Network Working Group   Z. Qiang
Internet Draft   A. Kavanagh
Intended status: Informational   Ericsson
Expires: April 16, 2014   October 16, 2013

Security Requirements of NVO3
draft-zu-nov3-security-requirements-01.txt

Status of this Memo

Copyright Notice

Abstract

This draft discusses the security requirements and several issues which need to be considered in securing a NVO3 network architecture based virtualized data center network for multiple tenants. In addition, the draft also discusses issues that could be addressed or mitigated.

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 highlights the security risks that a NVO3 network may encounter, and documents the lists the security requirements that a NVO3 network should fulfill. The purpose of the draft is to propose additional NVO3 network security requirement considerations which can be incorporated into the WG security requirement draft.

Note, it is not the intention to replace the Security Considerations section in each NVO3 draft by this document. This document provides the high level views of the security requirements when NVO3 network is developed. It only lists the architecture level security requirements which can be used as inputs at the design phase of the NVO3 network architecture, control plane and data plane. Each NVO3 drafts must have its security considerations which shall define the detail security solutions of a specific architecture and / or protocol. This document is only the input document when the Security Considerations section in each NVO3 draft is discussed.

2. Conventions used in this document

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. 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].

4. Security Risk

Overlay infrastructure increases security risks and introduces new threats. In a NVO3 network, there are security risks that the attack made on the underlying network, including the NVO3 control protocols, may be initiated from an exposed overlay virtual network; or the attack made on the encapsulated virtual networks may be initiated from the underlying network or a compromised overlay virtual network.

In a perfect world, virtualization is considered secure with no level of privilege within the virtualized guest environment that permits interference with the host system. There are really not much security issues if a tenant network is isolated as it is designed.

In practice, there are occasional misconfigurations and/or security vulnerabilities that allow an attacker to circumvent these protections and gain access to other virtual machines, or even worse the underlay network. While the misconfigurations or vulnerabilities are pretty rare, they do exist.

In NVO3 network, both the hypervisor and the NVE module is just a piece of software. Any software is vulnerable to a local privilege escalation attack. The vulnerability may be exploited for local privilege escalation or a guest-to-host virtual machine escape. So both hypervisor and NVE may be compromised due to any misconfigurations or software vulnerabilities. When the hypervisor and the NVE are compromised by the attacker, the NVO3 network and the underlay network architecture may be exposed to the attacker.

5. Security Control

5.1. Control Plane Protection

The DC service provider has the responsibility to protect the NVO3 control plane signaling against any attack.

5.1.1. Control Plane Availability

In a NVO3 network, the control plane is used to control the overlay data plane tunnels for the VNs. It must be available when it is needed. This means that the NVO3 control plane network components and the control plane interfaces must be functioning correctly and preventing any denial-of-service attacks.

R1. At NVO3 network design, any NVO3 network components, including NVA and NVE, SHALL not become the bottleneck of the control plane traffic. This is to avoid any DoS/DDoS attacks and to provide control plane availability.

R2. The control plane design SHALL minimize the amplification effects which have the potential to be used by attackers to carry out reflection attacks. For instance, the usage of the NVE broadcast address MUST be avoided or restricted in the control plane protocol. And the NVE MUST discard any control plane traffic received from any non-participating NVEs or unknown network addresses. This can minimize the amplification effects which a compromised NVE or a compromised network component to initiate a distributed reflection DoS attack by sending request message to the broadcast address of the NVEs, where the NVEs are exploited to act as reflectors of the amplification attacks.

R3. Some overlay data plane tunnel protocols may use endpoint addresses which is algorithmically derived from some known values. These addresses are structured, and the fields contained in them can be fairly predictable. If the control plane and data plane are sharing the same address space, it reduces the search space for an attacker and reduces the resistance of the address in scanning attacks. Therefore the NVE SHOULD have separated address space for data plane tunnel end point and control plane traffic in order to minimize security exposure of the control plane addresses, as recommended in [RFC6169].

