                      Security Requirements of NVO3
                 draft-ietf-nvo3-security-requirements-00

Abstract

   This draft discusses the security requirements and several issues
   which need to be considered in securing a virtualized data center
   network for multiple tenants (a NVO3 network for short). In
   addition, the draft also attempts to discuss how such issues could be
   addressed or mitigated.

Requirements Language

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in RFC 2119 [RFC2119].

Copyright Notice

Table of Contents

1. Introduction

   Security is the key issue which needs to be considered in the design
   of a data center network. This document first lists the security
   risks that a NVO3 network may encounter and the security requirements
   that a NVO3 network need to fulfill. Then, this draft discusses the

essential security approaches which could be applied to fulfill such
requirements.

The remainder of this document is organized as follows. (Section 4)
introduces the attack model of this work and the properties that a
NOV3 security mechanism needs to enforce. Section 5 describes the
essential security mechanisms which should be provide in the
generation of a NVO3 network. Then, in Section 6, we analyze the
challenges brought by the new features mentioned
in[I-D.ietf-nvo3-overlay-problem-statement].

2. Terminology

This document uses the same terminology as found in the NVO3
Framework document [I-D.ietf-nvo3-framework] and
[I-D.kreeger-nvo3-hypervisor-nve-cp]. Some of the terms defined in
the framework document have been repeated in this section for the
convenience of the reader, along with additional terminology that is
used by this document.

Tenant System (TS): A physical or virtual system that can play the
role of a host, or a forwarding element such as a router, switch,
firewall, etc. It belongs to a single tenant and connects to one or
more VNs of that tenant.

End System (ES): An end system of a tenant, which can be, e.g., a
virtual machine(VM), a non-virtualized server, or a physical
appliance. A TS is attached to a Network Virtualization Edge(NVE)
node.

Network Virtualization Edge (NVE): An NVE implements network
virtualization functions that allow for L2/L3 tenant separation and
tenant-related control plane activity. An NVE contains one or more
tenant service instances whereby a TS interfaces with its associated
instance. The NVE also provides tunneling overlay functions.

Virtual Network (VN): This is a virtual L2 or L3 domain that belongs
to a tenant.

Information Mapping Authority (IMA). A back-end system that is
responsible for distributing and maintaining the mapping information
for the entire overlay system. Note that the WG never reached
consensus on what to call this architectural entity within the
overlay system, so this term is subject to change. In [I-D.ietf-nvo3
-overlay-problem-statement], such a back-end system is referred to as
a "oracle".

Hartman, et al.   Expires April 02, 2014   [Page 3]

3. NVO3 Overlay Architecture

        Please view in a fixed-width font such as Courier.

        Please view in a fixed-width font such as Courier.


```
                    ...................................
                    .                                 .
                    .                                 .
                    .                                 .
              +-+--+                        +--+-++--------+
 +--------+   | NV |                        | NV || Tenant |
 | Tenant +------+Edge|      L3 Overlay      |Edge|| System |
 | System |   +-+--+        Network         +--+-++--------+
 +--------+       .                             .
                  .                             .
                  .                             .
                  ...................................
```
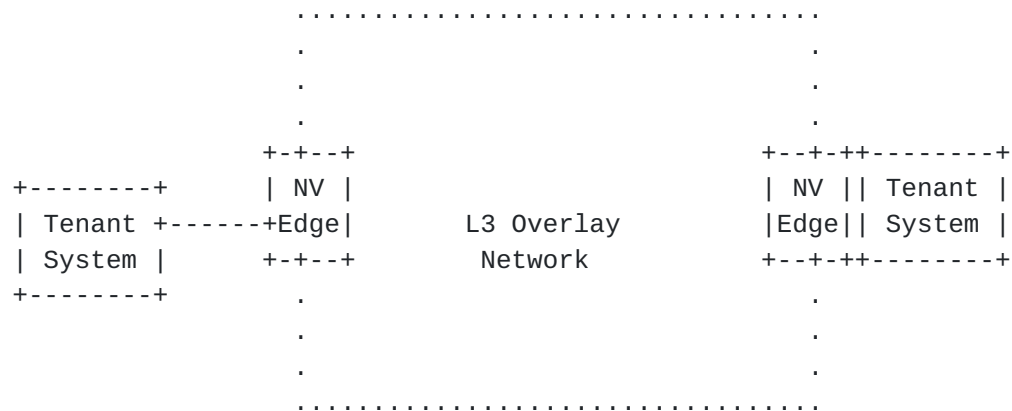

   This figure illustrates a simple nov3 overlay example where NVEs
   provide a logical L2/L3 interconnect for the TSes that belong to a
   specific tenant network over L3 networks. A packet from a tenant
   system is encapsulated when they reach the egress NVE. Then
   encapsulated packet is then sent to the remote NVE through a proper
   tunnel. When reaching the ingress NVE, the packet is decapsulated
   and forwarded to the target tenant system. The address
   advertisements and tunnel mappings are distributed amonge the NVEs
   through either distributed control protocols or by certain
   centralized servers (called Information Mapping Authorities).

