

DIRECT CONNECT EXADATA-X8 TO PCA

Procedure to create a direct network connection between Oracle Private Cloud Appliance and Exadata X8M/X8/X7.

January, 2021
Copyright © 2021, Oracle and/or its affiliates
Public

PURPOSE STATEMENT

This document provides an overview of features and enhancements included in release PCADirectConnect 1.0.1-7 release. It is intended solely to help you assess the business benefits and procedure of using PCADirectConnect functionality for projects involving Exadata-X8/X7 and PCA/PCC 2.4.x.

DISCLAIMER

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle.

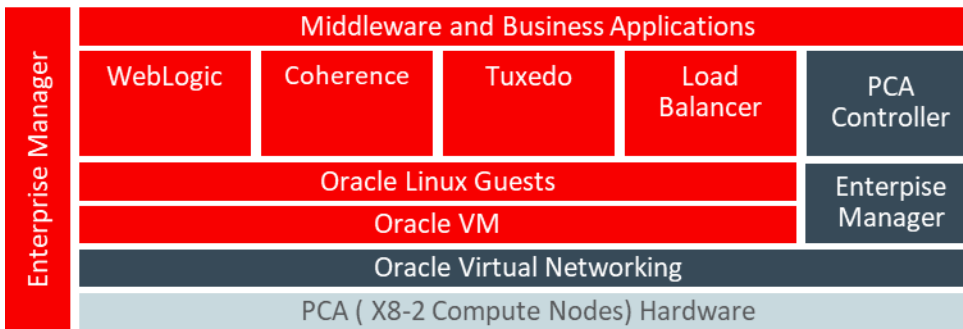
Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

TABLE OF CONTENTS

Purpose Statement	2
Disclaimer	2
Overview	4
Logical/Physical Network connectivity	5
Configuring the Oracle Private Compute Appliance	8
Step-1: Verify ports 7 and 8 are not in use and install the PCADirectConnect scripts	8
• Step-2: Creating Custom Configuration	9
A. Run the pca-create-uplink script to setup any/all of ports 7-10 to connect to Exadata dbnodes. Verify using pca-list-uplink script.	10
B. Run the pca-create-network script to setup the direct_connect network and assign to it ALL the ports created using pca-create-uplink script. Verify using pca-list-network script.	11
C. Run the pca-add-network script to add the direct_connect network to desired Compute Nodes or Tenent Groups. The network should be added ONE-AT-A-TIME to each Compute Node	12
D. (Optional) Update (add/delete ports) the direct_connect network AFTER it has been added to ComputeNodes	13
Step-3: Verify the direct_network is seen by OVMM	15
Configuring the Oracle Exadata Database Machine	16
Frequently Asked Questions	18

OVERVIEW

The Oracle Private Cloud Appliance provides a simple out-of-the-box converged infrastructure solution and a key enabler for Oracle-As-A-Service platform.



Benefits:

- Intelligent and agile infrastructure
- Flexibly supports any mix of Oracle and non-Oracle workloads
- Supports your choice of Oracle or 3rd party storage
- Automation speeds install
- OVM virtual appliances and migration tool for rapid Oracle app deployment
- Built for rapid and cost-effective private cloud deployment
- Zero Downtime Rolling Upgrades
- Software Define Networking – over 100Gb/s Ethernet backplane
- Migrate legacy Exalogic workloads to PCA
- Live migration
- No infrastructure license cost

The 100Gb Ethernet network inside Oracle Private Cloud Appliance, is implemented as spine and leaf topology. Each Compute Node has a connection to both leaf switches. Each leaf, is in turn, connected to both the spine switches. Each spine switch has connectivity to the Storage Nodes and the Management nodes, as well as a group of ports for external connectivity. Ports 1-4 on both the spine switches are reserved for custom networking requirements specified by the customer, while port 5 is for the default uplink connection.

Note: When referring to a port, such as 'Port 5', we are actually referring to both port 5 on the first spine switch located at Rack Unit 22 and the second spine switch, located at Rack Unit 23.

Each of the four customer reserved ports, numbered 1 through 4, may be configured in a number of ways. Each port may be broken down as:

- A single 100Gb Ethernet port
- A single 40Gb Ethernet port
- Four 25Gb Ethernet sub ports
- Four 10Gb Ethernet sub ports

The Oracle Exadata X8M product line introduces a new and improved networking infrastructure built on high-speed, low latency 100Gb Ethernet optimized for RDMA over Converged Ethernet (RoCE). Connectivity to Applications and Middleware tier is provided via physical Ethernet interfaces on the database-servers. Each Oracle Exadata Database Machine X8M-2 database server consists of the following network components and interfaces:

- 1 embedded 1 GbE (eth0 RJ-45)
- 2 embedded 10GBaseT ports (RJ-45) or 2 embedded 10GbE/25GbE ports (SFP+/SFP28)
- 1 dual-port QSFP28 100Gb/s RDMA Network Fabric Card
- 1 embedded Ethernet port for Integrated Lights Out Manager (ILOM) remote management
- 1 dual-port 25GbE PCIe 3.0 network card with Broadcom BCM57414 10Gb/25Gb Ethernet Controller

Note: The SFP28 modules for the 25GbE PCIe 3.0 network cards are purchased separately.

