Validated Reference Design
NetScaler SDX Clustering

This guide focuses on providing guidelines to customers on implementing SDX Clustering in NetScaler based on their use cases.
# Table of Contents

**Section 1: SDX Clustering Overview**  
Pre-Requisites: NetScaler SDX Clustering  
Layer 2 and Layer 3 Connectivity  
Cluster Traffic Distribution Methods  
Inter-node Communications  
Internal Traffic Distribution Methods  
Cluster Coordinator (CCO)  
IP Topology Requirements  
IP Requirements  
Static Route Management  
SVM and NSIP Requirements  
Dynamic Routing Requirements  

**Section 2: Design - NetScaler SDX Cluster**  
NetScaler Physical Design  

**Section 3: NetScaler Appflow for Clustering and HA**  
Introduction  
Proxy User Use Case  
LAN User Use Case  
NetScaler LAN Configuration  
Storefront Configuration  
HDX Insight Requirements for Cluster Environment  
Session Reliability Behavior  
Cluster Information  
HA Information  
Appflow Records  
Configuration  

**Section 4: Configuration: NetScaler SDX Cluster**  
SDX Build  
Cluster Build  
Route Management  
Create Spotted SNIP(s)  
Configure Virtual Server  
Configure Dynamic Routing  
Configuration Example: NetScaler SDX Cluster  

**Appendix A: Cluster Troubleshooting**  

**Appendix B: SVM Management**  

**Appendix B: OSPF Reference**
Section 1: SDX Clustering Overview

This handbook is intended to guide resources through the practical implementation of NetScaler SDX appliances configured in a Cluster Mode. This configuration is based on a multi-tenant deployment scenario which requires data plane and traffic plane isolation between tenants using unique VLAN ID’s per tenant. The details in this document are SDX model specific and caveats may apply to VPX and MPX appliance editions.

Pre-Requisites: NetScaler SDX Clustering

- A NetScaler cluster can only include NetScaler nCore appliances. Clustering of NetScaler Classic appliances is not supported.
- FIPS 140-2 Level 2 is not supported in NetScaler Clustered appliances. Thales nShield Connect appliances can be used for this requirement.
- All appliances must have the same software version and build.
- All appliances must be of the same platform type. This means that a cluster must have either all hardware appliances (MPX) or virtual appliances (VPX) or SDX NetScaler instances.
- All appliances must be on the same network.
- All appliances must have the same licenses. Also, depending on the NetScaler version, there are some additional aspects to address:
  - Be initially configured and connected to a common client-side and server-side network.
  - Pooled Licensing Mode is not supported.


Layer 2 and Layer 3 Connectivity

Dynamic Routing Requirements

If the NetScaler SDX Cluster requires that more than one VLAN be carried over a network interface on the SDX appliance, the ECMP topology is required. (As a rule of thumb, always use ECMP routing with SDX clustering.) The ECMP topology uses the dynamic routing method to achieve equal cost traffic distribution across the member nodes of the SDX cluster for Client traffic distribution.

Nexus Switch(s) Requirements

NOTE:

- Some Nexus Switch models do not support Dynamic Routing protocols over VPC.
- Nexus 7K Series switches require a minimum firmware build in order to support dynamic routing protocols over VPC.
- Nexus 9K series switches do not support dynamic routing protocols over VPC.

Jumbo Frame Support

Jumbo frame support is required when an SDX cluster has backplane data that traverses more than one physical SDX chassis. This requirement is due to the additional tuple information that is added to the Ethernet header.
to include the cluster member interface id. The result is an Ethernet frame that is 1514 bytes, thus exceeding the standard 1500 MTU Ethernet frame.

EXAMPLE:
First node in cluster uses 2-tuple:
  i.e int 1/1/1   MTU1500

Subsequent nodes use 3-tuple:
  i.e int 1/1/2   MTU1514

Cluster Traffic Distribution Methods:

CLAG: Single VLAN per Interface

SDX can support Cluster Link Aggregation Groups on a per interface basis, so long as there is only one network per SDX interface. (This method is rarely used with SDX clustering due to the limitation to only one VLAN per interface.) CLAG is an L2 Channel that can be either Static or Dynamic and the upstream switch sees a single cluster MAC address in the ARP table.

CLAG NOTES:
1. A separate physical medium is required for Client connection steering and node-to-node communications.
2. Cluster Heartbeats cannot be exchanged over the CLAG interfaces.
3. VPX appliances are not supported with CLAG, some ESX and KVM versions can support CLAG.

ECMP Multiple VLAN’s per Interface

SDX requires the use of Equal Cost Multipath via dynamic routing protocols if more than one VLAN is traversing an SDX interface and will use the Flow Receiver (FR) and Flow Processor(FP) for Client connection steering across the cluster backplane.

