Introducing the NSX-T Platform

An Architecture Overview

TECHNICAL WHITE PAPER

Table of Contents

Introduction

This technical white paper is focused on VMware NSX-T architecture, components, and related capabilities. We also explain the technical advantages and benefits of the NSX-T architecture. This white paper is primarily for data center and cloud architects and engineers.

Networking and security today

In the digital business era, nearly every organization is building custom applications to drive its core business and gain a competitive advantage. The speed with which development teams deliver new applications and capabilities directly impacts the success and bottom line of an organization. This has placed increasing pressure on organizations to innovate quickly, and has made developers central to this critical mission. As a result, the way developers create apps, and the way IT provides services for those apps and the wider business, has been evolving.

Application explosion – With applications quickly emerging as the new business currency, developers are under immense pressure to deliver apps in record time. This increasing need to deliver more apps—in less time—can drive developers to use public clouds or open source technologies, where they can write and provision apps in a fraction of the time.

Heterogeneity – The application explosion has given rise to heterogeneous environments, with application workloads being run inside VMs, containers, clouds, or bare metal. IT departments still need to have governance, security and visibility for application workloads in public clouds and clouds managed by third-parties, in addition to on-premises infrastructure.

Cloud-centric architectures – Cloud-centric architectures and approaches to building and managing applications are increasingly common because of their efficient development environments and fast delivery of applications. But cloud architectures can put pressure on networking and security, as they must integrate with private and public clouds. Logical networking and security must be highly extensible to adapt and keep pace with ongoing change.

Against this backdrop of increasing application needs, and greater heterogeneity and complexity of environments, IT must still protect applications and data, whose attack surface is continuously expanding. To quote VMware CTO, Bruce Davie, “The expectation is to meet the needs of two distinct communities: developers deploying applications on a heterogeneous collection of infrastructure, and IT managers trying to apply an appropriate level of control to ensure compliance and security without impeding developer productivity.”

How NSX-T addresses these trends and challenges

VMware NSX-T is designed to address emerging application frameworks and architectures that have heterogeneous endpoints and technology stacks. In addition to vSphere, these environments may also include other hypervisors, containers, bare metal, and public clouds. NSX-T allows IT and development teams to choose the technologies best suited for their particular applications. NSX-T is also designed for management, operations and consumption by development organizations – in addition for IT.

Architecture Overview

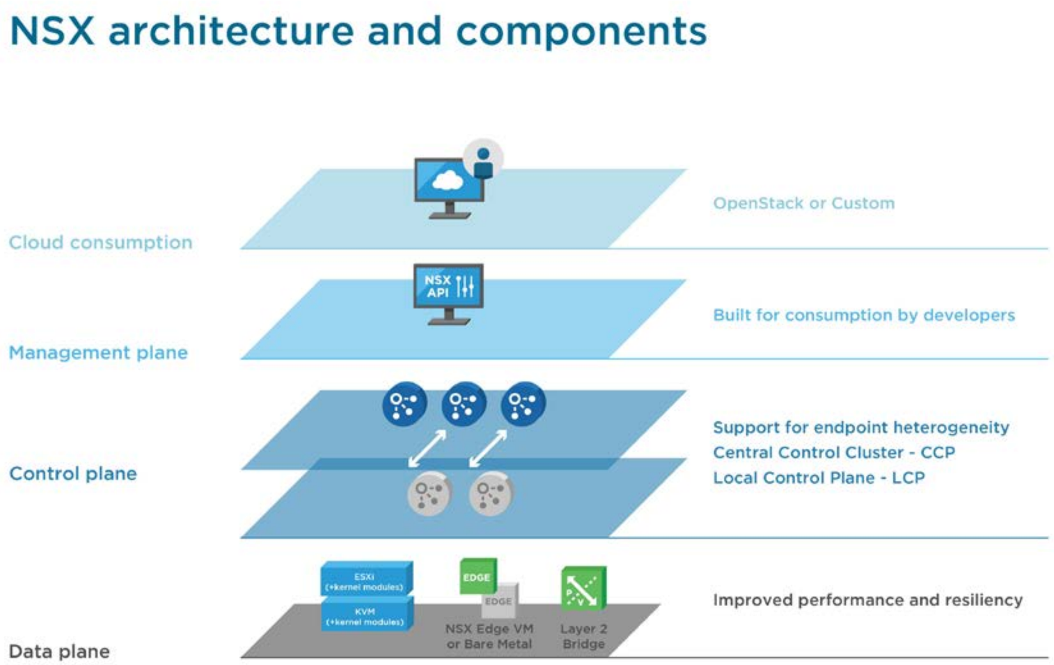
The NSX-T architecture has built-in separation of the data plane, control plane and management plane. This separation delivers multiple benefits, including scalability, performance, resiliency and heterogeneity. In this section, we’ll walk you through each layer and its specific role and characteristics.