This document focuses on network attachment and configuration steps needed to create secure, private and high-bandwidth connectivity between the VMs/Containers, comprising the apps/middleware tier running on Oracle Private Cloud Appliance, and the database tier running on the Oracle Exadata X8M Database Machine.

Directly connecting the database-servers in the Exadata-X8M rack to the spine switches in the PCA rack is ideal for migrating Oracle Exalogic workloads to PCA. This configuration will eliminate the need to acquire and manage dedicated top of rack switches. This option will also provide superior ROI, tighter integration, lowest latency for database transactions and provisions the maximum available bandwidth.

Important Considerations:

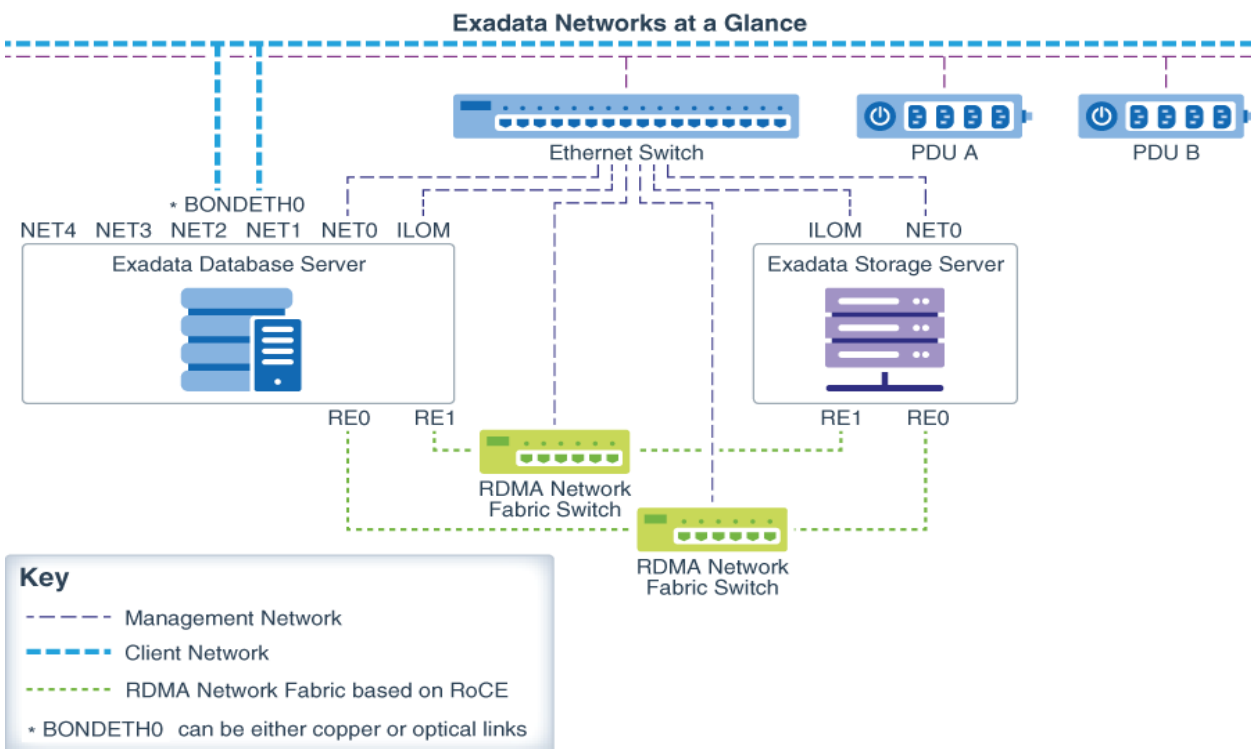
- Spine switches inside PCA rack have vPC configuration to make the 2 spines appear as a single-virtual-switch
- LACP **must** be enabled for the Client-Network interfaces to enable maximum bandwidth and failover
- The connection between PCA-spine-switch and database-servers is setup as a private network
- Access from the customer’s network should be enabled via other available interfaces on the database-server
- DefaultGateway/DNS/NTP servers must be reachable via some other interface on the database-servers
- The subnet-mask for Client/Other-network should be chosen to accommodate the ip addresses for the VMs in the Oracle Private Compute Appliance.
 - /22 mask provides 1024 addresses, /23 mask provides 512 addresses

This paper addresses the following topics:

- Logical/Physical network connectivity
- Network configuration on Oracle Private Compute Appliance
- Network configuration on Oracle Exadata X8M database-servers

LOGICAL/PHYSICAL NETWORK CONNECTIVITY

Figure 1. Logical Network diagram of Exadata X8M networks



The Client-Network is used by the applications to access the database. The database-servers provide 2-pairs of physical interfaces configured by OEDA as bondeth0 and bondeth1, one of which can be used to create a direct connection to the spine switches in the Oracle Private Cloud Appliance.