4. Threat Model

   To benefit the discussion, in this analysis work, attacks are
   classified into two categories: inside attacks and outside attacks.
   An attack is considered as an inside attack if the adversary
   performing the attack (inside attacker or insider) has got certain
   privileges in changing the configuration or software of a NVO3 device
   (or a network devices of the underlying network where the overlay is
   located upon) and initiates the attack within the overlay security
   perimeter. In contrast, an attack is referred to as an outside
   attack if the adversary performing the attack (outside attacker or
   outsider) has no such privilege and can only initiate the attacks
   from compromised TSes. Note that in a complex attack inside and
   outside attacking operations may be performed in a well organized way
   to expand the damages caused by the attack.

This analysis assumes that security protocols, algorithms, and
implementations provide the security properties for which they are
designed; attacks depending on a failure of this assumption are out
of scope. As an example, an attack caused by a weakness in a
cryptographic algorithm is out of scope, while an attack caused by
failure to use confidentiality when confidentiality is a security
requirement is in scope.

4.1. Outsider Capabilities

The following capabilities of outside attackers MUST be considered in
the design of a NOV3 security mechanism:

1. Eavesdropping on the packets,

2. Replaying the intercepted packets, and

3. Generating illegal packets and injecting them into the network.

With a successful outside attack, an attacker may be able to:

1. Analyze the traffic pattern of a tenant or an end device,

2. Disrupt the network connectivity or degrade the network service
    quality, or

3. Access the contents of the data/control packets if they are not
    encrypted.

4.2. Insider Capabilities

It is assumed that an inside attacker can perform any types of
outside attacks from the inside or outside of the overlay perimeter.
In addition, in an inside attack, an attacker may use already
obtained privilege to, for instance,

1. Interfere with the normal operations of the overlay as a legal
    entity, by sending packets containing invalid information or with
    improper frequencies,

2. Perform spoofing attacks and impersonate another legal device to
    communicate with victims using the cryptographic information it
    obtained, and

3. Access the contents of the data/control packets if they are
    encrypted with the keys held by the attacker.

Note that in practice an insider controlling an underlying network
device may break the communication of the overlays by discarding or
delaying the delivery of the packets passing through it. However,
this type of attack is out of scope.

4.3. Security Properties

When encountering an attack, a virtual data center network MUST
guarantee the following security properties:

1. Isolation of the VNs: In
    [I-D.ietf-nvo3-overlay-problem-statement], the data plane
    isolation requirement amongst different VNs has been discussed.
    The traffic within a virtual network can only be transited into
    another one in a controlled fashion (e.g., via a configured
    router and/or a security gateway). In addition, it MUST be
    ensured that an entity cannot make use of its privilege obtained
    within a VN to manipulate the overlay control plane to affect on
    the operations of other VNs.

2. Spoofing detection: Under the attacks performed by a privileged
    inside attacker, the attacker cannot use the obtained
    cryptographic materials to impersonate another one.

3. Integrity protection and message origin authentication for the
    control packets: The implementation of an overlay control plane
    MUST support the integrity protection on the signaling packets.
    No entity can modify a overlay signaling packet during its
    transportation without being detected. Also, an attacker cannot
    impersonate a legal victim (e.g., a NVE or another servers within
    the overlay) to send signaling packets without detection.

4. Availability of the control plane: The design of the control plan
    must consider the DoS/DDoS attacks. Especially when there are
    centralized servers in the control plan of the overlay, the
    servers need to be well protected and make sure that they will
    not become the bottle neck of the control plane especially under
    DDOS attacks.

The following properties SHOULD be optionally provided:

1. Confidentiality and integrity of the data traffic of TSes. In
    some conditions, the cryptographic protection on the TS traffic
    is not necessary. For example, if most of the ES data is headed
    towards the Internet and nothing is confidential, encryption or
    integrity protection on such data may be unnecessary. In
    addition, in the cases where the underlay network is secure
    enough, no additional cryptographic protection needs to be
    provided.