5.1.2. NVA-NVE Control Plane

In [I-D.ietf-nvo3-arch], two different possibilities are allowed for VN context configuration and inner-outer address mapping table updating at VN connection / disconnection or vNIC association / disassociation:

- NVA Configuration only; or
- With Hypervisor/NVE Notifications

With the "NVA Configuration only" approach, the hypervisor is always configured by the VM Orchestration Systems. It is assumed that either the NVA is collocated with the VM Orchestration Systems, or there is an interface between the VM Orchestration Systems and NVA, which the NVA learns the VN connection / disconnection and vNIC association / disassociation from the VM Orchestration Systems. Then the NVA configures the attached NVE with the VN context of the connected / disconnected VN and the associated / disassociated vNIC addresses. The next step is that the NVA updates the inner-outer address mapping table of the VN at both the attached NVE and all the participating remote NVEs.

With the "With Hypervisor/NVE Notifications" approach, the hypervisor is always configured by the VM Orchestration Systems. The attached NVE is informed by the hypervisor using the Hypervisor-NVE protocol at VN connection / disconnection and vNIC association / disassociation. Then the attached NVE notifies the NVA with the received VN updating information. The next step is that the NVA updates the inner-outer address mapping table of the VN at both the attached NVE and all the participating remote NVEs.

In both above approaches, the NVA is the network entity that provides reachability and forwarding information to all participating NVEs. The NVA is the center control point of the NVO3 control plane network. If the NVA is compromised, the entire NVO3 control plane can be damaged. Therefore it is very important to protect the NVA from any possible attacks.

Comparing the above two approaches, the "With Hypervisor/NVE Notifications" approach may have additional security risks. The updating of the inner-outer address mapping table of the VN at the attached NVE and all the participating remote NVEs is triggered by the VN updating notifications received from the hypervisor, at VN connection / disconnection and vNIC association / disassociation. In both the split-NVE case and the NVE collocated with the hypervisor case, the updating of the inner-outer address mapping table may be triggered by incorrect VN updating information received from a compromised hypervisor or a compromised NVE. And it is difficult to detect it and prevent it, unless an additional validation procedure is supported in the NVA.

With the "NVA Configuration only" approach, the updating of the inner-outer address mapping table at the attached NVE and all the participating remote NVEs is triggered by the VN updating information learned from the VM Orchestration Systems. In such circumstance, a compromised hypervisor or a compromised NVE has limited security risks on the NVO3 control plane. For instance, the compromised NVE may send error notifications with incorrect error information to the NVA, which may trigger the error-handling procedure. But it should not trigger the inner-outer address mapping table updating procedure. Once the compromise is detected, the NVA may have the possibilities to mitigate security damages by informing the VM Orchestration Systems to relocate the attached VNs from the compromised NVE to other NVEs.

In both above approaches, there are other security risks in the NVA-NVE control plane which need to be avoided.

For instance, if the control plane traffic between the NVA and the NVE has been intercepted or modified, a compromised network component may attempt to learn the NVO3 network topologic in order to initiate an attack. Or a compromised network component may try to redirect the NVA-NVE traffic as a man in middle. If the control plane traffic between the NVA and the NVE has been redirected, the NVEs may not be updated correctly and timely at VN connection / disconnection or vNIC association / disassociation. And the NVA may not be able to receive any VN updating or error notifications from the NVEs.

Another security risk is that a compromised network component may try to impersonate as a NVA to update the NVEs with incorrect VN configuration. Or a compromised network component may try to impersonate as a NVE to notify the NVA with incorrect VN updating or error information. In those circumstances, the VN traffic may be redirected to a desired network point, or the data plane connectivity of a VN may be disabled by removing / redefining the overlay tunnel end point.

R4. At the NVA-NVE control plane, authentication and authorization of the NVA MUST be supported to prevent a compromised network component for impersonating as a NVA when communicate with NVEs, using incorrect VN updating information, e.g. an untrue inner outer address mapping table updating.

R5. At the NVA-NVE control plane, ingress filtering SHOULD be supported at the NVA. Any control plane traffic received from any unknown addresses MUST be discarded without processing. This is to prevent a compromised network component for impersonating as a NVE when communicate with the NVA using untrue network updating information or error notifications.

R6. At the NVA-NVE control plane, with the "With Hypervisor/NVE Notifications" approach, authentication of the NVE MUST be supported, otherwise it MAY be supported to prevent a compromised network component for impersonating as a NVE using a snooped NVE address as source address when communicate with the NVA with untrue network updating information or error notifications.