Some of the key architecture highlights discussed include the following:

- **Management plane:** NSX-T management plane is designed from the ground up with advanced clustering technology, which allows the platform to process large-scale concurrent API requests.
- **Control plane:** NSX-T control plane keeps track of the real-time virtual networking and security state of the system. NSX-T control plane separates the control plane into a central clustered control plane (CCP) and a local control plane (LCP). This simplifies the job of the CCP significantly and enables the platform to extend and scale for heterogeneous endpoints.
- **Data plane:** The NSX-T data plane introduces a host switch (rather than relying on the vSwitch), which decouples it from the compute manager and normalizes networking connectivity. All create, read, update, and delete (CRUD) operations are performed via the NSX-T Manager.

Next, we'll discuss several other characteristics of the NSX-T architecture components and features. For example, multi-tiered routing, DPDK powered data-plane, distributed firewalling, network services and operational tools like Traceflow to name a few.

The following graphic illustrates the detailed NSX-T architecture and components.


NSX architecture and components
Cloud consumption
OpenStack or Custom
Management plane
NSX API
Built for consumption by developers
Control plane
Support for endpoint heterogeneity
Central Control Cluster - CCP
Local Control Plane - LCP
Data plane
ESXi (+kernel modules)
KVM (+kernel modules)
EDGE
EDGE
NSX Edge VM or Bare Metal
Layer 2 Bridge
Improved performance and resiliency

As seen in the graphic, the NSX-T architecture is made up of the management plane, control plane and data plane. This separation allows the architecture to grow and scale without impacting workloads. We'll describe each layer and its components below.

Management Plane

The NSX-T management plane provides secure concurrent entry points to the system via a full-featured RESTful API or a browser-based, high-performance and modern HTML5 Graphical User Interface. The management plane is designed from the ground up for performance to process large-scale concurrent API calls from cloud management platforms (CMPs). The system is able to be integrated into any CMP and will ship with a fully supported OpenStack Neutron plugin. As the system scales, the management plane has the ability to scale out using advanced clustering technology.

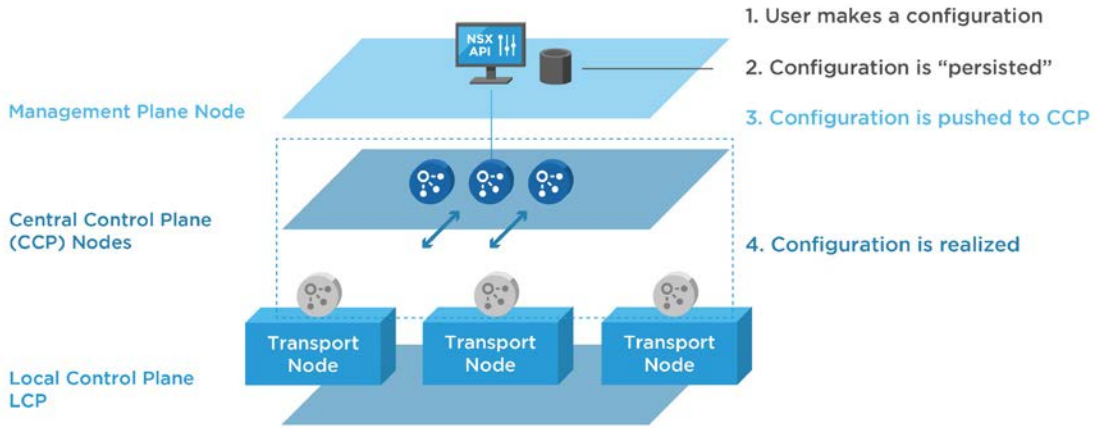
Another salient feature of the NSX-T management plane is that it is decoupled from vCenter and can be used to manage networking and security on other compute platforms, providing choice to the user. In spite of not being tightly coupled, vCenter will still provide greater value due to the ecosystem and functionality integration.