2. Confidentiality of the control plane. On many occasions, the
    signaling messages can be transported in plaintext. However,
    when the information contained within the signaling messages are
    sensitive or valuable to attackers (e.g., the location of a ES,
    when a VM migration happens), the signaling messages related with
    that tenant SHOULD be encrypted.

[5](#). Basic Security Approaches

This section introduces the security mechanisms which could be used
to provided in order to guarantee the security properties mentioned
in [section 4](#) when encountering attacks.

[5.1](#). Securing the Communications between NVEs and TSes

Assume there is a VNE providing a logical L2/L3 interconnect for a
set of TSes. Apart from data traffics, the NVE and the TSes also
need to exchange signaling messages in order to facilitate, e.g., VM
online detection, VM migration detection, or auto-provisioning/
service discovery [[I-D.ietf-nvo3-framework](#)].

The NVE and its associated TSes can be deployed in a distributed way
(e.g., a NVE is implemented in an individual device, and VMs are
located on servers) or in a co-located way (e.g., a NVE and the TSes
it serves are located on the same server).

In the former case, the data and control traffic between the NVE and
the TSes are exchanged over network. If the NVE supports multiple
VNs concurrently, the data/control traffics in different VNs MUST be
isolated physically or by using VPN technologies. If the network
connecting the NVE and the TSes is potentially accessible to
attackers, the security properties of data traffic (e.g., integrity,
confidentiality, and message origin authenticity) SHOULD be provided.
The security mechanisms such as IPsec, SSL, and TCP-AO, can be used
according to different security requirements.

In order to guarantee the integrity and the origin authentication of
signaling messages, integrated security mechanisms or additional
security protocols need to be provided. In order to secure the data/

control traffic, cryptographic keys need to be distributed to
generate digests or signatures for the control packets. Such
cryptographic keys can be manually deployed in advance or dynamically
generated with certain automatic key management protocols (e.g., TLS
[RFC5246]). The TSes belonging to different VNs MUST use different
keys to secure the control packets exchanges with their NVE.
Therefore, an attacker cannot use the keys it obtained from a
compromised TS to generate bogus signaling messages and inject them
into other VNs without being detected. For a better damage
confinement capability, different TSes SHOULD use different keys to
secure their control packet exchanges with NVEs, even if they belong
to the same VN.

In the co-located case, all the information exchanges between the NVE
and the TSes are within the same device, and no standardized protocol
need to be provided for transporting control/data packets. It is
also important to keep the isolation of the TS traffic in different
VNs. In addition, in the co-location fashion, because the NVE, the
hypervisor, and the VMs are deployed on the same device, the
computing and memory resources used by the NVE , the hypervisor, and
the TSes need to be isolated to prevents a malicious or compromised
TS from, e.g., accessing the memory of the NVE or affecting the
performance of the NVE by occupying large amounts of computing
resources.

5.2. Securing the Communications within Overlays

This section analyzes the security issues in the control and data
plans of a NVO3 overlay.

5.2.1. Control Plane Security

It is the responsibility of the NVO3 network to protect the control
plane packets transported over the underlay network against the
attacks from the underlying network. The integrity and origin
authentication of the messages MUST be guaranteed. The signaling
packets SHOULD be encrypted when the signaling messages are
confidential. To achieve such objectives, when the network devices
exchange control plane packets, integrated security mechanisms or
security protocols need to provided. In addition, cryptographic keys
need to be deployed manually in advance or dynamically generated by
using certain automatic key management protocols (e.g., TLS
[RFC5246]).

In order to enforce the security boundary of different VNs in the
existence of inside adversaries, the signaling messages belonging to
different VNs need to be secured by different keys. Otherwise, an
inside attacker may try to use the keys obtained within a VN to

impersonate the NVEs in other VNs and generate illegal signaling
messages without being detected. If we expect to provide a better
attack confinement capability and prevent a compromised NVE to
impersonate other NVEs in the same VN, different NVEs working inside
a VN need to secure their signaling messages with different keys.
When there are centralized servers providing mapping information
(IMAs) within the overlay, it will be important to prevent a
compromised NVE from impersonating the centralized servers to
communicate with other NVEs. A straightforward solution is to
associate different NVEs with different keys when they exchange
information with the centralized servers.