ECMP NOTES:
1. ECMP is an L3 Routing mechanism that has a limitation of the supporting router regarding the number of ECMP routes that can be supported.
2. Cluster Heartbeats can be exchanged over the ECMP interfaces (not recommended).

Inter-node Communications:

Cluster nodes within the same Cluster Instance communicate with each other by using the cluster backplane. The backplane is a set of connections in which one interface of each node is connected to a common switch or VLAN, which is called the cluster backplane switch or VLAN. Each node of the cluster uses a special MAC address to communicate with other nodes through the backplane.

Internal Traffic Distribution Methods:

An upstream L3 device will use a Hash based algorithm based on the routes received from ECMP to determine which cluster node will receive the incoming client connection. The cluster node that receives the client connection from the upstream L3 device is known as the Flow Receiver (FR).
Flow Processor (FP):

The Flow Processor is the cluster node that is assigned the packet for the initial client connection and establishes the traffic flow. The Flow Processor will process the client traffic and respond to the client using its own link.

Flow Receiver (FR):

The Flow Receiver is the cluster that receives the inbound client packet. Sometimes the Flow Receiver is also the Flow Processor in which case no further packet steering is required. When the Flow Receiver receives a packet that is part of a flow for which it is not the Flow Processor, it determines the Flow Processor (FP) for the traffic stream and steers the packet to the cluster node that owns the flow (FP). All packet steering happens via the data backplane, if the backplane is unavailable, the packet is dropped.

Spotted entity: Any configuration element that is Node specific and does not belong to all nodes in the cluster. FP = node in processing set.

Striped entity: processing set = multiple nodes. Hash computed on the packet parameters. (TCP and UDP protocols use -4 tuple computation, any other protocols other IP -2 tuple computation)

An ACTIVE node from processing set is selected as FP based on the hash. Global Key synchronized across all the nodes to compute consistent hash.

Cluster Coordinator (CCO)

All configurations on a NetScaler cluster are performed on the cluster IP address, which is the management address of the cluster. This cluster IP address is owned by a cluster node that is referred to as the cluster configuration coordinator.

When a node is added to a cluster, the configurations and the files (SSL certificates, licenses, DNS, and so on) that are available on the cluster configuration coordinator are synchronized to the newly added cluster node. When an existing cluster node, that was intentionally disabled or that had failed, is once again added, the cluster compares the configurations available on the node with the configurations available on the configuration coordinator. If there is a mismatch in configurations, the node is synchronized by using one of the following:

Full synchronization. If the difference between configurations exceeds 255 commands, all the configurations of the configuration coordinator are applied to the node that is rejoining the cluster. The node remains operationally unavailable for the duration of the synchronization.

Incremental Synchronization. If the difference between configurations is less than or equal to 255 commands, only the configurations that are not available are applied to the node that is rejoining the cluster. The operational state of the node remains unaffected.

Determining the CCO:

Every command propagated in cluster will be assigned a number called ARU number. It is a sequentially increasing number and is assigned by the CCO node. ARU number is used to detect configuration mismatch and trigger sync.

• The node which has remained CCO for a longer time (applicable only for 2-node)
• The node which has highest ARU (All Received Upto) value (see section below)
• The node which has better Operational View – has the most visibility of the cluster node neighbors.
• Priority – (0-31) numeric priority of the node members, with 0 being the highest priority.
• Previous CCO – the node of the cluster that was the previous elected CCO.
• The node having lowest IP.
All Received Upto (ARU)

When we issue force sync on a node, it resets its local ARU to 0, triggers a sync from the CCO node and sets local ARU to highest ARU after sync is complete.

Counters to view ARU: nsconmsg -g aru -d stats

SCENARIO:
Say we have a cluster with 3 nodes \{n1, n2, n3\} and n1 is the CCO. ARU starts from 0.

1. Command c1 is fired on the cluster IP address. CCO assigns ARU 1 to command c1 and propagates the command to other nodes.
2. Command c2 gets fired on cluster IP address. ARU 2 gets assigned to c2 and so on.
3. Say node n3 is offline when command c3 is fired. ARU 3 gets assigned to c3 and the command is applied on the online nodes n1 and n2. At this stage, the local ARU of n1 and n2 is 3 whereas that of node n3 is 2.
4. When n3 comes back online, it sees that the highest ARU number is 3 (ARU is advertised via heartbeat, so highest ARU can be derived), and it triggers a sync to catch up.

IP Topology Requirements

Data Plane (Cluster Back Plane)

The SDX cluster backplane is an isolated network and requires NSIP traffic on the the same network as the inter-cluster communication traffic for the SDX cluster. The data plane traffic is specific to the cluster heartbeat and traffic steering from the cluster coordinator (CCO).

