

BGP CONVERGENCE IN MILITARY COALITION NETWORKS

Ingvild Sorteberg
Baseline Communications as,
Havegt. 2, N-2010 Strømmen, Norway,
email: ingvild.sorteberg@baseline.no

ABSTRACT

The BGP interdomain routing protocol will play an important role in military coalition and national networks. BGP is used for routing between different autonomous systems or administrative domains and offers a well defined interface between networks.

However, the use of local routing policies to control distribution of routing information may cause the network routing to diverge resulting in instable network operations. The use of common guidelines for configuration of routing policy may prevent divergence problems, but at the cost of loss of connectivity in certain situations, even when communication's paths do exist. This is not a favorable situation, since access to communications services is extremely important in military networks and available paths should be accessible.

INTRODUCTION

The Internet routing architecture is based on a hierarchical structure where the networks are divided into Autonomous Systems (AS). An AS describes a collection of networks under a single administration which appears to other ASes to have a single coherent interior routing plan and presents a consistent picture of the destinations that are reachable through it. The most commonly used inter-AS or interdomain routing protocol is the Border Gateway Protocol version 4 (BGP4) [1]. BGP4 is basically a distance vector protocol. Routes are chosen based on the calculation of the shortest path given by the number of ASes that need to be traversed.

BGP is likely to become frequently used in military networks, both between different nations' military networks and inside nation's military networks. Since BGP operates at the AS level, use of BGP in international coalition operations limits the need for network planning and coordination. As an operation proceeds, changes in the number of ASes involved and in the interconnection of ASes can easily be supported. Also from a security perspective BGP offers an attractive solution since it prevents distribution of the AS internal topology outside the AS boundaries.

However, BGP has convergence problems, [3] and [4]. These are mainly due to the support of local policies overriding the shortest path selection. This problem has been reduced in commercial network through the inherent hierarchical structure of different ASes which is given by the provider, customer relations and peering relations. Additionally, a strong self-justice has been enforced, limiting the operators' freedom when configuring local policies and expelling operators causing problems. The result is that the inherent BGP convergence problems have been limited and the Internet BGP routing is fairly stable.

The use of BGP in military networks is more susceptible to convergence problems since there are no clear hierarchical relationships between different ASes. This makes it difficult to configure local policies not resulting in convergence problems. Also ASes may be mobile causing frequent changes to their interconnection points, and thereby changing their relations to the other ASes in the hierarchy.

This paper presents the configuration guidelines for local routing policies which are commonly used in the Internet and which have been proved in [2] to result in stable routing. An example of the use of these policy guidelines in a military coalition network scenario is presented and the consequences in terms of connectivity and flexibility are analyzed and discussed.

The structure of this paper is as follows, first a brief description of the BGP routing protocol is given before the recommended local policy guidelines for commercial IP networks are presented. Afterwards an example of a maritime coalition network is presented. It is shown that this network configuration may experience convergence problems and how it is possible by adopting the local policy guidelines to resolve these problems, but at a cost. We conclude with some ideas of how this may be implemented in order to avoid increased network configuration. Finally some additional ideas for future work which may prove to offer possible solutions to the convergence problems in military networks are presented.

BGP OPERATION

BGP is a distance vector protocol with added policy control and the basic element of connectivity is an AS. A path is given by a list of ASes to traverse.

In BGP every boarder router notifies its neighboring BGP peers whenever there is a change in the network topology. BGP is not aware of the finer levels of the network topology on a link-by-link basis within the local AS or within any remote AS. This prevents each AS from distributing information about their internal structure and topology changes within the AS are not necessarily announced.

The BGP UPDATE messages are used to notify neighboring peers. An UPDATE message includes a list of ASes to be traversed to reach a set of prefixes. In addition, it has information about the origin of the prefix and a set of path attributes.