R7. The NVA-NVE control plane protocol MUST be protected with integrity and confidentiality against any off-path or on-path attacks. This is to avoid the NVA-NVE control plane messages to be intercepted or modified by a compromised network component, which may attempt to learn the NVO3 network topologic in order to initiate an attack.

5.1.3. NVE-NVE Control Plane

Besides the approaches described in the previous section, it is also possible to use a NVE-NVE control plane protocol to update the peer NVEs' inner-outer address mapping table timely at VN connection / disconnection or vNIC association / disassociation. However, this approach may require the NVE-NVE control plane packets to be flooded to all NVEs when no mapping exists, which may have additional security risks compare to other approaches described in the previous section.

For instance, a compromised network component may attempt to learn the NVO3 network topologic by intercepting any NVE-NVE control plane messages. It may also try to modify the NVE-NVE control plane messages in order to redirect the control plane traffic to a desired network point. Or it may impersonate as a NVE using a snooped participating NVE address to update the peer NVEs’ inner outer address mapping table of the VN in order to redirect the VN traffic.

Moreover, if a NVE is compromised, it may attempt to send control plane messages to update the peer NVEs with untrue network updating information. In this circumstance, the VN traffic may be redirected to a desired network point.

Furthermore, a compromised network component or a compromised NVE may try to initiate DOS attack by flooding all NVEs with untrue network updating information. If the NVE-NVE control plane protocol requires a respond, all the NVEs are exploited to act as reflectors of the amplification attacks. When the number of involved NVEs is large enough, it can slow down the NVE-NVE control plane to the point of impossible to work on.

Especially when a NVE is compromised, it is very difficult to detect it and mitigate the damage without additional security mechanism.

Therefore it is very important to protect the NVE-NVE control plane from any possible attacks initiated from a compromised network component or a compromised NVE.

R8. At the NVE-NVE control plane, authentication of the NVE MUST be supported to prevent a compromised network component for impersonating as a NVE when communicate with other NVEs.

R9. The NVE SHOULD apply ingress controls at the NVE-NVE interface to filter the incoming control plane traffic and discard any control plane traffic received from non-participating NVEs without processing. This can prevent a compromised NVE sending any control plane messages which it is not supposed to send.

R10. The NVE-NVE control plane protocol MUST be protected with integrity and confidentiality against any off-path or on-path attacks.

R11. If the Inter-DC control plane traffic is crossing Public Internet, it MUST be protected by one or more security solutions to provide confidentiality, integrity and availability. This is to avoid the crossing DC NVE-NVE control plane messages to be intercepted or modified by an attacker from the public internet.

5.1.4. Hypervisor-to-NVE Control Plane

In [I-D.ietf-nvo3-arch], two different possibilities are allowed for NVE implementations: "Collocated with Hypervisor" or "Split-NVE". The Hypervisor-to-NVE control plane protocol is only needed at the "With Hypervisor/NVE Notifications" approach and the "Split-NVE" use case. It is used by the hypervisor to update the attached NVE at VN connection / disconnection and vNIC association / disassociation.

When the NVE is collocated with the hypervisor, there are additional security risks if the hypervisor may be compromised. As the NVE's configuration including the security keys may be exposed to the attacker, the security damages could be multiplied to other NVO3 control plane in some control plane approaches, e.g. using NVE-NVE for inner-outer address mapping table updates. And it is difficult to detect and prevent it.

In the Split-NVE case, there are security risks that the NVE may be polluted by a compromised hypervisor with incorrect network updating information. However in this circumstance, the security damages can be limited to the hypervisor and the VNs attached to the compromised hypervisor. There are still ways to protect the attached NVE itself and mitigate the damages.

Therefore, when Hypervisor-to-NVE control plane protocol is used, it is very important to protect the Hypervisor-NVE control plane from any possible attacks initiated from a compromised hypervisor.

R12. At the Hypervisor-to-NVE control plane protocol, authentication of the hypervisor SHOULD be provided to prevent a compromised hypervisor for impersonating as another hypervisor when communicate with the attached NVE.

R13. The Hypervisor-to-NVE control plane protocol MAY be protected with integrity to avoid the Hypervisor-to-NVE control plane messages to be intercepted and modified by a compromised hypervisor attached to the same NVE.

