

Ericsson Unicycle OVS-DPDK Validation HW, Networking and IP plan

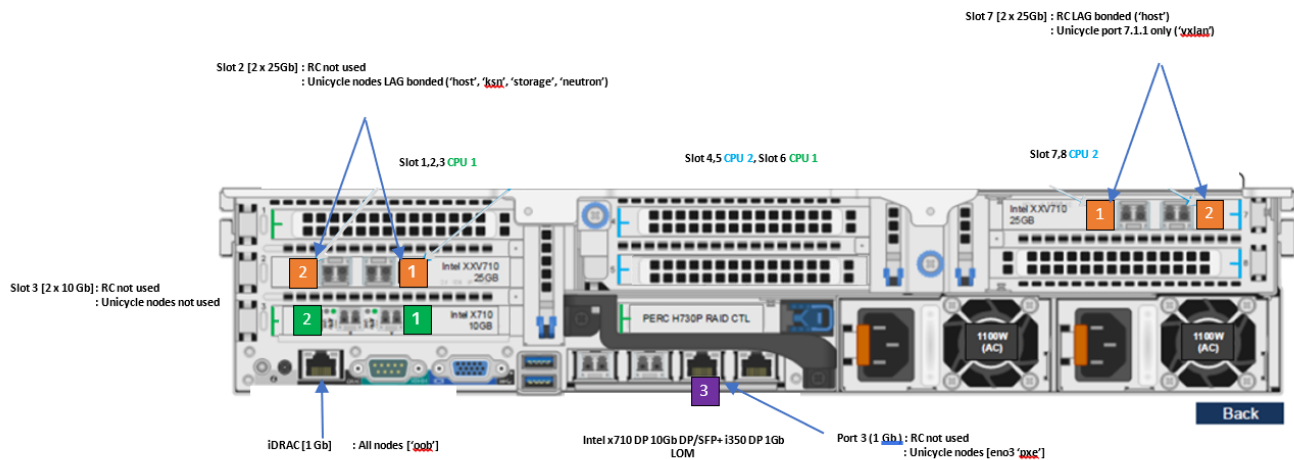
- [Unicycle OVS-DPDK Validation Servers](#)
- [Unicycle OVS-DPDK Validation Networking](#)
- [Unicycle OVS-DPDK Validation IP/VLAN Plan](#)
- [Unicycle OVS-DPDK BGP Plan](#)
- [Unicycle OVS-DPDK LAG Details](#)

Unicycle OVS-DPDK Validation Servers

Verification was based on a Build Server VM and four identical Dell 740XD servers for the Regional Controller and three Unicycle nodes.

Dell Purley 740XD Server

L2 networks 'oob', 'host', 'pxe', 'ksn', 'storage', 'neutron', 'vxlan' – Refer to IP plan spreadsheet for values and further details



PowerEdge R740XD Server



Components

- 1 PowerEdge R740/R740XD Motherboard
- 1 Intel Xeon Gold 6152 2.1G, 22C/44T, 10.4GT/s , 30M Cache, Turbo, HT (140W) DDR4-2666
- 1 iDRAC Group Manager, Enabled
- 1 iDRAC,Legacy Password
- 1 Chassis with Up to 24 x 2.5 Hard Drives for 2CPU, GPU Capable Configuration
- 1 Riser Config 6, 5 x8, 3 x16 slots
- 1 PowerEdge R740 Shipping Material
- 1 No Quick Sync
- 1 Performance Optimized
- 1 2666MT/s RDIMMs
- 12 32GB RDIMM 2666MT/s Dual Rank
- 1 Intel Xeon Gold 6152 2.1G, 22C/44T, 10.4GT/s , 30M Cache, Turbo, HT (140W) DDR4-2666
- 1 iDRAC9,Enterprise
- 4 480GB SSD SATA Read Intensive 6Gbps 512 2.5in Hot-plug AG Drive, 1 DWPD, 876 TBW
- 6 2.4TB 10K RPM SAS 12Gbps 512e 2.5in Hot-plug Hard Drive
- 1 PERC H730P RAID Controller, 2GB NV Cache, Adapter, Low Profile
- 2 C13 to C14, PDU Style, 10 AMP, 6.5 Feet (2m), Power Cord
- 1 Dual, Hot-plug, Redundant Power Supply (1+1), 1100W
- 1 No Trusted Platform Module
- 1 Order Configuration Shipbox Label (Ship Date, Model, Processor Speed, HDD Size, RAM)
- 1 GPU Ready Configuration Cable Install Kit
- 1 PE R740XD Luggage Tag
- 1 Intel X710 Dual Port 10Gb Direct Attach, SFP+, Converged Network Adapter
- 2 Intel XXV710 Dual Port 25GbE SFP28 PCIe Adapter, Full Height
- 1 Intel X710 DP 10Gb DA/SFP+, + I350 DP 1Gb Ethernet, Network Daughter Card
- 1 HS Install Kit,GPU Config.No cable
- 1 ReadyRails Sliding Rails Without Cable Management Arm
- 1 Unconfigured RAID
- 1 OME Server Configuration Management