Figure 2. Available physical interfaces on the database-servers inside Exadata X8M

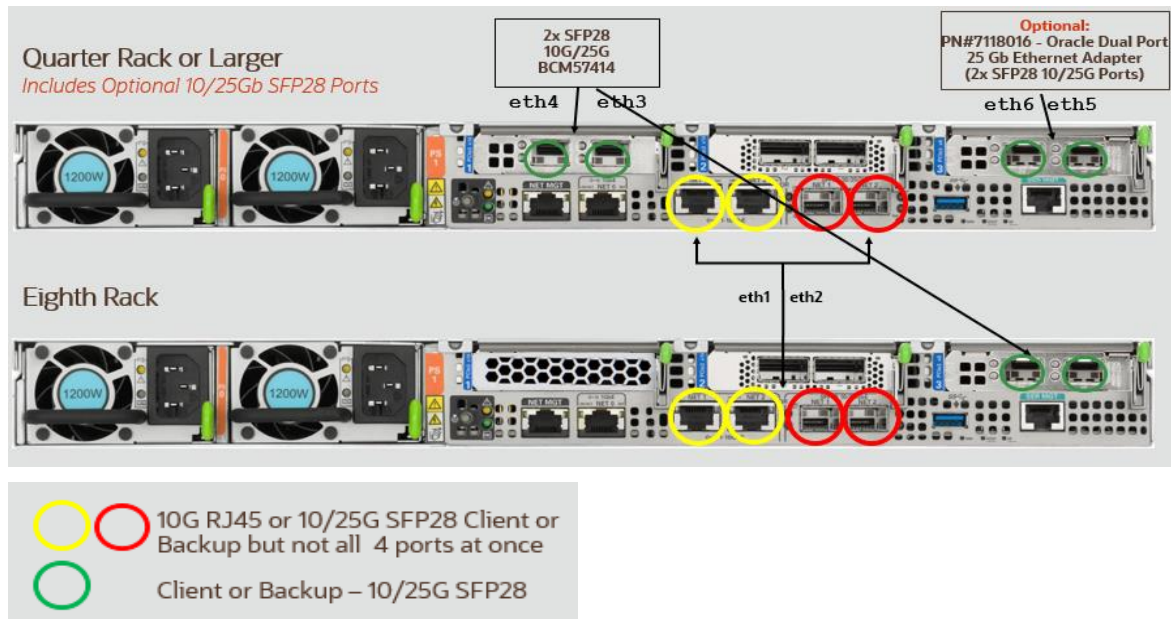
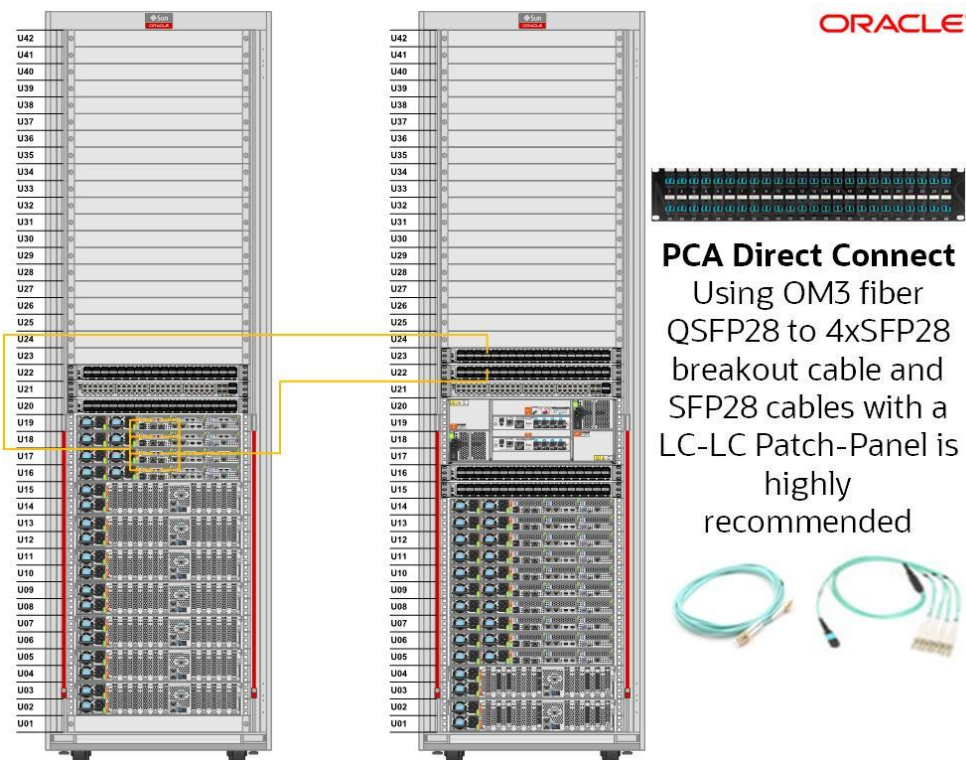


Figure 3. Physical cable connections between the PCA spine switches and database-servers