The shortest AS path is preferred, but an associated dimension of the BGP operation is the use of local policies. Policies give the local administrator the opportunity to control which routing information to accept from the neighboring ASes and which information to announce to its neighbors. For example, if an AS does not redistribute information about a route. This signals that it does not want to undertake the role of being a transit network for traffic towards the prefixes reached by this route. Additionally, the commonly used path attributes Local Preference (`local_pref`) and Multi Exit Discriminator (MED) are also used to implement local policies. For example, the path attribute `local_pref` only has local significance and provides a relative local ranking of routes to the same destination. This enables use of local policies to overwrite the distance based metric by adding a local weight to the different paths. MED is used to discriminate between multiple entry and exit points to the same AS. The path with the lowest MED is preferred. The MED value is not propagated to neighboring BGP peers.

The connectivity information is stored in the Routing Information Base (RIB) and each entry in the RIB refers to a distinct route. The routes may include an attribute list and a set of local policy constraints. This information is used to determine the best path from the local AS to the AS where the route originates. The selection of the best route is done by choosing the routes with the highest `local_pref` value and if several routes have the same value, then the route with the shortest AS path is preferred. If several routes still are equal, additional selection criteria are used. The selection process may differ slightly between different BGP implementations. Finally, the preferred route is stored in the local RIB (Loc-RIB).

PROBLEM STATEMENT

The BGP protocol does not guarantee convergence since each AS administrator may configure its own local policies. The combination of these policies may result in diverging routes as shown in [3] and [4]. Figure 1 shows an example of a non-converging network which has been copied from [3].

To reach destination d in AS 0, AS 1 prefers the route via AS 2 before the direct route. This is done by giving this route a higher `local_pref` value compared to the direct route. Similarly, local policies in AS 2 are configured to ensure that AS 2 prefers the route via AS 3 and AS 3 prefers the route via AS 1. As shown in [3], routes to d will flap between using the indirect routes with the highest `local_pref` and the direct route. Consequently, the network never reaches a stable situation.

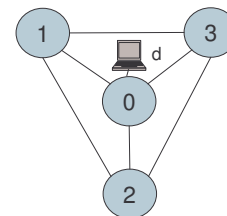


Figure 1: Non-converging network

The BGP route flap dampening mechanism is used to limit the impact of instable routes, but it does not prevent instability. By gradually increasing the time between sending UPDATE messages for the same prefix, the route flap dampening mechanism makes the problem of instable routing run in slow motion.

As shown above, lack of coordination when configuring local policies may result in unreachable networks or instable routes. A combination of the inherent hierarchical structure of the Internet and a set of policy configuration guidelines for Internet Service Providers (ISP) has prevented major instability in the Internet routing.

The hierarchical structure of the Internet is based on the business relationships between network operators. A typical example is that local providers are customers of regional providers which again are customers of national or multinational providers. In addition, ISPs establish peering relations. ISPs that cause routing problems, by not adhering to the configuration guidelines, are prevented from using dynamic routing.

The problem of non-converging routes is not as easily dealt with in military networks. There are no clear provider/customer relations and the fact that networks are mobile and constantly change their interconnection points make it difficult to configure local policies that always result in stable routing. By enforcing a hierarchical structure to a military coalition network topology and applying the commercial policy guidelines, we analyze the consequences of the interdomain routing. Particularly we focus on the cost with respect to connectivity and the need for reconfiguration of policies as ASes move and change connection points to the rest of the network.

POLICY GUIDELINES

In the Internet a hierarchical relationship exists between the autonomous systems. This is based on the customer/provider relationships, peering relations and backup relationships. This inherent hierarchical structure together with a set of guidelines for policy configuration has ensured that the BGP operation results in a converged network.

Formal specification and verification of a set of guidelines for configuring local policies have been presented in [2] and these guidelines will in the next chapter be applied to a maritime coalition network.

A. Import policy guidelines

The following import policy guidelines should be applied to prevent convergence problems:

- Classify routes based on next_hop AS (e.g. customer, peer, provider or backup route)
- Assign local_pref based on classification
 - Customer routes are preferred over peer and provider routes, i.e. the customer routes are given a higher local_pref value compared to the provider and peer routes.
- Allow any local preference (local_pref) within a class
 - It is left to the local administrator to define the ranking between ASes within the same class.

