the full membership knowledge assumption of HBM where each node potentially knows all other nodes.

By default, metric evaluation consists in issuing `ping ECHO_REQ/ECHO_REPLY` messages within the IPsec tunnels. It assumes the presence of a fully meshed VPN, otherwise some destinations could not be reached. This is a reasonable assumption since each tunnel between two EDs is in fact shared by all the unicast or multicast traffic between them, and there is a high probability that both sites have already exchanged some packets before.

The metrics are periodically and asynchronously collected by the ED, and sent to the VNOC/RP using the secure SSL channel. Once again, using SOAP is in line with the XML approach used by HBM for control messages.

An ED is rather different from the end-host assumed in HBM:

- an ED is rather stable when compared to a traditional end-host (usually a PC). An ED is a well administered router, that rarely reboots or crashes (at least in theory). This feature greatly improves the overlay multicast solution, since node stability, especially in case of a forwarding node, is of high importance.
- an ED is a small embedded PC, usually running a dedicated Linux OS, and has less processing power

than a traditional end-host. The VPRN approach adds some processing on the ED (metric evaluation, packet forwarding), yet most of the work (topology creation, database management and configuration distribution) is performed by the VNOC, not the ED.

4.2.2 VNOC/RP Functionalities

Periodically the VNOC/RP calculates a new topology, taking into account new networking conditions (e.g. a congested path between two EDs can lead the VNOC/RP to find an alternate path). This topology update is also performed in case of membership modification (e.g. when a new site joins the group). The new topology is then communicated to the concerned ED. Each topology update message consists in an updated VPRN configuration, instead of the new list of neighbors of a node as in HBM.

A major difference with the initial HBM proposal is the fact that group departure is by default implicit, since a site always subscribes to a VPN for a limited span of time. Because of this soft-state approach, each ED must periodically subscribe to the group, sending a new JOIN_GROUP message to the VNOC, otherwise the site is automatically removed. This is different from the original HBM proposal which follows an explicit leave model, plus a partition recovery mechanism in case of ungraceful departures (e.g. after a crash). Having a soft-state model enables the VNOC/RP to ask an ED that implicitly leaves a group to keep on forwarding packets until the new topology has been updated.

5 Performance Evaluations

5.1 Experimental Conditions

We implemented both the IVGMP and HBM protocols in C++/Perl on PC/Linux machines, and carried out several experiments to assess the benefits of the VPRN approach. The experiments reported in this paper are simulations based on a large interconnection transit-stub network, composed of 600 core routers, 3rd generated by the Georgia Tech Model (GT-ITM) [16]. Some of these routers are interconnection routers, others, at the leaf of the topology, are access routers connecting the client sites. We then choose $N \in 10; 243$ sites randomly among the 243 possible leaves.

We compared (1) the multi-unicast solution, where the ED of the sending site unicasts a copy of each packet to the $N - 1$ remote sites of the VPN, which corresponds to the non routed VPN approach, and (2) the VPRN

solution where an optimized shared tree is created between the N sites of the VPN. This tree is constrained so that the maximum degree of each node is 6.

5.2 Metrics Considered

The quality of the topology is judged with several metrics [17]. Some of them evaluate the *resource consumption*:

Cost: this is the sum of the *delay* over all the physical links of the distribution topology. With the multi-unicast case, this is the average for all possible sources among the N possibilities since there is a different distribution topology per source. With the shared tree case, by definition the distribution topology remains the same for all possible sources.

Link Stress: this is number of copies of a given packet that cross a given physical link. The ideal stress, only achieved with native multicast routing is 1. The maximum stress is an important information since it highlights the presence of hot spots within the physical network.

while others evaluate *performances*:

Average Delay: this is the average delay between two sites over the distribution topology. With the multi-unicast case, this is the average for all possible sources.

Diameter: this is the maximum delay between the two farthest sites over the distribution topology. The optimal case is achieved with the multi-unicast solution since all paths between the source and receivers are direct.