In the cases where there are a large amount of NVEs working within a
NVO3 overlay, manual key management may become infeasible. First, it
could be burdensome to deploy pre-shared keys for thousands of NVEs,
not to mention that multiple keys may need to be deployed on a single
device for different purposes. Key derivation can be used to
mitigate this problem. Using key derivation functions, multiple keys
for different usages can be derived from a pre-shared master key.
However, key derivation cannot protect against the situation where a
system was incorrectly trusted to have the key used to perform the
derivation. If the master key were somehow compromised, all the
resulting keys would need to be changed. In addition, VM migration
will introduce challenges to manual key management. The migration of
a VM in a VN may cause the change of the NVEs which are involved
within the NV. When a NVE is newly involved within a VN, it needs to
get the key to join the operations within the VN. If a NVE stops
supporting a VN, it should not keep the keys associated with that VN.
All those key updates need to be performed at run time, and difficult
to be handled by human beings. As a result, it is reasonable to
introduce automated key management solutions such as EAP [RFC4137]
for NVO3 overlays.

When an automated key management solution for NVO3 overlays is
deployed, as a part of the key management protocol, mutual
authentication needs to be performed before two network devices in
the overlay (NVEs or IMAs) start exchanging the control packets.
After an authentication, an device can find out whether its peer
holds valid security credentials is is the one who it has claimed.
The authentication results is also necessary for authorization; it is
important for a device to clarify the roles (e.g., a NVE or a IMA)
that its authentication peer acts as in the overlay. Therefore, a
compromised NVE cannot use it credential to impersonate an IMA to
communicate with other NVEs. Only the control messages from the
authenticated entity will be adopted. In addition, authorization MAY
need to be performed. For instance, before accepting a control
message, the receiver NVE needs to verify whether the message comes
from one which is authorized to send that message. If the

authorization fail, the control message will be discarded. For
instance, if a control packet about a VN is sent from a NVE which is
not authorized to support the VN, the packet will be discarded.

The issues of DDOS attacks also need to be considered in designing
the overlay control plane. For instance, in the VXLAN
solution[I-D.mahalingam-dutt-dcops-vxlan], an attacker attached to a
NVE can try to manipulate the NVE to keep multicasting control
messages by sending a large amount of ARP packets to query the
inexistent VMs. In order to mitigate this type of attack, the NVEs
SHOULD be only allowed to send signaling message in the overlay with
a limited frequency. When there are centralized servers (e.g., the
backend oracles providing mapping information for
NVEs[I-D.ietf-nvo3-overlay-problem-statement], or the SDN
controllers) are located within the overlay, the potential security
risks caused by DDOS attack on such servers can be more serious.

In addition, during the design of the control plane, it is important
to consider the amplification effects which may potential be used by
attackers to carry out reflection attacks.

5.2.2. Data Plan Security

[I-D.ietf-nvo3-framework] specifies a NVO3 overlay needs to generate
tunnels between NVEs for data transportation. When a data packet
reaches the boundary of a overlay, it will be encapsulated and
forwarded to the destination NVE through a proper tunnel. It is
normally assume that the underlying network connecting NVEs are
secure to outside attacks since it is under the management of DC
vendor and cannot be directly accessed by tenants. However, when
facing inside attacks, conditions could be complex. For instance, an
inside attacker compromising a underlying network device may
intercept an encapsulated data packet transported a tunnel, modify
the contents in the encapsulating tunnel packet and, transfer it into
another tunnel without being detected. When the modified packet
reaches a NVE, the NVE may decapsulated the data packet and forward
it into a VN according to the information within the encapsulating
header generated by the attacker. Similarly, a compromised NVE may
try to redirect the data packets within a VN into another VN by
adding improper encapsulating tunnel headers to the data packets.
Under such circumstances, in order to enforce the VN isolation
property, signatures or digests need to be generated for both data
packets and the encapsulating tunnel headers in order to provide data
origin authentication and integrity protection. In addition, NVEs
SHOULD use different keys to secure the packets transported in
different tunnels.

[6](#). **Security Issues Imposed by the New Overlay Design Characteristics**

[6.1](#). **Scalability Issues**

   NOV3 WG requires an overlay be able to work in an environment where
   there are many thousands of NVEs (e.g. residing within the
   hypervisors) and large amounts of trust domains (VNs). Therefore,
   the scalability issues should be considered. In the cases where a
   NVE only has a limited number of NVEs to communicate with, the
   scalability problem brought by the overhead of generating and
   maintaining the security channels with the remote NVEs is not
   serious. However, if a NVE needs to communicate with a large number
   of peers, the scalability issue could be serious. For instance,
   in[I-D.ietf-ipsecme-ad-vpn-problem], it has been demonstrated it is
   not trivial to enabling a large number of systems to communicate
   directly using IPsec to protect the traffic between them.