Backplane Traffic Includes:
- NSIP to NSIP traffic - cluster heartbeats, cluster protocol, config propagation/synchronization, NNM
- Data traffic / steering (srcmac:steering mac, destmac:backplane mac, IP tuple remains same as client/server packet)

Traffic Plane (Client / Server Traffic Plane)

The SDX traffic plane is an isolated traffic network that routes directly to the Client network and is where the Virtual Server IP’s will be terminated for incoming client requests. The Client and Server networks can be the same or different networks in ECMP mode. A single network is used for a one armed configuration and two or more networks are used for a multiple armed configuration.

VLAN Separation

There is not any limitation on VLAN separation in regards to the Data plane or the Traffic plane. Jumbo frames (MTU increase) will still be required on the data plane regardless of VLAN tagging due to the cluster interface tuple and the changes made to the MTU (1514).

The SDX appliance has a limitation of 63 VLAN’s per 10 Gpbs interface if the filtering mechanism is configured in the SVM. If there are not any VLAN filters configured in the SVM, the SDX appliance can accommodate 4096 VLAN’s per interface.
IP Requirements

Cluster IP: (CLIP) The CLIP is owned by the cluster coordinator (CCO) and this IP address will float to the active CCO node. The CCO is responsible for configuration synchronization, command propagation, and file synchronization (certificates, CRL, etc...). Think of the Cluster IP as similar to the NSIP for the whole cluster.

Striped Subnet IP: (SNIP) The striped SNIP address are owned by multiple nodes members in the cluster and will service all server proxy connections to the backend services. When the cluster initiates traffic such as service monitors, it will usually be sourced from the Striped Cluster SNIP.

Spotted Subnet IP: (SNIP) The spotted SNIP is the network host address that is used to advertise routes to the upstream L3 device for each node member of the cluster to allow for Hash distribution of incoming Client connections.

Striped Virtual IP: (VIP) The striped IP addresses are owned by multiple nodes members in the cluster and will service all Client connections based on the Hash table of the upstream distribution L3 device. This is the address that is being advertised via ECMP.

Spotted Virtual IP: (VIP) The spotted IP address is owned by a single node in the cluster and will utilize the cluster steering of the Flow Receiver via the Cluster Data plane to process inbound Client connections to the cluster node owner of the spotted VIP. A Spotted Cluster VIP would be used for features which are available in node only mode.

Virtual Server load balancing methods can be applied to the VIP's for service distribution as well.

Static Route Management

Overview

When the NetScaler instances are provisioned, they will only have an NSIP address present on the Cluster Backplane VLAN and a default route pointed to the gateway on the Cluster Backplane VLAN. After the cluster has been created, additional VLANs will be added to the cluster and configured on different interfaces. This will require static routing changes on the cluster.

Policy Based Routes

After the cluster is created, a PBR needs to be created on the cluster with a source IP range starting with the NSIP of each member node and including the CLIP. The next hop gateway will be the IP address of the gateway on the cluster backplane VLAN.

Default Routes

After the PBR has been created and applied, the default route for the cluster needs to be changed. The default route should be through the gateway on the client traffic VLAN.

VLAN's

VLAN Data / NSIP and CLIP (Cluster Backplane)
VLAN Traffic / Striped and Spotted SNIPs (Traffic)
VLAN SVM/ XS (Chassis Management)
VLAN LOM (ILO VLAN)
**SVM and NSIP Requirements**

**SVM and NSIP Communication Requirements**

While it is possible to run non-clustered NetScaler instances on an SDX where the SVM does not have communication with the NSIP of the instances, it is not recommended. When clustering is being used, this is not an option. The SVM must be able to communicate with the NSIP of all cluster instances at all times. Failure to maintain SVM to NSIP communication will lead to orphaned cluster instances which will exist in a bad state. In some cases, the best way to remedy the situation after SVM to NSIP communication is restored, is to delete the instance, recreate it and rejoin it to the cluster.

**Cluster Heartbeats**

Node to Node cluster heartbeats from NSIP to NSIP should be carried over the cluster data plane whenever possible. When ECMP mode is used, Node to Node cluster heartbeats can be carried over the Traffic plane via routing, however, special care should be taken in this model as upstream layer 3 devices can potentially interfere with cluster heartbeats. (Not recommended)

**SVM for Provisioning and Management of Cluster Node Groups**

Do not attempt to provision the cluster from the administration GUI of the NetScaler instances. The Cluster must be created from the SVM and Nodes must be added or removed from the SVM. If you create a cluster or add cluster elements from the NetScaler instance administration GUI, the resultant cluster will be unstable.