Each point in the figures (except for the stress distribution figure) is the average over N sites chosen randomly among the 243 possibilities. Note that during simulations, the physical links are only limited by their unidirectional delay (assumed constant), not by their bandwidth, which leads to underestimate the effect of the link stress.

5.3 Results and Discussion

Figures 3 (a)-(d) and 4 (a)-(d) compare performance metrics (average delay and diameter) and resource consumption metrics (cost and link stress) with the multi-unicast and constrained shared tree topologies. Figure 4 (c) is a zoom that emphasizes the tail distribution, but that does not show that most links have a stress at most equal to 5, and 325 links a stress equal to 1.

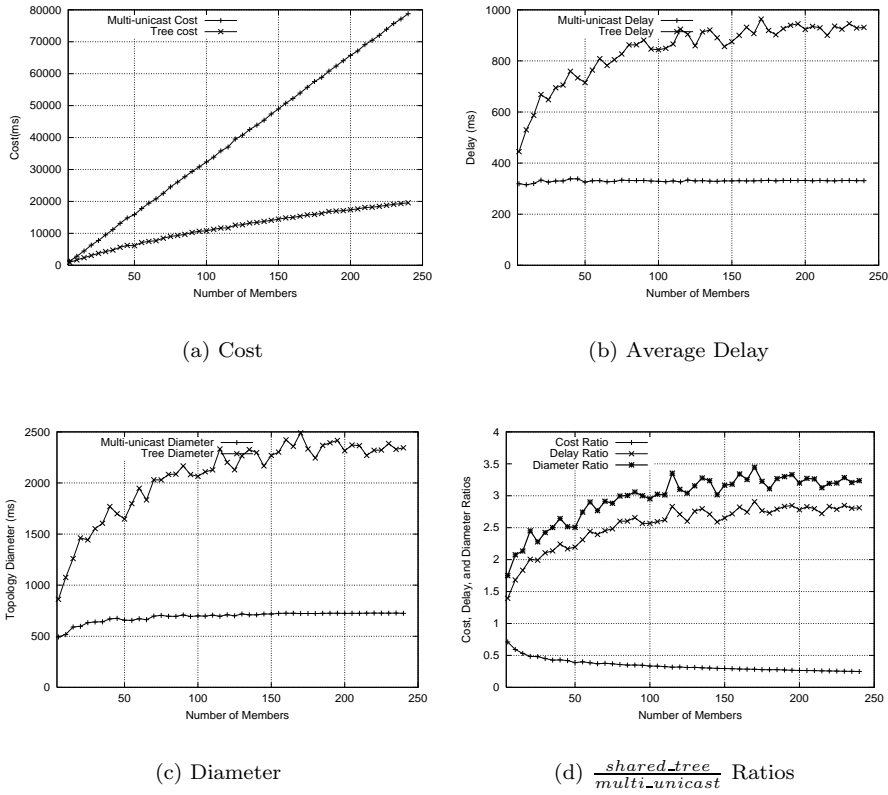


Figure 3: Multi-unicast versus constrained shared tree comparison

The multi-unicast topology is quickly limited by the fan-out of the sending site (equal to $N - 1$) which creates a high stress on the first few links. This is visible on figures 4-(a) (look at the maximum stress) and (d) (stress distribution, showing a tail distribution at $x = N - 1$). Consequently, the multi-unicast approach is definitely not a reasonable solution when there are more than a few tens sites (especially as these tests underestimates the effects of the link stress).

On the opposite the constrained tree, by construction, limits this maximum stress (at most 6 neighbors in these experiments), no matter what is the number of sites, N . The price to pay is a higher maximum delay/diameter of the topology, but experiments show that the stretch factor when compared to the unicast case remains inferior to 3.5 which is fairly reasonable.

Note that in case of a single source application, the distribution topology can be optimized using a per-source tree. This is made possible by the total control over the topology at the VNOC/RP. The only difficulty is inform the VNOC of this application specific feature.