[6.2](#). **Influence on Security Devices**

   If the data packets transported through out an overlay are encrypted
   (e.g., by NVEs), it is difficult for a security device, e,g., a
   firewall deployed on the path connecting two NVEs to inspect the
   contents of the packets. The firewall can only know which VN the
   packets belong to through the VN ID transported in the outer header.
   If a firewall would like to identify which end device sends a packets
   or which end device a packet is sent to, the firewall can be deployed
   in some place where it can access the packet before it is
   encapsulated or un-encapsulated by NVEs. However, in this case, the
   firewall cannot get VN ID from the packet. If the firewall is used
   to process two VNs concurrently and there are IP or MAC addresses of
   the end devices in the two VNs overlapped, confusion will be caused.
   If a firewall can generate multiple firewalls instances for different
   tenants respectively, this issue can be largely addressed.

[6.3](#). **Security Issues with VM Migration**

   The support of VM migration is an important issue considered in NVO3
   WG. The migration may also cause security risks. Because the VMs
   within a VN may move from one server to another which connects to a
   different NVE, the packets exchanging between two VMs may be
   transferred in a new path. If the security policies deployed on the
   firewalls of the two paths are conflict or the firewalls on the new
   path lack essential state to process the packets. The communication
   between the VMs may be broken. To address this problem, one option
   is to enable the state migration and policy confliction detection
   between firewalls. The other one is to force all the traffic within
   a VN be processed by a single firewall. However this solution may
   cause traffic optimization issues.

7. IANA Considerations

   This document makes no request of IANA.

   Note to RFC Editor: this section may be removed on publication as an
   RFC.

8. Security Considerations

   TBD

9. Acknowledgements

10. References

10.1. Normative References

   [RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119, March 1997.

10.2. Informative References

   [I-D.ietf-ipsecme-ad-vpn-problem]
               Manral, V. and S. Hanna, "Auto Discovery VPN Problem
               Statement and Requirements", draft-ietf-ipsecme-ad-vpn-
               problem-09 (work in progress), July 2013.

   [I-D.ietf-nvo3-framework]
               Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y.
               Rekhter, "Framework for DC Network Virtualization", draft-
               ietf-nvo3-framework-03 (work in progress), July 2013.

   [I-D.ietf-nvo3-overlay-problem-statement]
               Narten, T., Gray, E., Black, D., Fang, L., Kreeger, L.,
               and M. Napierala, "Problem Statement: Overlays for Network
               Virtualization", draft-ietf-nvo3-overlay-problem-
               statement-04 (work in progress), July 2013.

   [I-D.kreeger-nvo3-hypervisor-nve-cp]
               Kreeger, L., Narten, T., and D. Black, "Network
               Virtualization Hypervisor-to-NVE Overlay Control Protocol
               Requirements", draft-kreeger-nvo3-hypervisor-nve-cp-01
               (work in progress), February 2013.

   [I-D.mahalingam-dutt-dcops-vxlan]
               Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger,
               L., Sridhar, T., Bursell, M., and C. Wright, "VXLAN: A
               Framework for Overlaying Virtualized Layer 2 Networks over

                  Layer 3 Networks", draft-mahalingam-dutt-dcops-vxlan-04
                  (work in progress), May 2013.

     [RFC4137] Vollbrecht, J., Eronen, P., Petroni, N., and Y. Ohba,
                  "State Machines for Extensible Authentication Protocol
                  (EAP) Peer and Authenticator", RFC 4137, August 2005.

     [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security
                  (TLS) Protocol Version 1.2", RFC 5246, August 2008.

Authors' Addresses

     Sam Hartman
     Painless Security
     356 Abbott Street
     North Andover, MA 01845
     USA

     Email: hartmans@painless-security.com
     URI: http://www.painless-security.com

     Dacheng Zhang
     Huawei
     Beijing
     China

     Email: zhangdacheng@huawei.com

     Margaret Wasserman
     Painless Security
     356 Abbott Street
     North Andover, MA 01845
     USA

     Phone: +1 781 405 7464
     Email: mrw@painless-security.com
     URI: http://www.painless-security.com