B. Export policies guidelines

These export policies guidelines should be applied to prevent convergence problems:

- Exporting to a provider: In exchanging routing information with a provider, an AS can export its routes and the routes of its customers, but it can not export routes learned from other providers or peers.
- Exporting to customer: In exchanging routing information with a customer, an AS can export its routes and the routes learned from its providers and peers.
- Exporting to peers: In exchanging routing information with a peer, an AS can export its routes and its customer's routes, but it can not export routes learned from other providers or peers.

C. Backup lines

Backup lines are not normally used unless a failure occur and the favored paths are not available any longer. Therefore backup lines should have lower local_pref than other lines.

A MILITARY NETWORK SCENARIO

The policy guidelines presented above assume a hierarchical network structure. A hierarchical structure based on customer/provider relations is not as evident in military networks and at the same time mobile networks will

change their position in the hierarchy and consequently require changes in local policies.

In this chapter, we present an example of a network topology for a maritime coalition operation, based on the ideas presented in [5]. The example is used to show that by enforcing a hierarchy and using the local policy configuration guidelines, the network transforms from an instable, non-converging network to a stable network.

However, this method does cause other problems related to connectivity and there will be a need for mechanisms to ensure a consistent view of the relationships as mobile networks change connection points or the network partition.

A. The maritime coalition network

The maritime network infrastructure consists of national WANs (UK WAN, CA WAN and US WAN), an allied WAN, a Task Group Area Network (TGAN), one or more coalition Maritime Marine Forces (MMF) and coalition Maritime Air Groups (MAG).

The national WANs provide access for the individual nations participating in the operation to the allied networks. It also provides an interconnection point for the national units participating in the operation. The latter is not really needed, but often national units do want to establish direct contact with their national HQ and have additional bandwidth available for soldiers communicating back home. The national WAN can also be used as a backup link to the MMF and MAG networks or the TGAN.

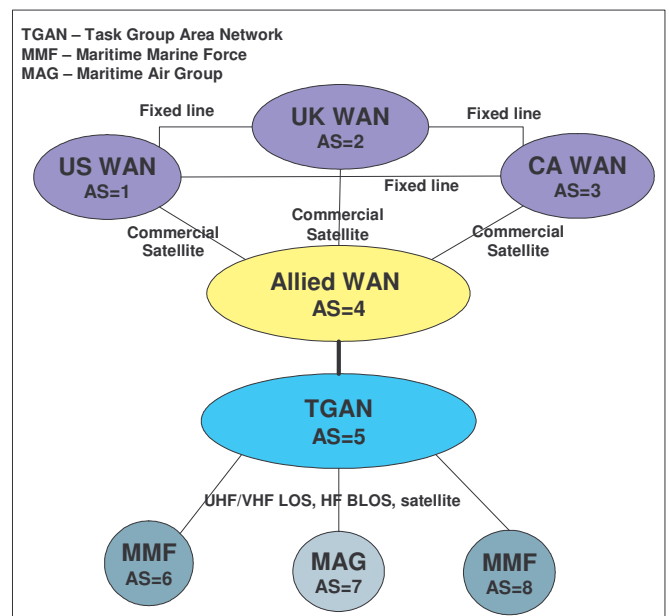


Figure 2: Initial maritime coalition network configuration

The BGP protocol is used for distribution of routing information between the autonomous systems (ASes), see

Figure 2. During normal operation, the route via the allied WAN is the preferred route for communication between the national WANs and the TGAN. For the national CA WAN, this is achieved by configuring the following local policies:

```
as_path = [4,5] => local_pref := default + 2
as_path = [2,?,5] => local_pref := default + 1
true => local_pref := default
```

The preferred route to TGAN via the allied WAN is given a local_pref of 2 + default value. If the route between the CA WAN and the allied WAN fails, the route via UK's national WAN is used to reach the allied WAN. The reason why the wildcard is used in the AS path and not only a direct reference to AS=4 is because the CA WAN administrator wanted to make sure he could use any additional backup lines UK had to AS=5. Similar policies are defined in all the national WANs, UK has an agreement with US to use their connection to the allied WAN in case of their own link failing. Similarly, the US has an agreement with CA to use their link to the allied WAN. This initial configuration results in a stable routing.