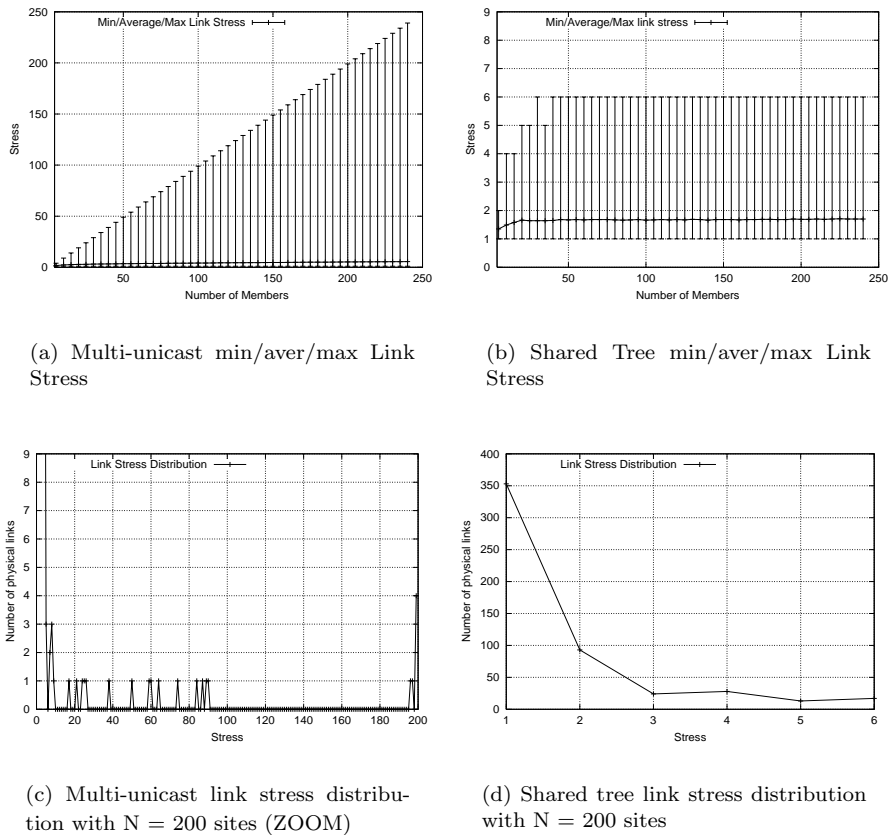


Figure 4: Multi-unicast versus constrained shared tree link stress comparison.

6 Related Works

In this paper we only considered degree-bounded shared tree overlay topologies. Other overlay topologies are possible. For instance [15] discusses several algorithms/heuristics to optimize both the topology diameter and bandwidth usage at the overlay nodes.

[14] describes a solution to offer a multicast service over MPLS/BGP VPNs, using PIM within the VPN and customer routers at different sites. This solution differs from ours by the fact (1) it is aimed to be used in the service-provider backbones specified in [5], while our approach is based on edge technology, and (2) it does not address the problem of using PIM along with IPSec.

As mentioned before, a PPVPN can easily and efficiently offer a group communication service. [12] describes, at a high level, how the VPN can exploit either a multicast routing service in the provider's network, or an MPLS-enabled infrastructure.

Finally, our approach shares some similarities with the Centralized Multicast (CM) approach [10]. In CM, the data forwarding and control functions are kept separated, and the control part is centralized in distinct

control elements. The control elements are arranged in a two-level hierarchy within autonomous systems and are used to set up multicast trees. In our approach too, the control part is centralized in the VNOG. The major difference yet it that CM does not address security.