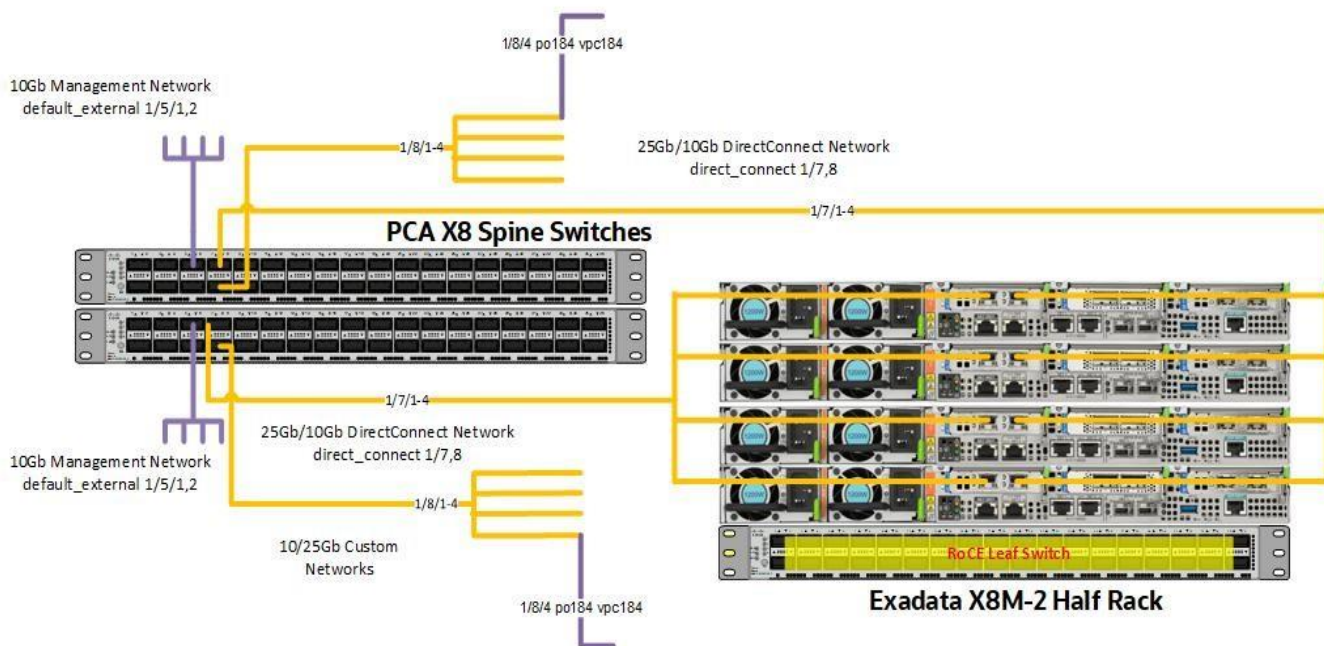


ORACLE

ORACLE

PCA Direct Connect

Sample cable connections for Eighth, Quarter and Half Exadata-X7/X8 racks



CONFIGURING THE ORACLE PRIVATE COMPUTE APPLIANCE

The Client-network on Exadata X8-2 will always use 25Gb Ethernet cards in the database-servers. Hence, the PCA spine-switch ports 7 (and 8, if connecting more than 4 database-servers) must be configured as 4x25Gb.

Connecting the Exadata Client-network directly to the PCA spine-switches has the benefits of isolating the Client-network from other workloads and providing a greater level of control. Dedicated 50Gbps of bandwidth capacity to each database-server for large workloads and the ability to use jumbo frames, enables higher application efficiency.

Figure 3 depicts the network cabling diagram for a full-rack Exadata-X8M with 8 database-servers. Smaller configurations, like eighth, quarter and half rack will only require connecting port 7 from both switches. The configuration on PCA is done in multiple steps which are done using the script/rpm attached to this MOS note.

- The ports must be first reserved to ensure they can be used exclusively for direct Exadata connection.
- Configuration on each spine-switch has to be updated using the switch CLI
- QSFP28 to SFP28 breakout cables are then used to make the physical connections. Please refer to Table-2 for the cable and transceiver part-numbers.

Step-1: Verify ports 7 and 8 are not in use and install the PCADirectConnect scripts

```
PCA> list uplink-port-group
```

Port_Group_Name	Ports	Mode	Speed	Breakout_Mode	Enabled	State
default_5_1	5:1 5:2	LAG	10G	10G-4X	TRUE	(UP) * NOT ALL PORTS ARE UP
default_5_2	5:3 5:4	LAG	10G	10G-4X	FALSE	DOWN

- **Download and install the latest version of scripts rpm – Requires Oracle VPN access**

The latest version was 1.0.1-7 at the release of this document

https://systemsweb.us.oracle.com/vse/pca-direct-connect/pca-direct-connect-1.0.1-7.x86_64.rpm

The rpm needs to be copied and installed on the management node as the root user. The rpm includes several scripts which provide a modular way to configure the PCA switch ports.