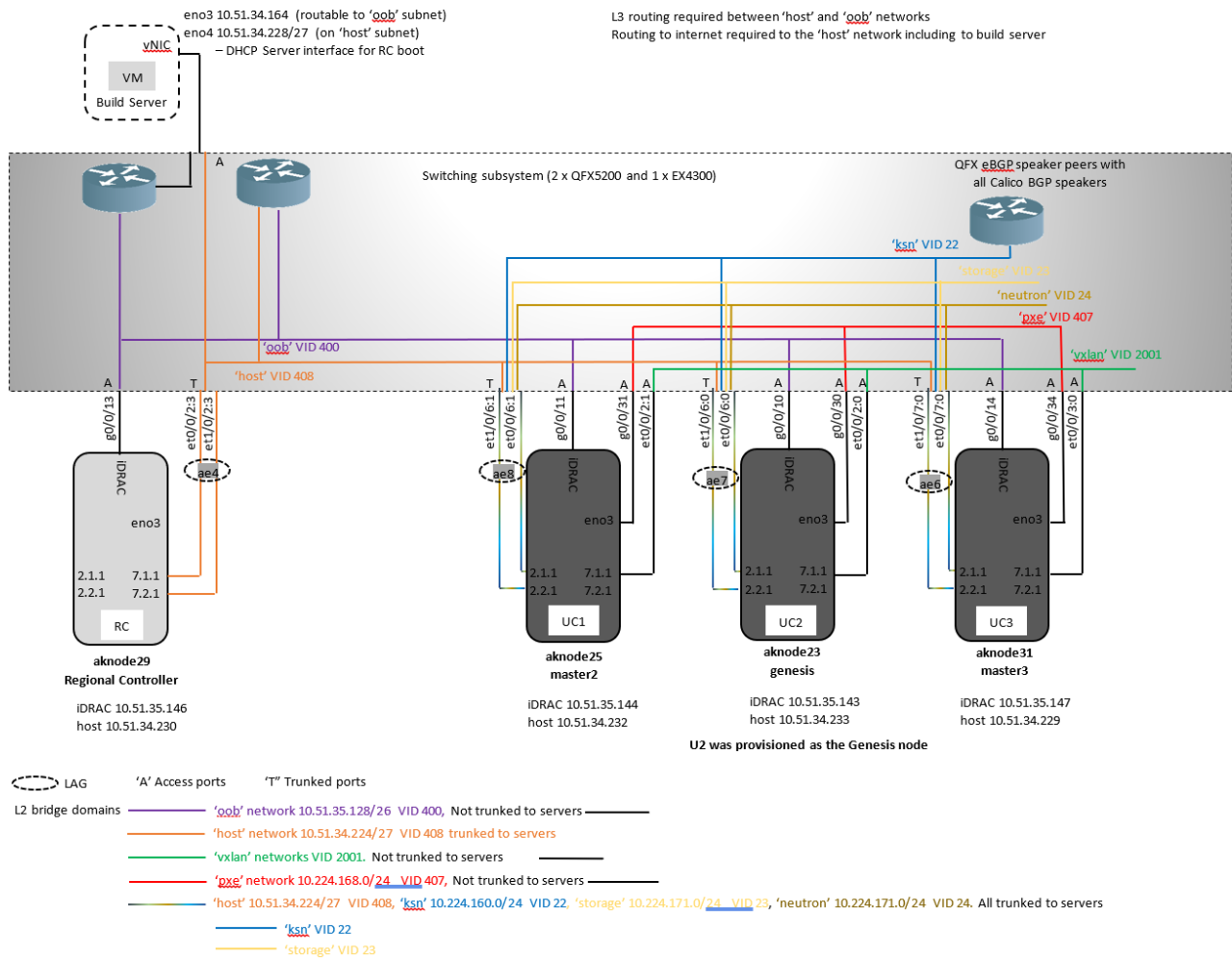
Software

- 1 6 Performance Fans forR740/740XD
- 1 Performance BIOS Settings
- 1 No Operating System
- 1 No Systems Documentation, No OpenManage DVD Kit

Unicycle OVS-DPDK Validation Networking

The diagram below shows an example physical and L2 connectivity, the IP subnet plan, host and iDRAC addressing scheme similar to that used in the validation.