7 Conclusions

In this paper we show how to build a fully secure and efficient group communication service between several sites. It details both the underlying motivations and the architecture proposed. It is a follow-up of work we performed on offering a group communication service in an IPSec VPN environment [1] and on the HBM application-level multicast protocol [13]. We show that these proposals, that both follow a centralized approach, naturally fit with one-another and lead to the concept of Virtual Private Routed Network, or VPRN. This concept enables us to largely improve distribution efficiency, in particular by reducing the stress laid on the physical infrastructure, over the multi-unicast approach used so far to distribute packets between the sites. It is worth noting that centralized overlay multicast approaches, often criticized, find a perfect field of application in VPN environments, and the security benefits brought by the whole architecture largely compensate any possible scalability limitation.

Finally IVGMP and HBM have both been implemented and simulations carried out to quantify the gains made possible by the VPRN approach.

References

- [1] L. Al-Chaal, V. Roca, and M. Habert. Offering a multicast delivery service in a programmable secure ip vpn environment. In *Fourth International Workshop on Networked Group Communication (NGC'02)*, Boston, USA, Oct. 2002.
- [2] D. Box, D. Ehnebuske, G. Kakivaya, A. Layman, N. Mendelsohn, H. F. Nielsen, S. Thatte, and D. Winer. *Simple Object Access Protocol (SOAP) 1.1*, May 2000. W3C Note, <http://www.w3.org/TR/SOAP/>.
- [3] C. Diot, B. Levine, B. Lyles, H. Kassem, and D. Balensiefen. Deployment issues for the ip multicast service and architecture. *IEEE Network*, pages 78–88, Jan. 2000.

- [4] A. El-Sayed, V. Roca, and L. Mathy. A survey of proposals for an alternative group communication service. *IEEE Network, special issue on multicasting*, Jan. 2003.
- [5] E. Rosen and Y. Rekhter. *BGP/MPLS VPNs*, Mar. 1999. IETF Request for Comments, RFC 2547.
- [6] FreeS/Wan org. *FreeS/Wan project home page: an open-source implementation of IPSEC and IKE for Linux*. <http://www.freeswan.org/>.
- [7] B. Gleeson, A. Lin, J. Heinanen, G. Armitage, and A. Malis. *A Framework for IP based Virtual Private Networks*, Feb. 2000. IETF Request for Comments, RFC 2764.
- [8] D. Harkins and D. Carrel. *The Internet Key Exchange (IKE)*, Nov. 1998. IETF Request for Comments, RFC 2409.
- [9] S. Kent and R. Atkinson. *Security Architecture for the Internet Protocol*, Nov. 1998. IETF Request for Comments, RFC 2401.
- [10] S. Keshav and S. Paul. Centralized multicast. In *7th Int. Conference on Network Protocols (ICNP'99)*, Toronto, Canada, Oct. 1999.
- [11] D. Kosiur. *Building and Managing Virtual Private Networks*. John Wiley & Sons Inc., ISBN 0-471-29526-4, 1998.
- [12] D. Ooms and J. D. Clercq. *Overview of Multicast in VPNs*, Feb. 2002. work in progress, <draft-ooms-ppvnpn-mcast-overview-00.txt>.
- [13] V. Roca and A. El-Sayed. A host-based multicast (hbm) solution for group communications. In *First IEEE International Conference on Networking (ICN'01)*, Colmar, France, July 2001.
- [14] E. Rosen, Y. Cai, D. Tapan, I. Wijnands, Y. Rekhter, and D. Farinacci. *Multicast in MPLS/BGP VPNs*, February 2002. work in progress, <draft-rosen-vpn-mcast-03.txt>.
- [15] S. Shi and J. Turner. Routing in overlay multicast networks. In *IEEE INFOCOM'02*, June 2002.
- [16] E. Zegura, K. Calvert, and S. Bhattacharjee. How to model an internetwork. In *IEEE INFOCOM'96*, Mar. 1996.

[17] B. Zhang, S. Jamin, and L. Zhang. Host multicast: a framework for delivering multicast to end users.

IEEE INFOCOM'02, New York, USA, June 2002.