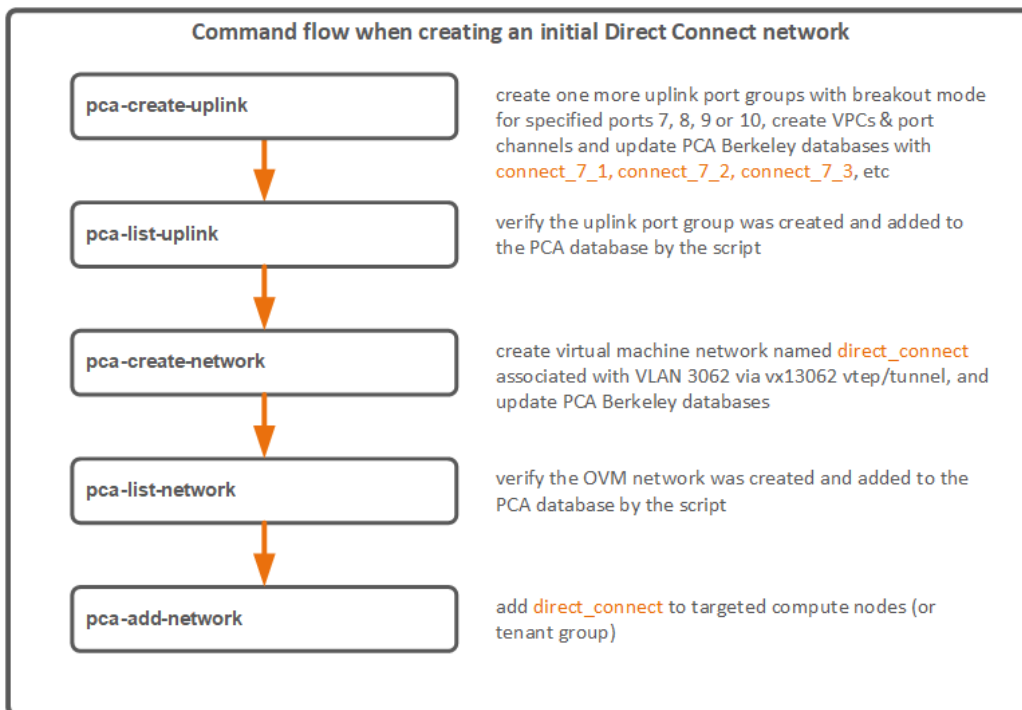
```
root@ovcamn06r1-bx8pca# rpm -qlp ./pca-direct-connect-1.0.1-7.x86_64.rpm
/usr/sbin
/usr/sbin/pca-add-network
/usr/sbin/pca-create-network
/usr/sbin/pca-create-uplink
/usr/sbin/pca-delete-network
/usr/sbin/pca-delete-uplink
/usr/sbin/pca-list-network
/usr/sbin/pca-list-uplink
/usr/sbin/pca-remove-network
/usr/sbin/pca-update-network

root@ovcamn06r1-bx8pca# rpm -ivh ./pca-direct-connect-1.0.1-7.x86_64.rpm
preparing... ##### [100%]
1:pca-direct-connect ##### [100%]

root@ovcamn06r1-bx8pca# rpm -qa | grep direct
pca-direct-connect-1.0.1-7.x86_64
```


- **Step-2: Creating Custom Configuration**

- Parameters can be passed to the script to select a subset of ports or to configure the ports as 10Gb to support older Exadata models
- The script will do the following changes:
 - Each port-pair (one from each spine-switch) will be in its own port-channel/vpc
 - All port-channels connecting to the dbnodes will have the same vlan-id
 - FEC will be set to off
 - Configuration changes will be done on both spine-switches
 - Create the direct_connect network with selected ports
 - Add the direct_connect network to the Compute Nodes so it can be used by the VMs



ORACLE VM Manager

Health Servers and VMs Repositories **Networking** Storage Reports

Networks VLAN Interfaces Virtual NICs

Name	ID	Network Channels				
		Server Management	Cluster Heartbeat	Live Migrate	Storage	Virtual Machine
192.168.32.0	c0a82000	√		√		
192.168.40.0	10ee950564		√		√	
default_external	1037661503					√
default_internal	10f57bf8da					√
direct_connect	107fb3bf93					√
external_network	10766aeb16					√
etc_nw1	10235a4a52					√

- A.** Run the `pca-create-uplink` script to setup any/all of ports 7-10 to connect to Exadata dbnodes.
Verify using `pca-list-uplink` script.

```
# pca-create-uplink -h
usage: pca-create-uplink [-h] [-p PORTS] [-m BREAKOUT_MODE] [-f FEC_MODE] [-v]

Direct Connect uplink groups creation script

optional arguments:
  -h, --help            show this help message and exit
  -p PORTS, --ports PORTS
                        specify ports
  -m BREAKOUT_MODE, --mode BREAKOUT_MODE
                        specify the breakout mode
  -f FEC_MODE, --fec FEC_MODE
                        set fec mode
  -v, --verbose         enable verbosity
```

Example: Create DirectConnect Uplink Port Group on port 7

```
root@ovcamn06r1-bx8pca# pca-create-uplink -p "7"

root@ovcamn06r1-bx8pca# pca-list-uplink
Port Name      Ports      Speed  Breakout  Network
-----
default_5_1    5:1 5:2    10g    10g-4x    default_external
default_5_2    5:3 5:4    10g    10g-4x    None
dconnect_7_1   7:1        25g    25g-4x    None
dconnect_7_2   7:2        25g    25g-4x    None
dconnect_7_3   7:3        25g    25g-4x    None
dconnect_7_4   7:4        25g    25g-4x    None
-----
6 rows selected.
```