**Dynamic Routing Requirements**

**OSPF**

OSPF routing is required to have the customer’s network infrastructure distribute traffic between the different nodes in the NetScaler cluster. In order to configure the NetScaler to participate in OSPF routing with the customer, several steps are required. OSPF routing must be enabled for the cluster globally. The vServers that we want to advertise to OSPF must have route health injection configured. The VIP must be configured in the NetScaler System IP menus to participate in OSPF routing using Type 1 LSAs for the proper OSPF area assigned by the customer. ZebOS must be configured from VTYSH to advertise OSPF routes learned from the NetScaler kernel. Each node in the cluster must be given a distinct router ID in VTYSH. OSPF routing can be validated by running the show ip ospf neighbor command from VTYSH. OSPF route advertisement can be disabled by showing the ip routes learned from the NetScalers on the customer infrastructure.
Section 2: Design - NetScaler SDXCluster

NetScaler Physical Design

Physical Cabling Diagram:

Physical To Virtual Example Diagram
Node Groups

Default Node Group

Member Nodes

N+1 Default Clusters:
- 1 node on Appliance (1)
- 1 node on Appliance (2)
- 1 node on Appliance (3)

NetScaler Logical Diagram
Section 3: NetScaler Appflow for Clustering and HA

Introduction

A NetScaler Cluster solution refers to a group of NetScalers which can be configured and managed as a single system. It provides scalability and availability.

A NetScaler High Availability solution (Active-Passive) refers to two NetScalers forming a pair in which one is actively processing the traffic while the other is a stand-by. Both the netscaler are in sync with respect to configuration. When the primary netscaler goes down for some reason, the secondary immediately takes-over, thus providing the user a minimal disruption.

The HDX Insight solution consists of Appflow, which is a method of logging protocol and flow related information. The Netscaler exports this information and a collector (external to NetScaler) parses and stores this information for subsequent viewing in required format.

This document describes the design for supporting appflow on a group of NetScalers that belong to a cluster or two NetScalers that form a HA pair.

Proxy User Use Case

Provides HDX Insight Deep Packet Inspection (DPI) for all ICA Client connections from the External Internet (WAN) when scaling cluster deployments in a single tier architecture, once an HA Platform has reached the maximum number of ICA connections per location.

Additionally, using NetScaler Clusters will provide capacity for Dynamic Scaling requirements without the need to rebuild the network topology or placement of the NetScaler appliances as the environment grows.

LAN User Use Case

Provide HDX Insight Deep Packet Inspection (DPI) for all ICA Client connections from the Internal Intranet (LAN) when scaling cluster deployments in a single tier architecture, once an HA Platform has reached the maximum number of ICA connections per location.
NetScaler LAN Configuration

• Create a CR vserver with type as HDX.
• Bind an appflow policy of the type ICA_REQ_DEFAULT globally.
• Below is an example configuration:
  
  ```
  add appflow collector col1 -IPAddress <collector IP>
  add appflow action act1 -collectors col1
  add appflow policy pol2 true act2
  bind appflow global pol2 1 END -type ICA_REQ_DEFAULT
  add cr vserver crvs HDX <crvserver IP> <Port> -cacheType FORWARD -cltTimeout 180
  ```

Storefront Configuration

• Edit “default.ica” file of the store that is used to launch apps/desktops.
• This file will be present in “C:\inetpub\wwwroot\Citrix\<Store>\App_Data”.
• Following configuration needs to be added under WFClient and Application section
  
  ```
  ProxyType=Socks
  ProxyHost=<crvserver IP>:<Port>
  ICASOCKSProtocolVersion=0
  ICASOCKSProxyHost=<crvserver>
  ICASOCKSProxyPortNumber=<Port>
  ```

HDX Insight Requirements for Cluster Environment

The following prerequisites are required when using ICA Proxy and HDX Insight features in a NetScaler Cluster:

• All Nodes in the Cluster must be of the same Model of NetScaler Platform
• All Nodes in the Cluster must have the same License Mode for Cluster Members
  (NetScaler Platinum Edition is required for HDX Insight)
• Standard ICA Cluster is supported with the Enterprise (EE) and Standard (STD) License Modes

Session Reliability Behavior

Session Reliability Behavior: (when all Nodes are working in the Cluster)

• All Session Contexts are shared among the Healthy Node members and will use the Back Plane for connection communication.
• If, a Node member has died, the Session Reliability information becomes invalid, and Auto Client Reconnect is the reconnect method for disconnected users via the Citrix Receiver Client for Windows and MAC.
• If the ACR Feature is not supported for that users Citrix Receiver Client, the user will need to reconnect the session.

The Reconnect info is shared via the Back Plane, via the Inter-Node communication
This is standard traffic overhead for the reconnect operation.
The Backplane Overhead will reduce once the connections are made via disconnect and reconnect
Cluster Information

When multiple NetScalers belonging to a cluster export appflow records, it becomes necessary to correlate/aggregate them to get a system-level view. Since each NetScaler exports the data records independently, all the nodes must export an additional information in all of their records indicating which cluster they belong to.

This field, called the observation domain ID, will contain the cluster ID as configured by the administrator. Observation domain ID is part of the ipfix header.

It is also desirable to display a name when the aggregated data is presented. This name, called the observation domain name, is associated with an observation domain and exported in the observation domain information record.

