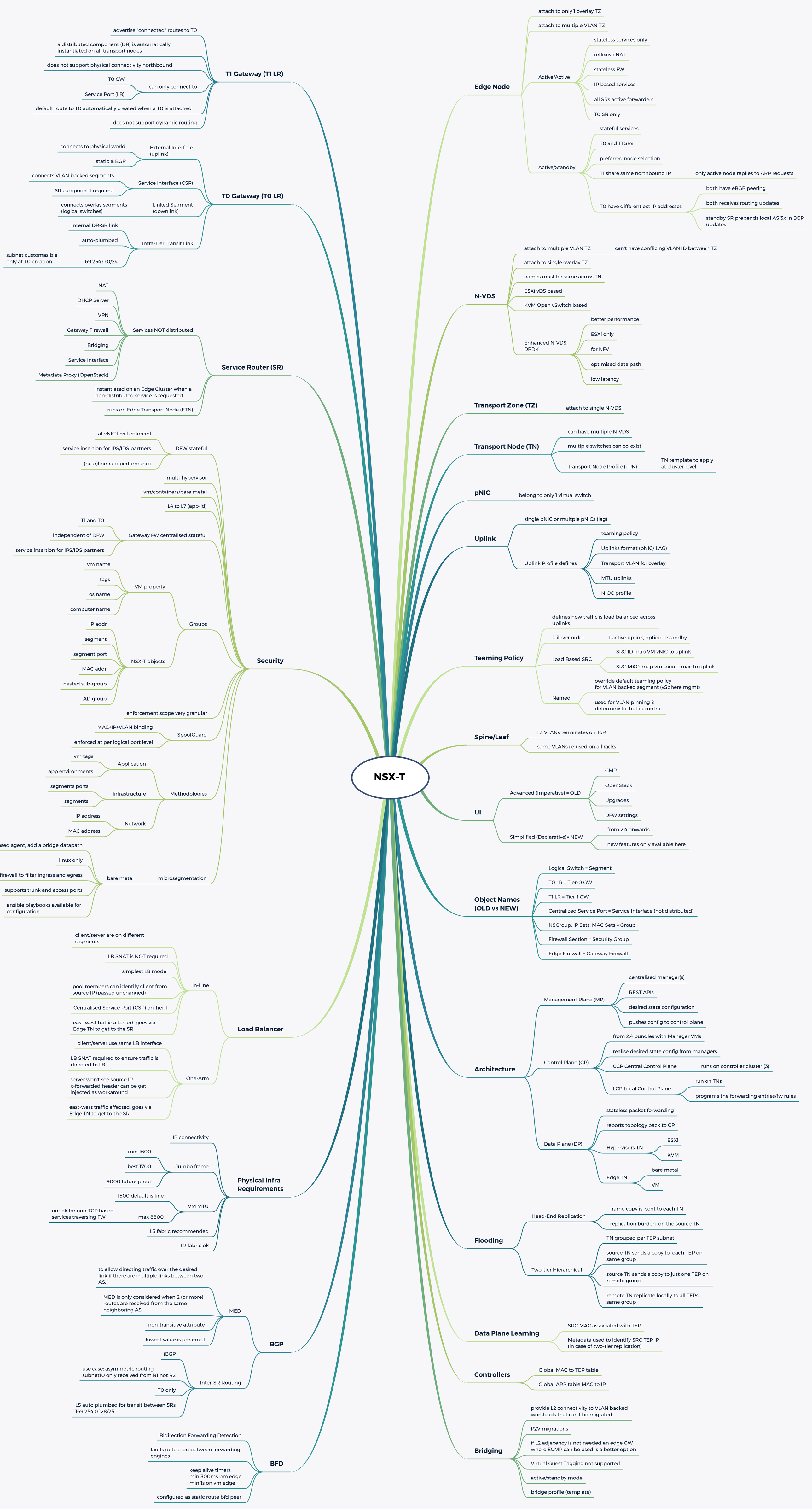
NSX-T
T1 Gateway (T1 LR)
advertise "connected" routes to T0
a distributed component (DR) is automatically instantiated on all transport nodes
does not support physical connectivity northbound
can only connect to
T0 GW
Service Port (LB)
default route to T0 automatically created when a T0 is attached
does not support dynamic routing
T0 Gateway (T0 LR)
External Interface (uplink)
connects to physical world
static & BGP
Service Interface (CSP)
connects VLAN backed segments
SR component required
Linked Segment (downlink)
connects overlay segments (logical switches)
Intra-Tier Transit Link
internal DR-SR link
auto-plumbed
169.254.0.0/24
subnet customisible only at T0 creation
Service Router (SR)
Services NOT distributed
NAT
DHCP Server
VPN
Gateway Firewall
Bridging
Service Interface
Metadata Proxy (OpenStack)
instantiated on an Edge Cluster when a non-distributed service is requested
runs on Edge Transport Node (ETN)
Security
DFW stateful
at vNIC level enforced
service insertion for IPS/IDS partners
(near)line-rate performance
multi-hypervisor
vm/containers/bare metal
L4 to L7 (app-id)
Gateway FW centralised stateful
T1 and T0
independent of DFW
service insertion for IPS/IDS partners
Groups
VM property
vm name
tags
os name
computer name
NSX-T objects
IP addr
segment
segment port
MAC addr
nested sub-group
AD group
enforcement scope very granular
SpoofGuard
MAC+IP+VLAN binding
enforced at per logical port level
Methodologies
Application
vm tags
app environments
Infrastructure
segments ports
segments
Network
IP address
MAC address
microsegmentation
bare metal
sed agent, add a bridge datapath
linux only
firewall to filter ingress and egress
supports trunk and access ports
ansible playbooks available for configuration
Load Balancer
In-Line
client/server are on different segments
LB SNAT is NOT required
simplest LB model
pool members can identify client from source IP (passed unchanged)
Centralised Service Port (CSP) on Tier-1
east-west traffic affected, goes via Edge TN to get to the SR
One-Arm
client/server use same LB interface
LB SNAT required to ensure traffic is directed to LB
server won't see source IP x-forwarded header can be get injected as workaround
east-west traffic affected, goes via Edge TN to get to the SR
Physical Infra Requirements
IP connectivity
Jumbo frame
min 1600
best 1700
9000 future proof
VM MTU
1500 default is fine
max 8800
not ok for non-TCP based services traversing FW
L3 fabric recommended
L2 fabric ok
BGP
MED
to allow directing traffic over the desired link if there are multiple links between two AS.