If a failure occurs between the allied WAN and the TGAN, the national WANs will try and reroute via each other, but will eventually realize that there are no alternative routes to TGAN and the TGAN is declared unreachable.

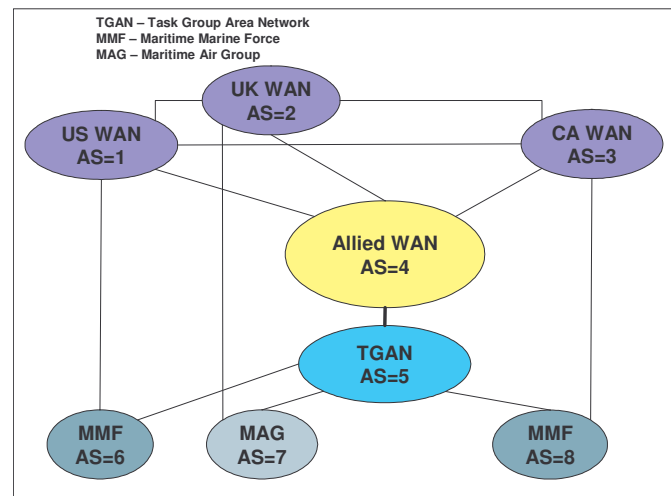


Figure 3: Maritime network topology after the establishment of the MMF and MAG networks

As the maritime operation proceeds, the MMF and MAG networks become operational and establish radio/satellite connections to their national WANs in addition to radio or satellite connections to TGAN, see Figure 3. For the national WAN the configuration is unchanged and the initial backup via the other national WANs are preferred since the connections between the MMF/MAG and the TGAN in many cases have very low bandwidth (BLOS radio). Nor-

mally, the MMF and MAG networks are not used for transit traffic, except in special cases where no other alternatives exist. The network still has a stable BGP routing.

The national administrators have optimized their configuration to handle failures in the connection between the national WAN and the allied WAN. However, other failure situations may occur. If the link between the allied WAN and the TGAN fails, the CA WAN will route its traffic to the UK WAN. The UK WAN network knows that it can not reach the TGAN via its direct route to the allied WAN and uses the alternative route via the US WAN since this has higher local_pref than the direct route via the MMF. The US and UK WANs are also informed about the failure between the allied WAN and the TGAN. Similarly, they switch to their alternative route via the UK and CA WAN respectively since these have a higher local_pref than the direct route via the MMF and MAG networks. This new route is then announced.

Receiving the update message from UK, CA WAN realizes that the new route loops. Therefore the CA WAN switched over using its link via the MMF network. The same thing happens with the other national WANs and a new BGP routes are announced. This causes a new situation, where the CA WAN realizes that new the link via the UK WAN and MAG network has a higher local_pref than its own. The CA WAN again changes its route and announces this. The same thing happens in all the national WANs and the network is unable to reach a converged state even though none of the alternative routes loop.

A similar situation occurs if all the national WANs lose their connection to the allied WAN. This may happen if for example the commercial satellite connecting the national WANs to the allied WAN fails.

The process of flapping between different routes will continue until one or more of the original direct connections to the allied WAN are restored. This route instability will result in a degraded service quality. The BGP route flap dampening mechanism will try to limit the impact of the route oscillation by increasing the time interval between consecutive routing announcements to the TGAN. However, this can not stop the route flapping, but slows down the frequency of the route changes.

B. Deployment of policy guidelines

By applying a hierarchy and using the policy guidelines presented above, it is possible to ensure that the maritime coalition network converges also in the presence of failures. First, a hierarchical structure is established and the relationships between the different ASes must be defined. As seen in Figure 4, the network can be converted into a four tier network.

The national WANs are all tier 1 networks with peering relations. The allied WAN is a tier 2 network and is a cus-

customer of the national WANs and a provider to the TGAN. The TGAN is a tier 3 network and has a customer-provider relationship with the allied WAN. The TGAN serves as a provider for the MMF and MAG networks. The MMF and MAG networks are tier 4 networks. If they are connected to the national WANs or to the allied WAN, they are also customers of these WANs. MMF and MAG networks which are directly connected have peering relationships.