Example: Create DirectConnect Uplink Port Group on port 8 with 10g-4x breakout

```
root@ovcamn06r1-bx8pca# pca-create-uplink -p "8" -m 10g-4x

root@ovcamn06r1-bx8pca# pca-list-uplink
Port Name      Ports      Speed  Breakout  Network
-----
default_5_1    5:1 5:2    10g    10g-4x    default_external
default_5_2    5:3 5:4    10g    10g-4x    None
dconnect_7_1   7:1        25g    25g-4x    None
dconnect_7_2   7:2        25g    25g-4x    None
dconnect_7_3   7:3        25g    25g-4x    None
dconnect_7_4   7:4        25g    25g-4x    None
dconnect_8_1   8:1        10g    10g-4x    None
dconnect_8_2   8:2        10g    10g-4x    None
dconnect_8_3   8:3        10g    10g-4x    None
dconnect_8_4   8:4        10g    10g-4x    None
-----
10 rows selected.
```

- B.** Run the `pca-create-network` script to setup the `direct_connect` network and assign to it ALL the ports created using `pca-create-uplink` script. Verify using `pca-list-network` script.

```
# pca-create-network -h
usage: pca-create-network [-h]

This script is used ONLY to create 'direct_connect' network. Please note, this
script takes no arguments.

optional arguments:
  -h, --help  show this help message and exit

Example: Create DirectConnect network

root@ovcamn06r1-bx8pca# pca-create-network
....
[12/17/2020 17:00:33 3490075] INFO (pca-create-network:108) Details:
{
  "route_dest": null,
  "status": "ready",
  "underlay_device": "tun-ext",
  "reserved": false,
  "subnet_block": null,
  "name": "direct_connect",
  "default": false,
  "vlan": 3062,
  "route_gw": null,
  "zfssa_ip": null,
  "vni": 13062,
  "vlan_enabled": true,
  "devs": [
    "vx13062"
  ],
  "multicast_group": "239.1.1.1",
  "prefix": null,
  "bond": null,
  "sub_interface": null,
  "type": 6,
  "ports": [
    "7:2",
    "7:4",
    "8:2",
    "8:4",
    "7:1",
    "7:3",
    "8:1",
    "8:3"
  ],
  "UUID": null
}

root@ovcamn06r1-bx8pca# pca-list-network
```

Network Name	Type	Default	Status
oracle-database-storage	storage_network	False	ready
underlay_external	unknown	True	ready
default_internal	rack_internal_network	True	ready
direct_connect	external_network	False	ready ←-----
underlay_internal	unknown	True	ready
default_external	external_network	True	ready

- C. Run the `pca-add-network` script to add the `direct_connect` network to desired Compute Nodes or Tenant Groups. The network should be added ONE-AT-A-TIME to each Compute Node

```
# pca-add-network -h
usage: pca-add-network [-h] [-n NETWORK_NAME]
                        (-c COMPUTE_NODE | -t TENANT_GROUP)

Example: Add direct connect network to Compute Nodes

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn07r1
..
...
[12/17/2020 17:16:10 3549184] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn07r1.

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn08r1
..
...
[12/17/2020 17:17:54 3553896] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn08r1.

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn09r1
..
...
[12/17/2020 17:19:31 3559384] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn09r1.

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn10r1
..
...
[12/17/2020 17:21:00 3563890] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn10r1.

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn11r1
..
...
[12/17/2020 17:27:33 3575940] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn11r1.

root@ovcamn06r1-bx8pca# pca-add-network -c ovcacn12r1
..
...
[12/17/2020 17:28:48 3579489] INFO (network_manager:731) Network: direct_connect, has
been added on compute_node: ovcacn12r1.

root@ovcamn06r1-bx8pca# pca-list-uplink
Port Name           Ports      Speed  Breakout  Network
-----
default_5_1         5:1 5:2    10g     10g-4x    default_external
default_5_2         5:3 5:4    10g     10g-4x    None
lab1_pg             1 2       40g     40g       lab1_net
dconnect_7_1        7:1       25g     25g-4x    direct_connect
dconnect_7_2        7:2       25g     25g-4x    direct_connect
dconnect_7_3        7:3       25g     25g-4x    direct_connect
dconnect_7_4        7:4       25g     25g-4x    direct_connect
dconnect_8_1        8:1       10g     10g-4x    direct_connect
dconnect_8_2        8:2       10g     10g-4x    direct_connect
dconnect_8_3        8:3       10g     10g-4x    direct_connect
dconnect_8_4        8:4       10g     10g-4x    direct_connect
-----
```

D. (Optional) Update (add/delete ports) the direct_connect network AFTER it has been added to ComputeNodes

```
# pca-update-network -h
usage: pca-update-network [-h] -p PORTS (-a | -d) [-f] network_name

This script is used ONLY to modify direct_connect network i.e. add/remove
uplink ports.

positional arguments:
  network_name          Name of the network which is to be modified i.e.
                        'direct_connect'

optional arguments:
  -h, --help            show this help message and exit
  -p PORTS, --port-names PORTS
                        Ports which are to be associated or dissociated.
                        Please note, the port value must be enclosed in
                        single/double quotes and when specifying multiple
                        ports, separate them with space e.g. -p "9" or -p "7
                        9".
  -a, --add            Specify this option when associating uplink ports to
                        direct_connect network
  -d, --delete        Specify this option when dissociating uplink ports
                        from direct_connect network
  -f, --force         Force flag for destructive command. Use this flag to
                        disable the confirmation prompt when you run this
                        command with -d/--delete option.
```

Please note, -a/--add & -d/--delete are mutually exclusive options and --confirm option is only valid for -d/--delete option.

