3GPP TSG SA WG3 (Security) –S3 #31 **S3-030741**
18-21 November 2003
Munich, Germany

Agenda Item:	**6.10 – WLAN Interworking**
Source:	**Nortel Networks, Siemens AG, Nokia**
Title:	**End-to-end tunneling: Security Considerations on resolution gateways**
Document for:	**Discussion and Decision**

1. Introduction

SA3 has received a liaison from SA2 (S3-030676) asking for feedback on the security properties of a proposed architecture for end-to-end tunnelling for WLAN Scenario 3.

This contribution provides an analysis of the security properties of this architecture and some alternatives/options within it. We focus primarily on the question of which packets are admitted onto the inter-PLMN backbone network, since this has been the main focus of concern in WLAN discussions within SA2. This corresponds to concerns about potential DoS attacks in which packets from a WLAN Interworking system cause disruption of inter-PLMN traffic (either inter-GSN traffic or other WLAN Interworking traffic) or resources connected to the inter-PLMN backbone (GSNs and PDGs).

2. Working assumptions

SA2 have taken a number of working assumptions for the Scenario 3 architecture for WLAN. We briefly examine the consequences of these from a security perspective.

The working assumptions are:

- An end-to-end tunnelling architecture: the UE transparently establishes a tunnel over the inter-PLMN backbone directly towards a Packet Data Gateway in the HPLMN
- Tunnel establishment is independent of the initial WLAN Access Authentication and Authorisation (in particular, it is independently authenticated and authorised)
- After tunnel establishment, filters can be provided to the VPLMN which allow classification of traffic associated with a single tunnel in the VPLMN

The first consequence of these assumptions is that it is not possible to reliably identify the source of packets at the border of the inter-PLMN backbone (the WAG in the VPLMN). This is because there is no cryptographic association between the UE and VPLMN. In particular, the source of Tunnel Establishment packets cannot be verified by the VPLMN – they must always be admitted to the inter-PLMN backbone for authentication at the PDG. (Verification of packet source based on source IP address does not work without assumptions about anti-spoofing measures in the WLAN and more importantly the absence of NA(P)Ts).

It should be remembered, though, that only 3GPP WLAN UEs which have been authorised for connection to the WLAN should be able to send packets to the WAG. So, whilst the exact source of packets cannot be identified, it is at least known that they are from 3GPP WLAN UEs which are authorised for WLAN access (rather than from the public internet, for example).

The second consequence is that it is possible to apply a policy which by default blocks tunnel data packets. After successful tunnel establishment, the filters provided to the VPLMN allow a path for these packets to be opened.

3. Confidentiality of a PDG address

It seems reasonable to assume that there will be no more than a handful of PDGs in a PLMN, and that the number of users allowed to access one of these PDGs could be very large. It may also be assumed that PDG addresses will not change very frequently. This makes it very likely that PDG addresses could not be kept confidential for very long:

It is easy for an authorised user to discover the PDG address. He could then pass the address to an attacker – wittingly or unwittingly. Therefore, authentication of a user, either by the WAG, or a new element such as the resolution gateway RGW described below, would not help to preserve the confidentiality of a PDG address. Consequently all analysis of DoS attacks should start from the assumption that PDG addresses are publicly known.

Further, we re-iterate the conclusion of SA3's response to SA2's previous liaison on PDG addresses (S3-030475). Specifically:

"SA3 believes that hiding the IP address of the PDG on GRX using NAT or other techniques would not be useful from a security point of view."

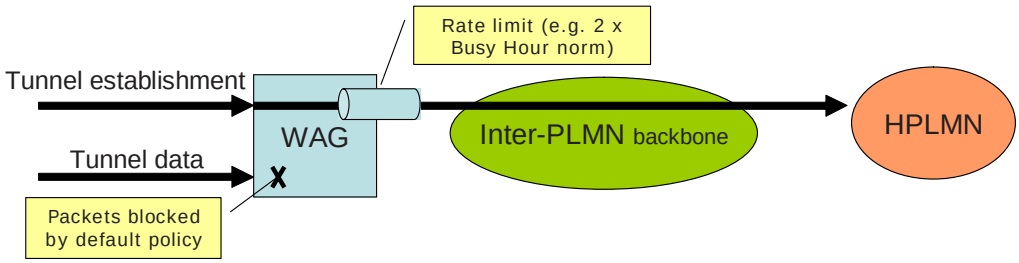
4. DoS mitigation

As described in the attachment to S3-030428, measures are required at the border of the inter-PLMN backbone to mitigate DoS attacks. This can be done by limiting the rate at which traffic of a particular type is admitted to the inter-PLMN backbone such that traffic through a single WAG can never be sufficient to cause a Denial of Service.