Secondly, the routes are given different local_pref values depending on their classification (customer/provider/peer). Since customer routes are preferred over routes via providers or peers, the following local_prefs will be configured in the CA WAN when using the guidelines presented above:

```

as_path = [4,5] => local_pref := default + 2
as_path = [2,?,5] => local_pref := default + 1
true          => local_pref := default

```

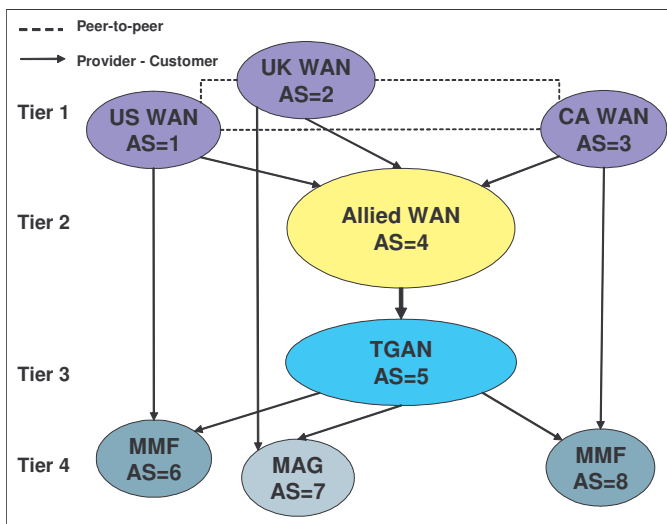


Figure 4: Hierarchical structure of maritime network

The alternative route via the UK network is kept as before and given a higher local_pref value compared to the route via the US national WAN. Similar configuration of local_pref values are done for all the national networks.

By deploying the export policy guidelines presented above, the following local policies apply:

- Routes distributed between national WANs (to peers) only include customer routes, i.e. routes received from the allied WAN and from the MMF and MAG networks directly connected.
- Routes distributed from the national WANs to the allied WAN (to customer) include all national WAN routes.
- Routes distributed from the MMF and MAG networks only include their internal routes

Traffic from a national WAN to the TGAN goes via the directly connected allied WAN. National WANs without a direct connection to a MMF network or MAG network will always route traffic via the allied WAN since customer routes are preferred over peer routes. For networks with direct connections to one or more MMF and MAG networks, two customer routes exist. In this case, local policies can be used to direct the traffic to these MMF and MAG networks either via the allied WAN and TGAN path or over the direct path. If local policies are not configured, the direct route is preferred since it offers the shortest AS path.

Initially, this gives the same situation as in the previous case. However, in case of a failure between the allied WAN and the TGAN, the network will be able to reach a stable situation. The MMF and MAG networks can not distribute routes received from the TGAN. Therefore, the TGAN will never be reachable via the MMF and MAG networks. The result is that BGP will first try to reach the TGAN via another national WAN, but this results in a routing loop which will be detected and BGP will conclude that no viable alternative paths exist. The outcome is that a failure in the connection between the allied WAN and the TGAN or between the national WANs and the allied WAN will cause the TGAN to become unreachable.

The consequence of applying the policy guidelines is that the network will be stable, but the routing may not be optimal seen from an operational point of view since some networks will not be available even if a path does exist.

In situations where the only available routes to TGAN are via the MMF and MAG networks, applying a hierarchical structure and using the policy guideline may limit availability. This is not according to the requirements of always being able to exploit all possible options in order to reach a network or user. A possible cure to this problem is to install additional backup links to prevent networks from becoming unreachable in case of single link failures. For example establish a backup line between the allied WAN and the TGAN or between the TGAN and some of the national WANs, but with limited resources this may not always be possible.

C. Policy guidelines and mobile networks

Another problem with applying a hierarchical structure and using the policy guidelines in military networks is mobile networks. Particularly the MMF and MAG networks will change their connection points as the operation proceeds. An example is shown in Figure 5, where MMF3 loses its direct radio contact with the TGAN and instead uses the MMF2 for transit to the TGAN.