Example: Delete port 8 from direct connect network

```
root@ovcamn06r1-bx8pca# pca-update-network -p "8" -d -f direct_connect

[12/17/2020 17:32:24 3588193] INFO (pca-update-network:57) Validating rack type
..
[12/17/2020 17:32:24 3588193] INFO (pca-update-network:76) Dissociating uplink
port/s: ['8:4', '8:1', '8:3', '8:2'] from network [direct_connect] ..
[12/17/2020 17:32:24 3588193] INFO (pca-update-network:80) Access to all
external systems attached to PCA spine uplink ports : ['8:4', '8:1', '8:3',
'8:2'] have been successfully dissociated from network [direct_connect].

root@ovcamn06r1-bx8pca# pca-list-uplink
Port Name          Ports      Speed  Breakout  Network
-----
default_5_1       5:1 5:2    10g     10g-4x    default_external
default_5_2       5:3 5:4    10g     10g-4x    None
lab1_pg           1 2        40g     40g       lab1_net
dconnect_7_1      7:1        25g     25g-4x    direct_connect
dconnect_7_2      7:2        25g     25g-4x    direct_connect
dconnect_7_3      7:3        25g     25g-4x    direct_connect
dconnect_7_4      7:4        25g     25g-4x    direct_connect
dconnect_8_1      8:1        10g     10g-4x    None
dconnect_8_2      8:2        10g     10g-4x    None
dconnect_8_3      8:3        10g     10g-4x    None
dconnect_8_4      8:4        10g     10g-4x    None
-----
```

Example: Add port 8 back to direct connect network

```

root@ovcamn06r1-bx8pca# pca-update-network -p "8" -a -f direct_connect

[12/17/2020 17:33:27 3590448] INFO (pca-update-network:88) Validating rack type
..
[12/17/2020 17:33:27 3590448] INFO (pca-update-network:101) Associating uplink
port/s: ['8:4', '8:1', '8:3', '8:2'] to network [direct_connect] ..
[12/17/2020 17:33:27 3590448] INFO (pca-update-network:105) Access to all
external systems attached to PCA spine uplink ports : ['8:4', '8:1', '8:3',
'8:2'] have been successfully associated to network [direct_connect].

root@ovcamn06r1-bx8pca# pca-list-uplink
Port Name           Ports      Speed  Breakout  Network
-----
default_5_1         5:1 5:2    10g    10g-4x    default_external
default_5_2         5:3 5:4    10g    10g-4x    None
lab1_pg             1 2        40g    40g       lab1_net
dconnect_7_1        7:1        25g    25g-4x    direct_connect
dconnect_7_2        7:2        25g    25g-4x    direct_connect
dconnect_7_3        7:3        25g    25g-4x    direct_connect
dconnect_7_4        7:4        25g    25g-4x    direct_connect
dconnect_8_1        8:1        10g    10g-4x    direct_connect
dconnect_8_2        8:2        10g    10g-4x    direct_connect
dconnect_8_3        8:3        10g    10g-4x    direct_connect
dconnect_8_4        8:4        10g    10g-4x    direct_connect
-----

```

The following table shows the resulting port-channel configuration for the Client-Network for the database-servers in a full rack Exadata-X8.

Table 1: EXPECTED port channel configuration (Full rack)

PORT-CHANNEL	VPC NUMBER	VLAN	SWITCH 1 PORT	SWITCH 2 PORT	COMMENTS
Po171(SU)	171	3062	Eth1/7/1(P)	Eth1/7/1(P)	DirectConnect port-channel
Po172(SU)	172	3062	Eth1/7/2(P)	Eth1/7/2(P)	DirectConnect port-channel
Po173(SU)	173	3062	Eth1/7/3(P)	Eth1/7/3(P)	DirectConnect port-channel
Po174(SU)	174	3062	Eth1/7/4(P)	Eth1/7/4(P)	DirectConnect port-channel
Po181(SU)	181	3062	Eth1/8/1(P)	Eth1/8/1(P)	DirectConnect port-channel
Po182(SU)	182	3062	Eth1/8/2(P)	Eth1/8/2(P)	DirectConnect port-channel
Po183(SU)	183	3062	Eth1/8/3(P)	Eth1/8/3(P)	DirectConnect port-channel
Po184(SU)	184	3062	Eth1/8/4(P)	Eth1/8/4(P)	DirectConnect port-channel

- Verify the port/port-channel configuration matches the output below (shown for port 7, verify port 8 in case of a half/full rack)