The IP address of each node in the observation domain is also exported in the observation domain information record.

HA Information

When two NetScalers form a primary-secondary pair, one processes traffic and sends appflow records while the other is stand-by. However, when a switch-over happens, the new primary starts sending appflow records. Records from both the NetScalers are required to be correlated. When both the NetScalers use the same ID, the collector can merge the records at the HA-pair level irrespective of which NS is processing the traffic currently. This ID is configured by the administrator and is exported as observation domain ID in all the records. It is also required to send the NSIP of the netscaler in all the records because in case of SNIP, both the NSs use the same source IP. Similar to cluster, the observation domain name and the nodes (primary and secondary) IP addresses are exported.

Appflow Records

A new record called the observation domain information record will be added to export information such as the observation domain ID, observation domain name and the IP addresses of the nodes constituting the observation domain. The observation domain id is also part of every record as it appears in the ipfix header. Also, the observation point ID will carry the NSIP of the respective node that is exporting the records.

Observation Domain Information (Appflow Template ID: 291):

<table>
<thead>
<tr>
<th>IE-ID</th>
<th>Enterprise ID</th>
<th>Source of Inputs</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>149</td>
<td>0</td>
<td>observationDomainId</td>
<td>4-bytes</td>
</tr>
<tr>
<td>352</td>
<td>5951</td>
<td>numberOfObservationPoints</td>
<td>2-bytes</td>
</tr>
<tr>
<td>353</td>
<td>5951</td>
<td>observationPointId_1 (NSIP)</td>
<td>4-bytes</td>
</tr>
<tr>
<td>354</td>
<td>5951</td>
<td>observationPointId_2 (NSIP)</td>
<td>4-bytes</td>
</tr>
<tr>
<td>355</td>
<td>5951</td>
<td>observationPointId_3 (NSIP)</td>
<td>4-bytes</td>
</tr>
<tr>
<td>356</td>
<td>5951</td>
<td>observationPointId_4 (NSIP)</td>
<td>4-bytes</td>
</tr>
<tr>
<td>357</td>
<td>5951</td>
<td>observationPointId_5 (NSIP)</td>
<td>4-bytes</td>
</tr>
<tr>
<td>300</td>
<td>0</td>
<td>observationDomainName</td>
<td>String (variable)</td>
</tr>
</tbody>
</table>
Configuration

The administrator is required to configure a domain ID on the CCO or primary node. This is propagated to other nodes of the cluster or secondary node of HA-pair. Similarly, an observation domain name can also be configured via the following Set commands:

- set appflow param observationDomainID <num>
- set appflow param observationDomainName <name>

Sample AppFlow commands for an ICA PRoxt Cluster:

- add vpn vserver ICAProxy_vip SSL 192.168.x.x 443 -icaOnly ON -downStateFlush DISABLED -appflow Log ENABLED
- add appflow collector Cluster_Collector -IPAddress 10.x.x.x
- set appflow param -observationDomainId 1020 -observationDomainName Cluster_Domain
- add appflow action appflow_act_cluster -collectors Cluster_Collector
- add appflow policy appflow_pol_cluster true appflow_act_cluster
- bind appflow global appflow_pol_cluster 1 END -type ICA_REQ_DEFAULT

By default, this number is 0, indicating that the node does not belong to any cluster or HA-pair. To form an observation domain of more than one NetScalers, this ID must be more than 1000. The default name of the observation domain is “Default”.

Section 4: Configuration: NetScaler SDX Cluster

SDX Build

Configure NetScaler SVMs

Install the SVM bundle.

Configure NetScaler Instances

Provision the NetScaler instances from the SVM with NSIP and default gateway on the cluster backplane VLAN.

Cluster Build

Configure Cluster Instance

Create a Cluster instance from the SVM on appliance (1)
Assign the unique Cluster IP (CLIP) Address

Configure Additional Cluster Nodes

Join Cluster node members from each SVM on which they are hosted into the Cluster Group (Default _Group).

Configure the Cluster Node Priorities

The default SVM assigned priority for a cluster node is 31. This is set on all nodes and can affect the cluster coordinator (CCO) election process when a node member leaves or joins the cluster. Change
the priority for each node to give first, second and third priority for CCO elections. Setting the Cluster Priority ID allows for management of the Default CCO process.

**Route Management**

**Policy Based Routes**

Use Policy Based Routes to manage CLIP and SNIP communication to the SVM instance. This may require more than one PBR rule depending on the next hop definition and how many nodes are in the cluster.

**Adjust Default Route**

Change the Default Route to forward all VServer traffic to the production network, from the default route that was assigned by the SVM at time of provisioning.

**Create Spotted SNIP(s)**

From the Cluster IP, configure a spotted SNIP for each member node of the cluster with an IP address from the Traffic VLAN. These IP addresses all need to be on the same network segment. Enable Dynamic Routing. A Spotted SNIP is required for each cluster node that will participate in ECMP.