5.2. Data Plane Protection

Data plane protection is the primary concern for a NVO3 network. And it depends on the control plane security described at previous section.

5.2.1. Security Policies on Tenant Traffic

In a NVO3 network data plane, the overlay network could be exploited to act as reflectors of the amplification attacks, which can be used to initiate DDOS / DOS attack on some network services provided by the NVO3 architecture.

For instance, a compromised tenant system may try to send a broadcast message to all the VMs in a VN with an intended victim's spoofed source IP address. The victim could be one of the NVO3 network services, e.g. a L3NVE where the layer 3 routing or forwarding function of the VN is provided. Most VMs on the VN, by the default, may respond by sending a reply to the source IP address. If the number of VMs on the VN to be involved is large enough, the victim will be flooded. This can slow down the NVO3 network service to the point where it becomes impossible to work on.

Therefore it is important to apply proper security policies on the received VN data traffic before forwarding it to the next hop, e.g. an embedded firewall function in the NVE. Any disallowed data traffic shall be filtered and discarded at an early point.

R14. The NVO3 infrastructure SHOULD support VN based security policy management, i.e. security policy defined with a granularity down to VN ID. Additional granularity MAY be supported.

R15. When the security policy management is enabled for the data packets of a VN, the security policies MUST be applied on the un-tunneled data packets.

R16. When the security policy management is enabled for the data packets of a VN, the same security policies MUST be applied on the VN data traffic during and after VM mobility. The VM shall have the same security policies wherever it has been migrated.

R17. When the security policy management is enabled for the data packets of a VN, the security policies MUST be applied on the inter-VN traffic. This is to avoid a compromised VM trying to involve more VMs which belong to other VNs in amplification attacks.

R18. When Public Internet connectivity is allowed for a VN, it is often that some layer 3 network services may be provided by the NVO3 network, such as NAT. This opening created in the layer 3 network services increases its Internet attack surface area. If vulnerabilities are present, this increased exposure can be used by attackers and their programs. Therefore the security policies MUST be applied on the VN Public Internet traffic before forwarding between the VN and Internet. This is to ensure that IP traffic from the public Internet cannot be used to modify the configuration of the VMs, or to mount DoS attacks on them.

R19. The NVE SHOULD apply security policies on the data packets received from the End Devices before encapsulation. Any disallowed traffic MUST be discarded.

R20. The NVE SHOULD apply security policies on the data packets received from the remote NVEs after de-capsulation, and discard any disallowed data packets before forwarding to the End Devices.

5.2.2. Protect the Overlay Tunnel

In a NVO3 network, a compromised network component may impersonate as a NVE to send data traffic of a VN which it is not supposed to send. When impersonating as a NVE, the compromised network component may use a snooped NVE address as the overlay tunnel source point to skip the ingress filter control at the peer NVE. In this case, per-tunnel based signatures or digests may provide data origin authentication, non-repudiation, and integrity protection. However, in a larger DC, which may have millions of VNs and thousands of NVEs, the key management scalability can be a concern. Besides, if a NVE has been compromised, it is difficult to prevent the compromised NVE from sending data traffic which it is not supposed to send. Even with a per-VN based key, it is not guaranteed that a NVE will have the key deleted after a VN is migrated into other NVEs. In that circumstance, a compromised NVE may use the un-deleted key for generating data traffic of a VN which is not attached to it any more.

To avoid that, NVE authentication and ingress control on both the inner address and outer address of an encapsulation tunnel is important. The ingress control can mitigate the security damage within a smaller amount of NVEs, i.e. the participating NVEs.

R21. The NVE SHOULD filter on the outer source address of the tunneled data packets received from the remote NVEs, and discard any data packets received from any non-participating NVEs or unknown address. This is to prevent a compromised NVE or a compromised network component from sending data traffic of a VN which it is not attached to it.

R22. The NVE SHOULD filter on the inner source address of the tunneled data packets received from a remote participating NVE, and discard any data packets which the participating NVE is not supposed to send. In the case that a participating NVE is compromised, this can prevent the compromised NVE from sending data traffic of a VN which it is not attached to it.

R23. The digital signature MAY be supported in the NVE to prevent a compromised network component for impersonating as a NVE when generating tunneled data traffic of a VN using a snooped NVE address as the overlay tunnel source point. It also can reduce the risks of a man in middle attack.