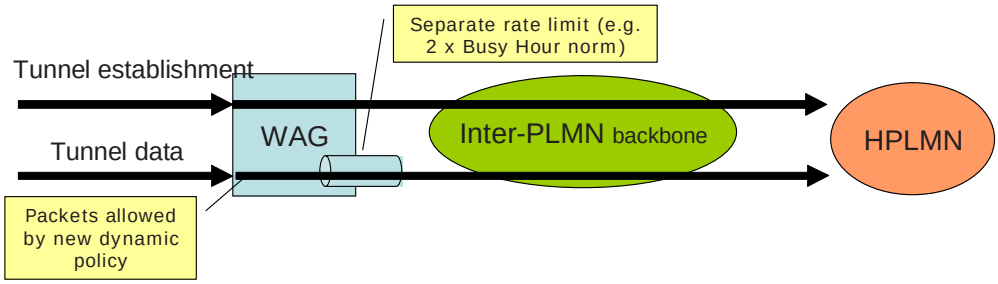
Based on the assumptions above, such limits can advantageously be applied separately to Tunnel Establishment traffic and Tunnel Data traffic as shown in the figure below:

Note: it is assumed that packets which are not addressed towards a PDG, or which are not recognised as Tunnel Establishment or Tunnel Data packets are always blocked by the WAG.

Before/during tunnel establishment



After tunnel establishment



Further attempts have been made during discussions within SA2 to improve on the above situation – for example by authenticating and authorising the user at the WAG before providing the address of the PDG to the user. However, these succeed only in making it more difficult for an attacker to discover the PDG address – cf. Section 3.

The fact remains that the eventual end-to-end tunnel establishment cannot be authenticated at the WAG without breaking the end-to-end assumption. It can be seen, therefore, that the Inter-PLMN backbone and the PDG in the HPLMN are exposed to Tunnel Establishment messages from the WLAN. In fact the inter-PLMN backbone becomes very much like a DMZ in this respect.

Further, whilst the dynamic filters at the WAG block Tunnel Data packets before Tunnel Establishment, afterwards they will admit any traffic with the correct IP header fields. These could be spoofed, although it is difficult to obtain the correct values at the correct time. Such spoofed packets would be quickly detected at the PDG and the Tunnel aborted as a result. Although this constitutes a DoS attack against the legitimate owner of the aborted tunnel, a much easier attack exists simply by flooding the WAG with packets.

We re-iterate that in SA3's response (S3-030477) to SA2's previous liaison on Denial Of Service attacks (S3-030428), SA3 agreed with the conclusions of the attached paper. Specifically:

- Rate limiting of tunnel establishment messaging at each WAG is sufficient to prevent traffic from any single WAG overloading a PDG or disrupting other users of the inter-PLMN backbone
- Separate rate limiting of tunnel data messages has a similar effect for this tunnel data traffic
- With such rate limiting measures, the exposure of the inter-PLMN backbone and HPLMN is no worse compared to interworking UE messaging in the WAG to AAA signalling – assuming authentication in the Home Network.

The above measures are adequate from SA3's point of view

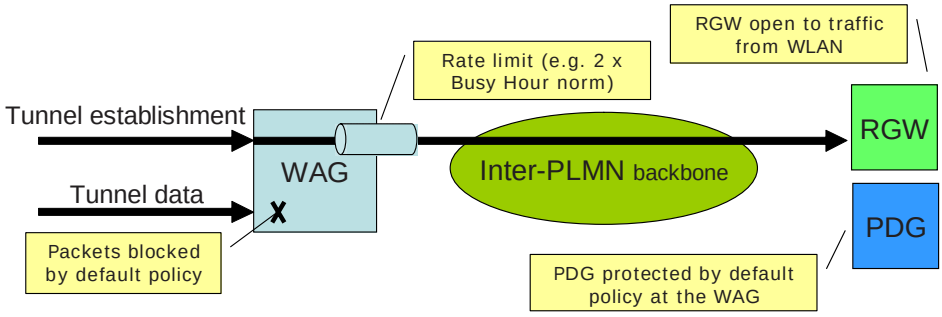
5. Functional decomposition of PDG

In the liaison from SA2 (S3-03xxxx, S2-033813), the possibility of separating the PDG into two components is suggested. The two components would be:

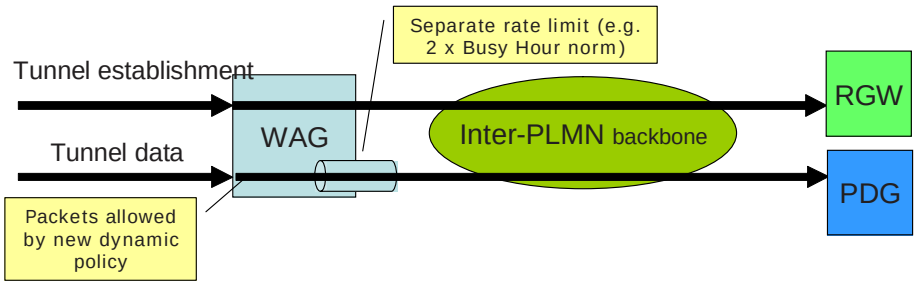
- The 'W-APN Resolution Gateway', which processes tunnel establishement messages, and
- The 'Packet Data Gateway', which processes tunnel data messages

The consequence of this separation is illustrated below:

Before/during tunnel establishment



After tunnel establishment



In summary, the PDG is protected from 'unauthorised' traffic from the WLAN until after Tunnel Establishment. It has been suggested that this results in less stringent robustness requirements on the PDG and is therefore a security advantage.

Additionally, the address of the PDG is not supplied to the user until they have been authenticated and authorised, which is seen by some as an advantage (but see comments in (3) above).

6. Comments on the proposal

We make the following comments on the proposal outlined above and in the SA2 liaison:

- Tunnel data packets will be addressed to a different Destination Address from Tunnel Establishment packets. As a result, they may appear to come from a different Source Address due

to presence of a NAPT. Any filters for Tunnel Data cannot therefore be based on Source Address, greatly simplifying the task of spoofing Tunnel Data packets for any attacker.

- The proposal represents a departure from the operation of standard VPN concentrators, which might be expected to be adapted for the PDG function – it therefore increases the expected costs of PDGs. In fact, this seems the main concern with the solution proposed by SA2. The most economical solution from an implementation point of view may in fact be to set up two tunnels, one between UE and RGW and one between UE and PDG. But this would clearly be undesirable from a performance point of view. There may also be interoperability problems if tunnel re-direction solution was allowed to co-exist with the two-tunnel solution as the UE would have to know which solution to apply.
- The proposal relies on separating the physical resources which deal with the 'unauthenticated' tunnel establishment messaging from that which deals with tunnel data. The objective is that overload of the former will not affect the operation of the latter. However, such segregation of resources is possible within a single piece of equipment as an implementation choice.
- The proposal suggests that a single RGW may serve many PDGs. However, such a RGW then becomes a single point of failure
- The rate limit applied at the WAG for Tunnel Data traffic is expected to be much higher than that applied to Tunnel Establishment packets. It is not clear that the relatively low volume of tunnel establishment packets is something that the PDG needs to be 'protected' from.
- The proposal requires the parameters of an IPsec security association, including the keys derived during tunnel establishment, to be passed from RGW to PDG, which although possible, may have security implications which are not usually considered in the design of tunnel establishment protocols.
- It is not clear what the protocol for the transport of security association parameters should be.
- Furthermore, the security association transferred from the RGW to PDG would have to be “patched” by replacing the RGW’s IP address with that of the PDG. Consideration would also have to be given to the SPIs. In order to ensure uniqueness of the SPIs at each PDG, the RGW would have to maintain SPI state across all PDGs. It is unclear how this could be achieved as the RGW would not be notified by the PDG about the deletion of an IPsec tunnel. If the PDG patched the SPI it is not clear how the new SPI could be communicated to the UE.
- The fact that the PDG address is not supplied to the user until after authentication/authorisation is a minor advantage, since it does not affect the vulnerability of the device, it just places an additional (small) hurdle in the way of an attacker

7. Conclusion

SA3 should communicate the following conclusions to SA2:

Based on the above considerations, we conclude that the proposal to physically separate the RGW and PDG has at most only marginal security advantages. We suggest that these advantages are not commensurate with the additional cost and complexity introduced. In particular, SA3 is concerned about the lack of available solutions for a separation of tunnel establishment protocol endpoint and tunnel endpoint.

As a result, a model in which users obtain a PDG address through DNS and establish a tunnel with this PDG using standard IP VPN procedures is acceptable to SA3 from a security perspective. SA3 notes that this model could result in a user requesting tunnel establishment to a Visited Network PDG although access to visited services is not allowed, and suggests that the request should simply be denied in this case. Afterwards the UE may establish a tunnel to a Home Network PDG by using an appropriate W-APN.