**Configure Virtual Server**

Two steps required:

- Create LBVserver for HTTP service
  - Enable RHI State on the LBVserver

- Configure the Striped VIP to the corresponding Vserver
  - Set the LSA value = 1
  - Set the OSPF AREA value (xx)
  - Set Host Route Gateway (Striped SNIP)

**Configure Dynamic Routing**

The dynamic routing configuration (OSPF in this example) for the SDX cluster should be completed on the Cluster Coordinator (CCO) via the Cluster IP address (CLIP).

**OSPF Configuration**

Defining the "-ownerNode" command for each member node of the cluster will allow the CCO to distribute the OSPF routes to all members of the cluster.
Configuration Example: NetScaler SDX Cluster

Cluster Diagram

Configuration Example

Created Instances from SVM

```bash
set ns config -IPAddress 192.168.7.11 -netmask 255.255.255.0
add vlan 301 -sdxVlan YES
set interface 10/1 -lacpMode ACTIVE -lacpKey 2 -ifnum LA/2
set interface 10/2 -lacpMode ACTIVE -lacpKey 2 -ifnum LA/2
set interface 10/3 -lacpMode ACTIVE -lacpKey 3 -ifnum LA/3
set channel LA/2 -tagall ON
add route 0.0.0.0 0.0.0.0 192.168.7.1 -routeType STATIC

set ns config -IPAddress 192.168.7.12 -netmask 255.255.255.0
add vlan 301 -sdxVlan YES
set interface 10/1 -lacpMode ACTIVE -lacpKey 2 -ifnum LA/2
set interface 10/2 -lacpMode ACTIVE -lacpKey 2 -ifnum LA/2
set interface 10/3 -lacpMode ACTIVE -lacpKey 3 -ifnum LA/3
set channel LA/2 -tagall ON
set channel 0/LA/3
add route 0.0.0.0 0.0.0.0 192.168.7.1 -routeType STATIC
```
Validated Reference Design Guide for NetScaler SDX Clustering

Initial Created Cluster Config

- add cluster instance 1
- add cluster node 0 192.168.7.11 -state ACTIVE -backplane 0/LA/3
- add cluster node 1 192.168.7.12 -state ACTIVE -backplane 1/LA/3
- add cluster node 2 192.168.7.13 -state ACTIVE -backplane 2/LA/3
- bind cluster nodegroup DEFAULT_NG -node 0
- bind cluster nodegroup DEFAULT_NG -node 1
- bind cluster nodegroup DEFAULT_NG -node 2
- enable cluster instance 1

- add ns ip 192.168.7.14 255.255.255.255 -type CLIP -vServer DISABLED -mgmtAccess ENABLED
- add route 0.0.0.0 0.0.0.0 192.168.7.1 -routeType STATIC

Add traffic VLAN to cluster

- add vlan 401

Add spotted SNIPs:

- add ns ip 192.168.6.11 255.255.255.0 -vServer DISABLED -gui DISABLED -dynamicRouting ENABLED -ownerNode 0
- add ns ip 192.168.6.12 255.255.255.0 -vServer DISABLED -gui DISABLED -dynamicRouting ENABLED -ownerNode 1
- add ns ip 192.168.6.13 255.255.255.0 -vServer DISABLED -gui DISABLED -dynamicRouting ENABLED -ownerNode 2

Add striped SNIP:

- add ns ip 192.168.6.14 255.255.255.0 -vServer DISABLED -gui DISABLED

BIND traffic VLAN to cluster interfaces and striped SNIP address

- bind vlan 401 -ifnum 0/LA/2 -tagged
- bind vlan 401 -ifnum 1/LA/2 -tagged
- bind vlan 401 -ifnum 2/LA/2 -tagged
- bind vlan 401 -IPAddress 192.168.6.14 255.255.255.0

Add PBR for cluster backplane traffic

- add ns pbr Traffic_PBR ALLOW -srcIP "!=" 192.168.7.11-192.168.7.14 -nextHop 192.168.7.1 -priority 10
- apply ns pbtrs

Change default route to use traffic VLAN

- add route 0.0.0.0 0.0.0.0 192.168.6.1 -routeType STATIC
- rm route 0.0.0.0 0.0.0.0 192.168.7.1

Enable Dynamic Routing

- enable ns feature OSPF

Add Server:

- add server HTTP-Test 127.0.0.1
Create Service or Service Group:
add service HTTP-Test-Svc HTTP-Test HTTP 80 -gslb NONE -maxClient 0 -maxReq 0 -cip DISABLED -usip NO -useproxyport YES -sp ON -cltTimeout 180 -svrTimeout 360 -CKA NO -TCPB NO -CMP NO

Create LB vServer
add lb vserver HTTP-Test-LBVS HTTP 192.168.5.1 80 -persistenceType NONE -cltTimeout 180 -RHIstate ACTIVE