R24. When Layer 3 routing/forwarding service is supported for a VN, the NVE SHOULD discard any tunneled IP packets that specify additional routing, as recommended in [RFC6169], though it may be allowed for the End Device to configure what source-routing types are allowed.

R25. Additional security mechanisms MAY be supported on the interworking function when supporting multiple encapsulation formats in a NVO3 network.

5.2.3. Protect the Tenant Traffic

In a NVO3 network, if the tenant traffic privacy is the concern, cryptographic measures must be applied in addition. Confidentiality and integrity on the tenant data plane traffic could avoid the tenant traffic to be redirected, intercepted or modified by a compromised underlay network component.

R26. All data plane packets MAY be protected in transit with confidentiality and integrity, including the un-tunneled traffic between the End devices and the NVEs, and the tunneled traffic between the NVEs.

R27. If the inter-DC data plane traffic is crossing Public Internet, it SHOULD be protected by one or more security solutions to provide confidentiality, integrity and availability.

5.3. Operation and Management

The Operation and Management data protection is also the concern for a NVO3 network.

R28. The NVO3 Operation and Management traffic MUST be isolated from any other underlay traffic in order to minimize security exposure of the Operation and Management traffic, and mitigate any damage due to an attack, as recommended in [RFC6169].

R29. The NVO3 Operation and Management data MUST be protected with confidentiality, integrity and availability while in transit.

5.4. Logging

Logging function is very important at network security risks detection.

R30. All NVO3 network components, e.g. NVA and NVE, SHOULD support collection of security logs and sending them to a centralized logging service.

R31. A centralized security logging and audit handling mechanism SHOULD be supported. Any access to the NVO3 resources SHOULD be recorded and stored in the centralized logging and audit storage.

5.5. Scalability

Scalability is a big concern in NVO3 network especially where a DC may have large amounts of VNs.

One example is that some security solutions may require a per-VN based key management. In a large data center, where the number of VNs can be huge, even there is no technology issue when generating that amount of security keys, but it may be a scalability issue at security credential management. Therefore optimized security credential management solution shall be allowed.

R32. The NVO3 network security solutions SHOULD minimize the impact on scalability and allow for simple configuration, e.g. shared security credential management.

5.6. Extensibility

R33. The NVO3 network security solution SHOULD be extensible to allow new security functionality to be introduced in the future.

R34. The NVO3 network security solution SHOULD be defined such that End Devices existing security solution can be supported without implementation impacts.

6. Security Considerations

This is a requirement document for the NVO3 network security and in itself does not introduce any new security concerns.

7. IANA Considerations

No actions are required from IANA for this informational document.

8. References

8.1. Normative References

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

[RFC2234] Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, Internet Mail Consortium and Demon Internet Ltd., November 1997.

[RFC6169] S. Krishnan, D. Thaler, J. Hoagland, "Security Concerns with IP Tunneling", RFC 6169, April 2011.

8.2. Informative References

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[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-03 (work in progress), May 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.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-arch] D. Black, J. Hudson, L. Kreeger, M. Lasserre, T. Narten, "An Architecture for Overlay Networks (NVO3)", draft-narten-nvo3-arch-00 (work in progress), July 2013.

9. Acknowledgments

Many people have contributed to the development of this document and many more will probably do so before we are done with it. While we cannot thank all contributors, some have played an especially prominent role. The following have provided essential input: Suresh Krishnan, David Allan I, Makan Pourzandi, Melinda Shore.

Appendix A. Change Log

A.1. Changes From -00 to -01

1. Numerous editorial and clarity improvements.

2. Adding analysis text and wording improvements in R1 ~ R8, R23, R25, R27 ~ R29, R30.

3. Adding more subclauses with more analysis text in section 5.1 and 5.2

Authors' Addresses

Zu Qiang
Ericsson
8400 Decarie Blvd.
Town of Mount Royal, QC
Canada

Phone: +1 514 345 7900 x47370
Email: Zu.Qiang@ericsson.com

Alan Kavanagh
Ericsson
8400 Decarie Blvd.
Town of Mount Royal, QC
Canada

Phone: +1 514 345 7900 x42871
Email: alan.kavanagh@ericsson.com