MED is only considered when 2 (or more) routes are received from the same neighboring AS.
non-transitive attribute
lowest value is preferred
Inter-SR Routing
iBGP
use case: asymmetric routing subnet10 only received from R1 not R2
T0 only
L5 auto plumbed for transit between SRs 169.254.0.128/25
BFD
Bidirection Forwarding Detection
faults detection between forwarding engines
keep alive timers
min 300ms bm edge
min 1s on vm edge
configured as static route bfd peer
Edge Node
attach to only 1 overlay TZ
attach to multiple VLAN TZ
Active/Active
stateless services only
reflexive NAT
stateless FW
IP based services
all SRs active forwarders
T0 SR only
Active/Standby
stateful services
T0 and T1 SRs
preferred node selection
T1 share same northbound IP
only active node replies to ARP requests
T0 have different ext IP addresses
both have eBGP peering
both receives routing updates
standby SR prepends local AS 3x in BGP updates
N-VDS
attach to multiple VLAN TZ
can't have conflicing VLAN ID between TZ
attach to single overlay TZ
names must be same across TN
ESXi vDS based
KVM Open vSwitch based
Enhanced N-VDS DPDK
better performance
ESXi only
for NFV
optimised data path
low latency
Transport Zone (TZ)
attach to single N-VDS
Transport Node (TN)
can have multiple N-VDS
multiple switches can co-exist
Transport Node Profile (TPN)
TN template to apply at cluster level
pNIC
belong to only 1 virtual switch
Uplink
single pNIC or multple pNICs (lag)
Uplink Profile defines
teaming policy
Uplinks format (pNIC/ LAG)
Transport VLAN for overlay
MTU uplinks
NIOC profile
Teaming Policy
defines how traffic is load balanced across uplinks
failover order
1 active uplink, optional standby
Load Based SRC
SRC ID map VM vNIC to uplink
SRC MAC: map vm source mac to uplink
Named
override default teaming policy for VLAN backed segment (vSphere mgmt)
used for VLAN pinning & deterministic traffic control
Spine/Leaf
L3 VLANs terminates on ToR
same VLANs re-used on all racks
UI
Advanced (Imperative) = OLD
CMP
OpenStack
Upgrades
DFW settings
Simplified (Declarative)= NEW
from 2.4 onwards
new features only available here
Object Names (OLD vs NEW)
Logical Switch = Segment
T0 LR = Tier-0 GW
T1 LR = Tier-1 GW
Centralized Service Port = Service Interface (not distributed)
NSGroup, IP Sets, MAC Sets = Group
Firewall Section = Security Group
Edge Firewall = Gateway Firewall
Architecture
Management Plane (MP)
centralised manager(s)
REST APIs
desired state configuration
pushes config to control plane
Control Plane (CP)
from 2.4 bundles with Manager VMs
realise desired state config from managers
CCP Central Control Plane
runs on controller cluster (3)
LCP Local Control Plane
run on TNs
programs the forwarding entries/fw rules
Data Plane (DP)
stateless packet forwarding
reports topology back to CP
Hypervisors TN
ESXi
KVM
Edge TN
bare metal
VM
Flooding
Head-End Replication
frame copy is sent to each TN
replication burden on the source TN
Two-tier Hierarchical
TN grouped per TEP subnet
source TN sends a copy to each TEP on same group
source TN sends a copy to just one TEP on remote group
remote TN replicate locally to all TEPs same group
Data Plane Learning
SRC MAC associated with TEP
Metadata used to identify SRC TEP IP (in case of two-tier replication)
Controllers
Global MAC to TEP table
Global ARP table MAC to IP
Bridging
provide L2 connectivity to VLAN backed workloads that can't be migrated
P2V migrations
if L2 adjecency is not needed an edge GW where ECMP can be used is a better option
Virtual Guest Tagging not supported
active/standby mode
bridge profile (template)