Management plane functions:

- Persist desired configuration from the cloud user for the entire networking and security topology via full-featured RESTful API
- Handle platform life-cycle management tasks for the management plane, control plane and data plane nodes in the system
- Provide operational tools and rich, REST-based APIs for handling user queries regarding system state
- Handle authentication, monitoring, grouping and inventory collection from the compute managers

The NSX Manager can be deployed in a VM form factor on either ESXi or KVM.

Control Plane


1. User makes a configuration
2. Configuration is "persisted"
3. Configuration is pushed to CCP
4. Configuration is realized
NSX API
Management Plane Node
Central Control Plane (CCP) Nodes
Local Control Plane LCP
Transport Node
Transport Node
Transport Node

The NSX-T control plane encompasses a clustered control-plane (CCP) running on controller nodes and a localized control plane (LCP) on compute endpoints.

The NSX-managed compute endpoints are known as transport nodes.

The CCP computes and disseminates the ephemeral runtime state based on configuration from the management plane and topology information reported by the data plane elements.

The LCP runs on the compute endpoints. It computes the local ephemeral runtime state for the endpoint based on updates from CCP and local data plane information. The LCP pushes stateless configuration to forwarding engines in the data-plane as well as reports the information back to the CCP. This simplifies the job of the CCP significantly and enables the platform to scale to thousands of heterogeneous endpoints (hypervisor, container host, bare metal or public cloud).

NSX controllers can be deployed in a VM form factor on either ESXi or KVM.

Data Plane

NSX-T includes a number of capabilities in the data plane that improve the performance and resiliency of the platform.

The NSX-T data plane can be enabled on ESXi, KVM hypervisors and appliances providing gateway functionality called edge nodes to provide rich networking and security services. Some of these services include logical switching, distributed logical routing, distributed firewalling and network services like NAT, DHCP Relay/Server and MetaData Proxy functionality. Data plane forwarding and transformation decisions are made based on the local tables populated by the control plane.

NSX-T Compute Independent Host Switch

The NSX-T data plane is powered by the vSwitch. In order to act as a common networking platform for a variety of compute and cloud-based connectivity, NSX-T normalizes the networking connectivity via the introduction of a host switch.

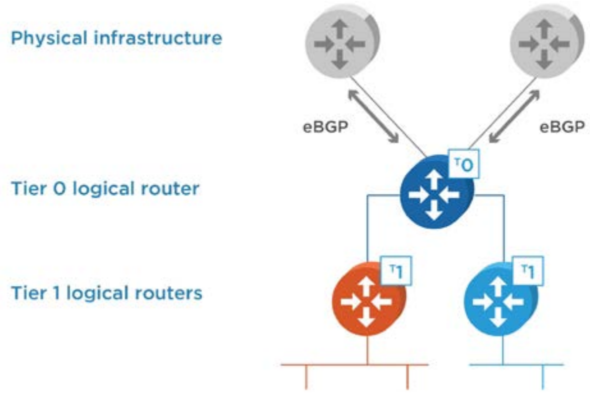
The provisioning of this high-performance host switch can be managed from the NSX management plane and is decoupled from the compute manager. All CRUD (create/remove/update/delete) operations on the host switch are performed from the NSX manager, including configuring uplink teaming profiles and defining switch security/QOS settings, etc.

The host switch enables the overlay network, as well as traditional VLAN-based topology. The NSX-T overlay topology control is enhanced by introduction of GENEVE-based encapsulation.
GENEVE is a tunneling mechanism which provides extensibility while still leveraging the traditional offload capabilities offered by NICs for performance. The ability to insert additional context into the overlay header unlocks doors for future innovations in context awareness, end-to-end telemetry, security, encryption and more. Additional information about the GENEVE industry standard can be found at https://tools.ietf.org/html/draft-gross-geneve-00 - section-1.2.