The Unicycle deployment does not configure the networking subsystem thus the choice of switch subsystem components is not restricted to those shown and they can be replaced with devices offering equivalent functionality.



Unicycle OVS-DPDK Validation IP/VLAN Plan

Below is an example of a detailed network and IP address plan similar to that used during validation.

NOTE: LAG groups always created on 2 ports of the SASE NIC card on compute. LAG not split across NIC card slots. 2 x SASE NIC card and host card											
NOTE: Build Server is only used to create the Regional Controller. Once the RC is built the build server has no further role in the deployment of the Row or Unicycle pods and can be removed. Note: The RC uses the LAG bonded interface on 1, 1 and 2, 1 to send traffic to the Row and Unicycle nodes connected to the host network that routing between the 'host' and 'oob' networks is required.											
Network name	Server	Network name	Compute HW	Present	NIC type	Port rate	LAG	VID	Subnet	Host address	Network name
'oob'	Build Server	oob	eno3	Yes	vmc	Any	No	native	Any	10.51.34.164	
	Regional controller	oob	eno3	Yes	BMC	10	No	native	Any	10.51.35.146	
	Unicycle [master2]	oob	eno3	Yes	BMC	10	No	native	10.51.35.128/27	10.51.35.144	'oob'
	Unicycle [master3]	oob	eno3	Yes	BMC	10	No	native	10.51.35.144	10.51.35.144	
'host' (single seen as 'host')	Build Server	host	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	host	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	host	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	host	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'vxlan' (single seen as 'vxlan')	Build Server	vxlan	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	vxlan	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'pxe' (single seen as 'pxe')	Build Server	pxe	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	pxe	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	pxe	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	pxe	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'storage' (single seen as 'storage')	Build Server	storage	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	storage	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	storage	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	storage	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'neutron' (single seen as 'neutron')	Build Server	neutron	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	neutron	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	neutron	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	neutron	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'vxlan' (single seen as 'vxlan')	Build Server	vxlan	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	vxlan	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'kns' (single seen as 'kns')	Build Server	kns	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	kns	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	kns	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	kns	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'storage' (single seen as 'storage')	Build Server	storage	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	storage	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	storage	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	storage	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'neutron' (single seen as 'neutron')	Build Server	neutron	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	neutron	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	neutron	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	neutron	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
'vxlan' (single seen as 'vxlan')	Build Server	vxlan	eno3	Yes	vmc	Any	No	native	Any	10.51.34.228	
	Regional controller	vxlan	eno3	Yes	BMC	10	No	native	Any	10.51.34.228	
	Unicycle [master2]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	
	Unicycle [master3]	vxlan	eno3	Yes	BMC	10	No	native	10.51.34.224/27	10.51.34.232	

Unicycle OVS-DPDK BGP Plan

The three unicycle nodes automatically peer with an external fabric BGP speaker using eBGP. The calico nodes do not peer using an internal iBGP mesh. Below is an example similar to that used in the validation.

BGP	ASN	IP address/subnet	Start IP	End IP		
Calico ('k8s' network)	65511	10.224.160.0/24	10.224.160.134	10.224.160.254		
external router (maybe QFX or other HW/SW router TBD)	65001	10.224.160.129	NA	NA		Any private ASN other than the Calico ASN can be used. The external router must be on the same L2 domain as the 'k8s' network but can be a physical router or a SW based router Dynamic BGP peering to the Calico subnet must be configured on the external peer router to accept BGP peering requests from any k8s Calico BGP speaker The external BGP router must be setup to advertise routes to BGP speakers of the same ASN that they are received from.

In addition to the eBGP approach, limited validation was also performed using a full iBGP mesh between calico nodes without the need for an external BGP router. To change the deployment to BGP peer in this manner follow the notes shown in the example yaml file [Example Configuration Input File - Unicycle Pods with OVS-DPDK Dataplane on Dell 740XD Servers](#).

Unicycle OVS-DPDK LAG Details

The RC and Unicycle genesis nodes boot via VLAN tagged 'host' interface which is pre-provisioned on the QFX switches with LAG bonding. Since booting occurs before the linux kernel can bring up its LAC-P signalling the QFX switches must be configured to pass traffic on their primary (first) link before the LAG bundle is up.

Note the Unicycle [master2] and unicycle [master3] nodes do not boot over the 'host' network but rather over the edge site 'pxe' network which is a single link to each unicycle server and thus not lag bonded.

LAG						
LAG config on QFX5200	FORCE UP	DHCP and http pxe boot occurs over interfaces configured for LAG on JPR switches				