Bind Service to LB vServer
bind lb vserver HTTP-Test-LBVS HTTP-Test-Svc

Enable Dynamic Routing on VIP for LB vServer
add ns ip 192.168.5.1 255.255.255.255 -type VIP -snmp DISABLED -hostRoute ENABLED -hostRtGw 192.168.6.14 -ospfLSAType TYPE1 -ospfArea 51

Configure OSPF in VTYSH
config t
!
log syslog
!
log record-priority
!
interface lo0
!
interface vlan0
!
interface vlan107
ip ospf priority 0
!

router ospf 1
owner-node 0
ospf router-id 192.168.6.11
exit-owner-node
owner-node 1
ospf router-id 192.168.6.12
exit-owner-node
owner-node 2
ospf router-id 192.168.6.13
exit-owner-node
auto-cost reference-bandwidth 100000
redistribute kernel
passive-interface vlan0
area 51 stub
network 192.168.6.0/24 area 51
!
exit
!
write mem
!
Appendix A: Cluster Troubleshooting

NetScaler provides several diagnostic tools to view console messages, event messages and download traces. Most of these utilities are available directly from the NetScaler GUI. This section focuses on external tools.

Nsclusterd: service daemon

Heartbeats:

NSIP: port 7000 UDP
Sent every hello interval (default 200ms), dead interval is 3 secs
ARP resolution occurs on all UP interfaces, including the backplane interfaces.
If ARP is resolved, the cluster ARP cache is updated.
Heartbeats are sent to peer via interfaces where ARP resolves.
L3 cluster use the routes for heartbeats, L2 cluster floods the enabled interfaces.

Cluster Communication:

Uses RPC (3011) or secure RPC (3009)
Cluster Communication happens via backplane. If the backplane is down, any other interface where ARP resolves is chosen.

CLI Commands:

add cluster instance <clid> -inc <ENABLED|DISABLED> -processLocal <ENABLED|DISABLED>
add cluster node <nodeid> <nodeip> -nodegroup <ng>
add ns pbr [-ownerGroup <string>]
add ns pbr6 [-ownerGroup <string>]
Show techsupport –scope cluster

Counters:
nsconmsg –d logfromnfw (Logs)
nconmsg –g clusterd_ –g cl_ -d current (Counters)

Heartbeats:

Cluster ARP cache (nsapimgr –ys call=nscl_ArpCacheDump)
Trace. (UDP port 7000)

NNM/config sync/propagation:

Node to Node Messaging for packet engine communication
• RPC NSIP:3011 or secure RPC 3009

RPC node mismatch between nodes (show rpcnode)
Trace. (RPC port 3011 (non-secure) or 3009 (secure))
nsapimgr –ys trace_nnm=1
nsapimgr –d nnm_messages
nsconmsg –g nnm –d current
Cluster Counters:

- `clusterd_tot_cu_enqueued`
  - Command has been received by clusterd (incremented on CCO)
- `clusterd_tot_cu_delivered`
  - Command is delivered to configd by clusterd (incremented on all nodes where command is delivered)
- `clusterd_tot_cuTimeout`
  - Propagation timeout (incremented on CCO)

VTYSH:

Ns#show ip route kernel
- Shows all vips to be advertised

Ns#show ip route
- Shows best routes in RIB

Ns#sh ip interface brief
- Shows all interfaces and their primary ip address along with up/down status

Ns#show ns sync-status
- Shows last sync status
  - `ns(config)#debug ini netscaler`
  - Enable logging of cluster events, config commands propagation and synchronization events: All logs will go to /var/log/ns.log file

ns(config)#debug nsm kernel
- Logs all cluster state changes received by nsm module, route updates sent to or received from kernel

Tools

Wireshark: Wireshark (Version 1.10.3 and above) is a free packet sniffer and capture tool; that allows for network troubleshooting and analysis in a networked environment. Wireshark provides a very similar approach to tcpdump by using a GUI front-end to view network traces. Wireshark also provides sophisticated filtering capabilities; allowing network administrators to filter on response codes, HTTP methods and payload. Wireshark can be downloaded from: WireShark

NetScaler Plugin for Wireshark: http://engwiki.eng.citrite.net/mynswiki/NsWireshark

Putty: Putty is a free tool that allows Windows users to connect to remote systems over the Internet via Telnet and SSH. While both Telnet and SSH allow you to connect to remote systems, SSH, supported in Putty, provides for a “Secure Shell”, encrypting information before it is transferred. Putty is the recommended tool for connecting via SSH to the NetScaler. Putty can be downloaded from: Putty

WinSCP: WinSCP is an open source free SFTP client and FTP client for Windows. WinSCP provide secure transfer of files between the NetScaler and a local host. WinSCP provides a graphical GUI to the traditional SCP (secure copy) functionality offered by *NIX operating systems. WinSCP can be downloaded from: WinSCP

PSCP: PSCP is the command line alternative to WinSCP that is identical to the “scp” functionality offered in traditional *NIX operating systems. PSCP is a lightweight DOS based application that can be used to copy files between the NetScaler and the local host. PSCP can be downloaded from: PSCP
Command Line Diagnostics

nstrace.sh: nstrace.sh is a NetScaler utility (script file) that allows network administrators to take NetScaler traces from the appliance. The NetScaler trace captures all traffic going through the box at any given time. The syntax for nstrace.sh is /netscaler/nstrace.sh -sZ0 -tcpdump 1.