The host switch is manifested as a variant of the VMware virtual switch on ESXi-based endpoints and as Open Virtual Switch (OVS) on KVM-based endpoints. Additional information on OVS can be found at http://openvswitch.org/.

Routing Scale and Performance

The NSX-T provides next-generation optimized routing, which allows minimal configuration while supporting both distributed and localized forwarding. The concept of tenancy is built into the routing model.



Multi-tenant routing model

NSX-T supports a multi-tiered routing model with logical separation between the provider router function (known in NSX as a Tier0 router) and the tenant router function (known in NSX as a Tier1 router).

The **Tier0** logical router, a routing layer which can be controlled by the cloud provider, is capable of peering with the physical infrastructure.

The **Tier1** logical routers, a routing layer provided to cloud tenants, can be provisioned by GUI/API/CMPs per tenant and be attached to the Tier0 routers.

NSX-T instantiates distributed routers on the hypervisors for optimal multi-tier E-W routing.

Connectivity between the tenant routers and the provider routers is managed by the NSX-T management plane and the NSX-T control plane, and there is no requirement to run complex routing protocols or control VM appliances to disseminate connectivity information.

High Performance Edge Nodes

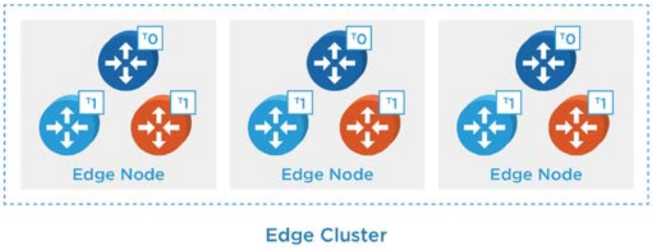
NSX-T enables high-performance edge nodes to provide Tier0 connectivity to the physical infrastructure. Edge nodes also provide capacity to run tenant- or provider-based centralized network services like NAT. Edge nodes can be deployed as both VM or bare metal form factor. These nodes leverage improvements in x86 forwarding technology like Intel's DPDK Libraries (DataPlane Development ToolKit) for faster packet processing. More information can be found at http://www.intel.com/go/DPDK.

These improvements enable line-rate gateway performance at small packet sizes. The edge nodes can further scale out to provide multiple Gbps of throughput.

Data Plane Scale Out

The edge nodes can be grouped into a pool of capacity (known as edge cluster) to provide scale out, redundant and high-throughput gateway functionality for the logical networks. Scale out from the logical networks to the edge nodes is achieved using equal cost multi path (ECMP). The edge nodes can also host stateful services and provide scale and redundancy to run these services at scale. The active standby model is also available to provide flexibility and choice.

Multiple tenant-based or provider-based services can run as individual contexts on any node in the edge cluster as shown in the figure.