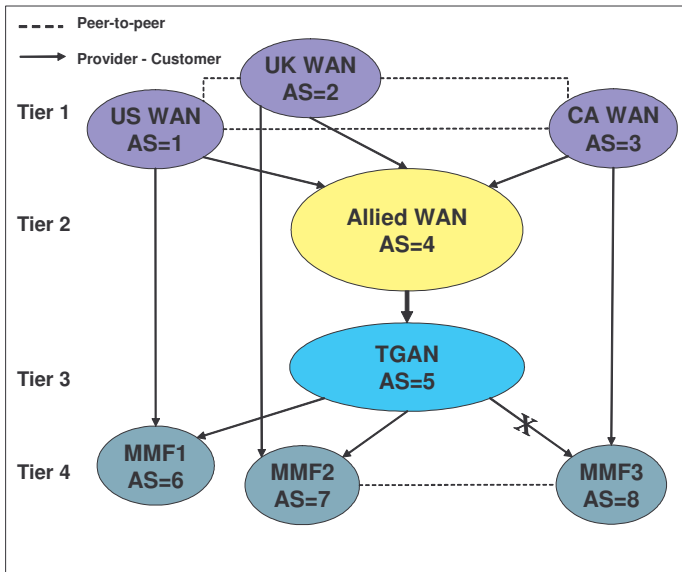


Figure 5: Initial topology when the link between MMF3 and TGAN fails

Previously the two networks were peering networks. However, this would prevent MMF2 from distributing routing information about the MMF3 network and MMF3 would no longer be reachable via the TGAN. For the MMF2 to be allowed to announce the MMF3 routes, the MMF3 network must be a customer of the MMF2 network. As a result of MMF3's move and its loss of the direct radio connection to the TGAN, MMF3 will become a tier 5 network, see Figure 6. The MMF2 network will have to reconfigure its distribution policies for the MMF3 network prefixes.

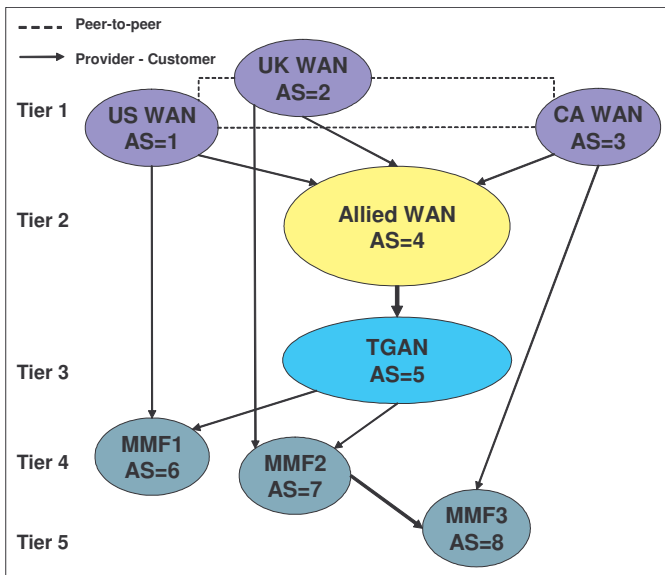


Figure 6: Topology after reconfiguration of local policies in MMF 2 and MMF3

DISCUSSIONS

The interdomain routing protocol BGP will play an important role in military networks, both within coalition and national networks. The BGP protocol, although being one of the cornerstones of the Internet, may cause stability problems due to inconsistent configuration of local policies. It has been shown in [2], that by exploiting the inherent hierarchical structure of the Internet and adhere to a set of configuration guidelines, convergence can be ensured. This paper analyzes the consequences of applying this method to a military maritime coalition network.

The maritime coalition network was built up of different ASes. Initially the routing converged and resulted in a stable network operation. However, during certain failure situation, the network experienced route instability due to non-converging route policies. By structuring the network into a hierarchy and applying the policy guidelines, it was shown that the network became stable even during failure. The cost was that connectivity to all parts of the network could no longer be ensured.