This syntax will automatically create a trace file in the /var/nstrace directory. The administrator can press Ctrl-C on his keyboard to stop the trace. This trace can then be downloaded to a local host (via PSCP, WinSCP) and viewed on any packet capture program (i.e. Wireshark).

nstcpdump.sh: The /netscaler/nstcpdump.sh script is a utility that emulates tcpdump syntax on NetScaler interfaces. The main benefit using nstcpdump.sh includes its filtering ability. Using an example below:

```
/netscaler/nstcpdump.sh -w /var/nstrace/ftp.pcap host 192.168.1.1 and host 192.168.1.2 and tcp.port==21
```

The above filter allows for all FTP traffic between 192.168.1.1 and 192.168.1.2 to be captured on the NetScaler and downloaded to /var/nstrace/ftp.pcap. This file can be downloaded to a local machine and viewed on Wireshark for easier analysis.

Local Syslog: NetScaler stores all log files locally on the appliance under /var/log/ns.log. Details regarding authentication errors or NetScaler can be viewed by running the following command:

```shell
# cat /var/log/ns.log
tail -f /var/log/ns.log
```

Authentication Debugging Tools: NetScaler provides a debugging utility to check for authentication successes and failures. Group extraction can also be validated using this utility. User details including group extraction can be viewed on screen by typing the following command, while a user is logging on to the VPN:

```shell
# cat /tmp/aaad.debug
```

The output of the command will provide helpful details on the authentication scheme used, success, failures and cause of failures (i.e. incorrect password, bad bindings etc.).

Console Message Diagnostics: The NetScaler provides useful console messages that can shed light on NetScaler performance. For example, using the following command below can easily identify IP address conflicts and duplex mismatches:

```shell
> shell
# nsconmsg -K newnslog -d consmsg (live)
or
# nsconmsg -K newnslog -K /var/nslog/newnslog -d consmsg (from latest file)
```

Event Message Diagnostics: The NetScaler provides useful event messages that can shed light on status of configured NetScaler services and high availability. For example, the status of a specific configured service or notification of a failover or a reboot can be identified by running the command below:

```shell
> shell
# nsconmsg -K newnslog -d event (live)
or
# nsconmsg -K newnslog -K /var/nslog/newnslog -d event (from latest file)
```
Appendix B: SVM Management

The Service Virtual Machine (SVM) and subsequent Management Service should always be used to perform management operations on NetScaler VPX instances. Once a NetScaler VPX instance is provisioned, the Management Service creates policies, instance administration profiles, and other configurations on the VPX instance (such as CPU allocation, SSL chips, memory, etc.).

After the conclusion of this engagement, it was found that the NetScaler VPX instance seemed to be named different from what was defined on the SVM. A memory allocation change resulted in the SVM pushing down the name that was defined on the SVM and a reboot (as expected to bring the parameters into effect). This caused a mismatch with the name as defined in the license file and thus prevented user connections (Citrix Technical Support Case # 71253542).

While it is possible to access many of the same properties directly on the NetScaler VPX instance, it is strongly recommended to only utilize the SVM for conducting such changes and configuration.

Appendix C: OSPF Command Reference

https://docs.citrix.com/content/dam/docs/en-us/netscaler/media/dynamic-routing-crgs/Citrix-ZebOSOSPFC-mdRef.pdf
About Citrix

Citrix (NASDAQ:CTXS) is a leader in mobile workspaces, providing virtualization, mobility management, networking and cloud services to enable new ways to work better. Citrix solutions power business mobility through secure, personal workspaces that provide people with instant access to apps, desktops, data and communications on any device, over any network and cloud. This year Citrix is celebrating 25 years of innovation, making IT simpler and people more productive. With annual revenue in 2013 of $2.9 billion, Citrix solutions are in use at more than 330,000 organizations and by over 100 million users globally. Learn more at www.citrix.com.

Copyright © 2014 Citrix Systems, Inc. All rights reserved. Citrix, NetScaler MPX, NetScaler SDX, NetScaler, CloudBridge and AppFlow are trademarks of Citrix Systems, Inc. and/or one of its subsidiaries, and may be registered in the U.S. and other countries. Other product and company names mentioned herein may be trademarks of their respective companies.