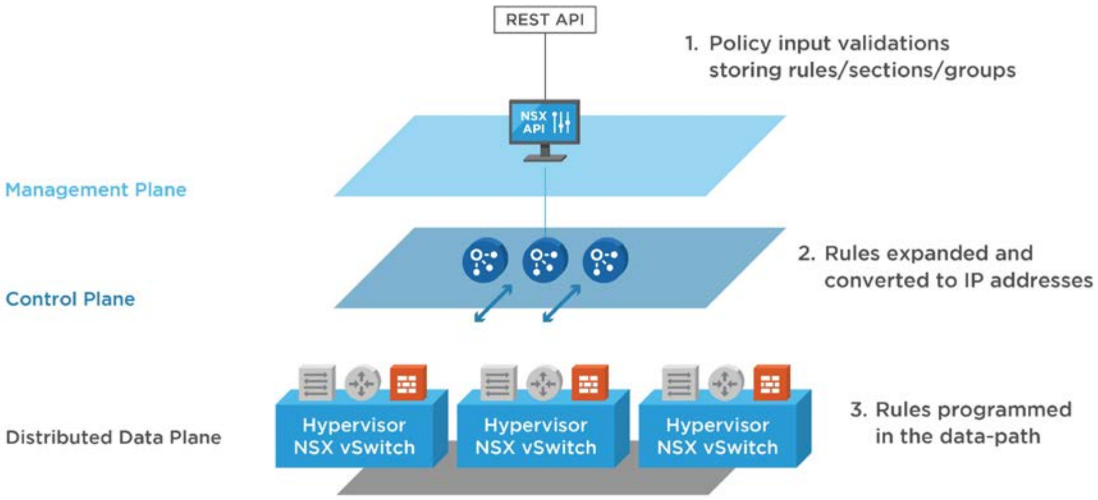


Routing and Convergence

NSX-T supports protocol-standards-based static and dynamic routing to peer with the physical networking world. The edge node is architected to run a large number of tenant routing contexts within a virtual/BM appliance and leverage advanced routing functionality, while also supporting BGP routing and attribute-manipulation capabilities using route maps and prefix lists.

For faster detection of link or node failures and fast routing convergence, bidirectional forwarding detection (BFD) towards the physical space is available. BFD is also used internally for fault detection and fast failover of nodes.

Security


REST API
1. Policy input validations storing rules/sections/groups
NSX API
Management Plane
Control Plane
2. Rules expanded and converted to IP addresses
Distributed Data Plane
Hypervisor NSX vSwitch
Hypervisor NSX vSwitch
Hypervisor NSX vSwitch
3. Rules programmed in the data-path

NSX-T introduces support for distributed firewall (DFW) functionality for workloads running on both ESXi and KVM hosts/hypervisors. The NSX DFW provides the capability to enforce firewalling functionality directly at the workload vNIC layer, providing an optimal micro-segmented environment. Both stateful and stateless firewall rules are supported. Firewall rule provisioning will be proliferated in the system by the CCP for more redundancy and scale. Users will have the ability to craft their micro-segmentation model based on IP address sets (L3), MAC identifier sets (L2), logical switches, logical ports or advanced security policies based on security groups. CMPs and PaaS platforms will be able to provision policies on NSX-managed workloads using the advanced REST API or pluggable modules. Operators will be able to view provisioned policies in a single pane of glass in the NSX Manager GUI or via the API.

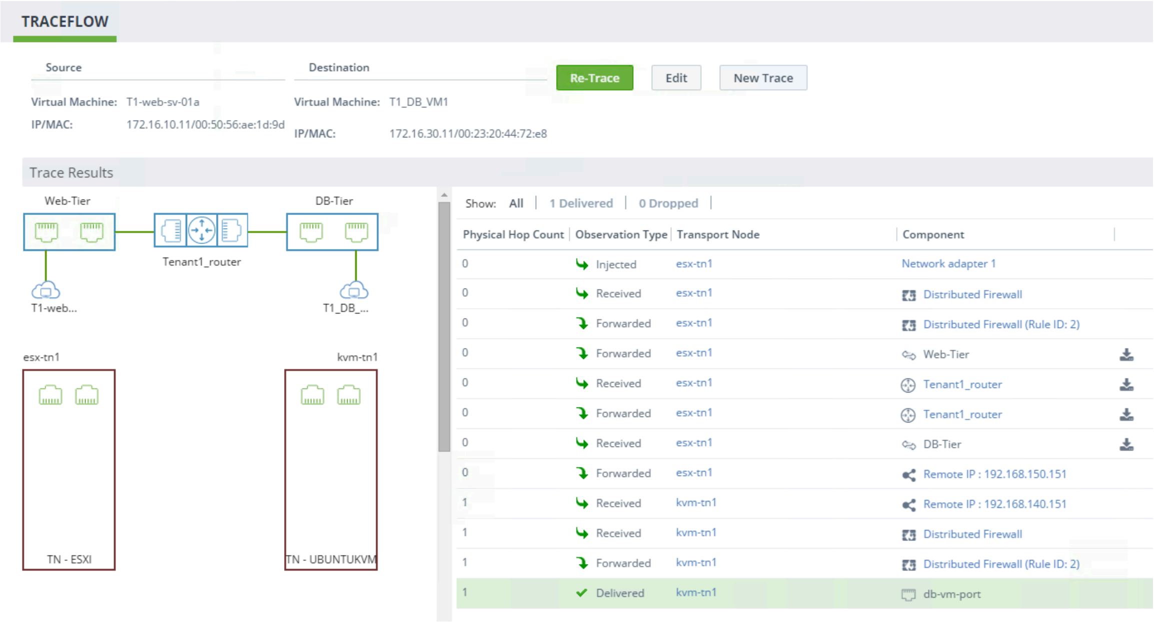
Services

NSX-T offers a set of native services such as DHCP relay, DHCP server, MetaData proxy, Stateful/Stateless NAT which can be leveraged by the CMP to offer both scale and functionality. The platform is architected to enable flexibility of running multiple services. The platform has an architectural design that allows it to support native load balancing, perimeter firewalling and service chaining functionality.