The problem of networks becoming unreachable even in situations where paths do exist is difficult to handle unless a flat network structure is deployed and with it the problems of non-converging routes. Possible solutions to this problem will be to change the BGP protocol, either by extending it to carry policy information allowing every BGP node to check for policy consistency and report non-consistent policies to the network administrator. This is a processing intensive task and with frequent topology changes it will put heavy burden the system. Another approach is to defer from the use of local policies and only rely on the use of a pure distance vector algorithm. This will ensure network convergence, but does not support the need of local administrators to control routing based on for example security policies, bandwidth availability or other needs. The use of a link state routing protocol between administrative domains is also a possibility. This ensures a stable network, but at the cost of having to expose the topology of the AS internal network. Also use of link state protocols in large networks may require a hierarchy in order to scale.

Another problem with deploying a hierarchical structure is that military networks frequently experiences topology changes due to mobile networks and network partitioning.

The result of this is changes to the hierarchy and the inter-network relationships. Initially, the network planning process may define a hierarchical structure, but this will not be possible to sustain over time. This will require re-configuration of the local policies when mobile networks change their interconnection point and thereby they position in the hierarchy. In operations with a high degree of mobile networks, it may be difficult to maintain a consis-

tent view of the network hierarchy and guarantee that the local policies are correctly configured.

Therefore, mechanisms for distribution of relationships may be needed. One possible solution is to use a central registry to hold information about the network relations. Whenever relations change, this must be communicated to the central register and the networks affected by it informed about the changes. This solution has been proposed for commercial networks. The problem with this solution is that it is difficult to realize in a highly dynamic military network where the network topology changes frequently and networks partition and loose contact with central servers.

Another approach may be to extend the BGP protocol to support distribution of relationships. This information can be used to check the consistency of the local policy configuration and may contribute to automating the process of local policy configuration. Exactly how this would be done and the pros and cons will have to be analyzed further.

CONCLUSIONS

No simple method exists to ensure consistent local policy configuration. The solution presented in this paper, even when ensuring convergence, causes operational problems and loss of connectivity. Therefore, it is not a recommended solution for highly dynamic networks. However, it may pose the most likely short term solution to the interdomain routing problem unless network administrators are deferred from using local policies.

In the long term, more research is needed to extend the existing BGP protocol or define a new interdomain routing protocol which is more suitable for highly dynamic networks.

FUTURE WORK

An area for future research is to define methods or protocol extensions for distributing and updating the network hierarchical relationships.

Use of the route flap dampening mechanisms to control instable routes do not differentiate route instability due to non-converging routes or route instability caused by for example lossy radio links or mobile networks. We are planning to do a set of simulations to get a better understanding of the effects of this mechanism and see if it may be tuned and optimized.

ACKNOWLEDGEMENTS

I would like to thank my colleague Øivind Kure for valuable comments and Thales Communication Norway which have sponsored parts of this work.

REFERENCES

[1] Y. Rekhter and T. Li (editors), *A Border Gateway Protocol 4 (BGP-4)*, RFC 1771, March 1995

- [2] Linix Gao, Jennifer Rexford, *Stable Internet Routing Without Global Coordination*, IEEE/ACM Transactions on Networking, Dec 2001, pp. 681-692
- [3] T. Griffin, G. Wilfong, *An Analysis of BGP Convergence Properties*, Proc of the ACM SIGCOMM, Aug 1999
- [4] K. Varadhan, R. Govindan, D. Estrin, *Persistent Route Oscillations in Inter-Domain Routing*, Technical Report USC CS TR 96-631 Department of Computer Science, University of Southern California, Feb 1996
- [5] R. Govindan, C. Alaettinoglu, G. Eddy, D. Kessens, S. Kumar, W.S. Lee, *An Architecture for Stable, Analyzable Internet Routing*, IEEE Network, 13(1):29-35, 1999
- [6] ACP 200, *Maritime Tactical Wide Area Networking (MTWAN)*, AUS, CAN, NZ, UK, US (Unclassified), 2003