Operations

In heterogeneous environments, having consistent operations, monitoring and troubleshooting becomes even more crucial than in homogeneous deployments. NSX-T has the operational toolbox for managing and troubleshooting complex environments running on heterogeneous infrastructures.

NSX-T has tools like Port Connectivity and Traceflow which help users trace connectivity through virtual and physical devices and detect failures.



TRACEFLOW

Source
Virtual Machine: T1-web-sv-01a
IP/MAC: 172.16.10.11/00:50:56:ae:1d:9d

Destination
Virtual Machine: T1_DB_VM1
IP/MAC: 172.16.30.11/00:23:20:44:72:e8

Trace Results

Show: All | 1 Delivered | 0 Dropped |

Physical Hop Count	Observation Type	Transport Node	Component
0	Injected	esx-tn1	Network adapter 1
0	Received	esx-tn1	Distributed Firewall
0	Forwarded	esx-tn1	Distributed Firewall (Rule ID: 2)
0	Forwarded	esx-tn1	Web-Tier
0	Received	esx-tn1	Tenant1_router
0	Forwarded	esx-tn1	Tenant1_router
0	Received	esx-tn1	DB-Tier
0	Forwarded	esx-tn1	Remote IP : 192.168.150.151
1	Received	kvm-tn1	Remote IP : 192.168.140.151
1	Received	kvm-tn1	Distributed Firewall
1	Forwarded	kvm-tn1	Distributed Firewall (Rule ID: 2)
1	✓ Delivered	kvm-tn1	db-vm-port

An example of Traceflow in action is shown in the figure where a trace is being run between 2 VMs running on ESX and KVM hypervisors which are attached to Web-Tier and DB-Tier logical switches.

NSX-T provides a robust logging framework with consistent log formats, trackable event identifiers and error codes, and tags across the distributed components to help build a correlated log-based dashboard on centralized log collectors such as VMware vRealize Log Insight.

NSX aggregation services provide a centralized view of information like statistics, routes and MAC table information from distributed components in a single pane of glass without logging into those individual components.

NSX-T provides granular flow and packet-level visibility through standards tools such as IPFIX and port mirroring. This enables customers to use their existing monitoring and troubleshooting tools for visibility when troubleshooting.

Conclusion

NSX-T is built to support the increasingly heterogeneous and app-driven realities of digital business environments. As IT's role in this landscape becomes more complex, NSX-T provides an approach to support, manage and secure workloads running inside multiple clouds, containers and hypervisors, distributed across microservice architectures. And by enabling developers to consume networking resources via APIs or natively via Openstack NSX Neutron Plugin, NSX-T lets them treat infrastructure-as-code (IaC) within the context of their build and CI/CD pipeline, and get more work done, faster.

Key Takeaways

- NSX-T architecture is designed for heterogeneity and choice, and will evolve to provide networking and security services to workloads running anywhere.
- NSX provides a rich set of networking and security services with enterprise-grade quality and scale for workloads running on ESXi or KVM hosts/hypervisors.
- NSX provides a single pane of glass and tools to configure, monitor and troubleshoot networking and security on heterogeneous infrastructures.

- The NSX REST API is powerful and comes with an OpenAPI specification that enables various language bindings and plugins to integrate with the CMP of your choice, including OpenStack.
- NSX OpenStack plugin is available for developers who want to build and maintain multi-tenant developer clouds with advanced services.
- NSX edge nodes provide extreme gateway and services performance using innovation in x86 forwarding and Intel DPDK technology.
- NSX is vCenter agnostic and can be consumed by vCenter, other compute managers as well as PaaS platforms.

What You Can Do Now

As a next step, you can contact your VMware sales representative for an overview and demonstration of NSX-T.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com