OVCASW22R1# show run int eth 1/7/1-4	OVCASW22R1# show run int po 171, po 172, po 173, po 174
<pre>interface Ethernet1/7/1 switchport mode dot1q-tunnel switchport access vlan 3062 mtu 9216 speed 25000 fec off channel-group 171 mode active no shutdown</pre>	<pre>interface port-channel171 description DirectConnect port-channel switchport mode dot1q-tunnel switchport access vlan 3062 spanning-tree port type edge spanning-tree bpduguard enable spanning-tree bpdufilter enable mtu 9216 speed 25000 vpc 171</pre>
<pre>interface Ethernet1/7/2 switchport mode dot1q-tunnel switchport access vlan 3062 mtu 9216 speed 25000 fec off channel-group 172 mode active no shutdown</pre>	<pre>interface port-channel172 description DirectConnect port-channel switchport mode dot1q-tunnel switchport access vlan 3062 spanning-tree port type edge spanning-tree bpduguard enable spanning-tree bpdufilter enable mtu 9216 speed 25000 vpc 172</pre>
<pre>interface Ethernet1/7/3 switchport mode dot1q-tunnel switchport access vlan 3062 mtu 9216 speed 25000 fec off channel-group 173 mode active</pre>	<pre>interface port-channel173 description DirectConnect port-channel switchport mode dot1q-tunnel switchport access vlan 3062 spanning-tree port type edge spanning-tree bpduguard enable spanning-tree bpdufilter enable mtu 9216 speed 25000 vpc 173</pre>
<pre>interface Ethernet1/7/4 switchport mode dot1q-tunnel switchport access vlan 3062 mtu 9216 speed 25000 fec off channel-group 174 mode active no shutdown</pre>	<pre>interface port-channel174 description DirectConnect port-channel switchport mode dot1q-tunnel switchport access vlan 3062 spanning-tree port type edge spanning-tree bpduguard enable spanning-tree bpdufilter enable mtu 9216 speed 25000 vpc 174</pre>

Step-3: Verify the direct_connect network is seen by OVMM

The screenshot shows the Oracle VM Manager interface with the 'Networking' tab selected. Under 'Networks', there is a table of 'Network Channels'. The 'direct_connect' row is highlighted with a red box.

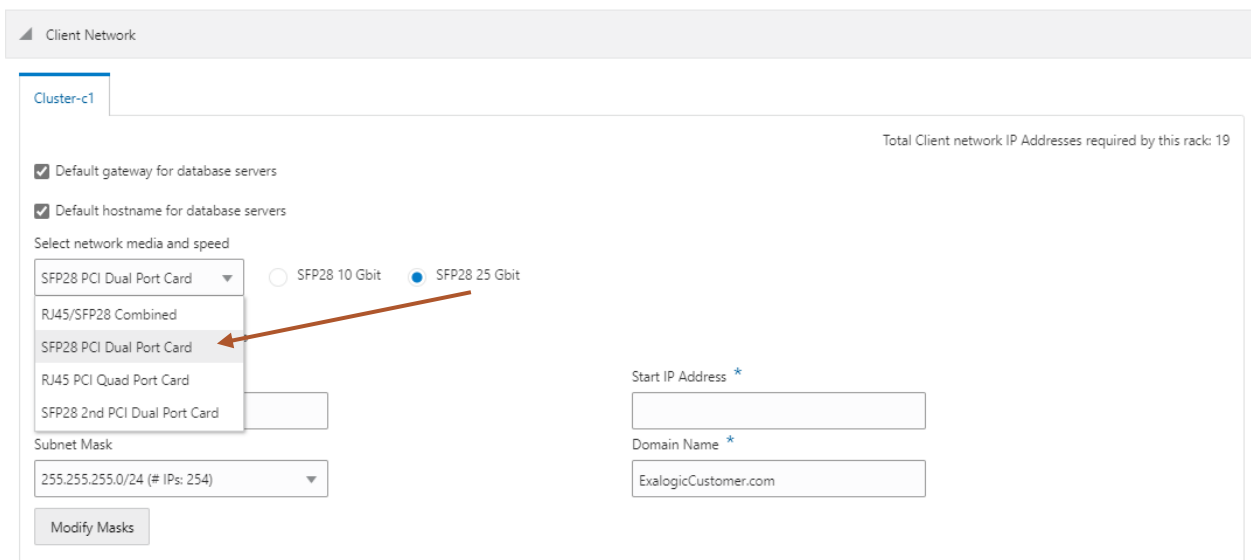
Name	ID	Server Management	Cluster Heartbeat	Live Migrate	Storage	Virtual Machine
192.168.32.0	c0a82000	✓		✓		
192.168.40.0	10ee950564		✓		✓	
default_external	1037661503					✓
default_internal	10f57bf8da					✓
direct_connect	107fb3bf93					✓
external_network	10766aeb16					✓
ota_svd	10225a4a52					✓

CONFIGURING THE ORACLE EXADATA DATABASE MACHINE

The Client-network is represented by the green or red 25Gb SFP28 ports in Figure-2.

The Client-network is configured as bondeth0 or bondeth1 interface on each database node. In our example, we will use bondeth1 comprising eth3 and eth4 and should be configured with Linux bonding in active-active/LACP (mode=4). The configuration steps in this solution brief require the use of LACP. Oracle Exadata Deployment Assistant (OEDA) can assist with the configuration of the Client-network.

LACP should be configured to use a policy of layer3+4 and a fast polling rate. The default MTU is 1500 bytes, but configuring jumbo frames (9000 bytes) is highly recommended. It reduce the level of IP fragmentation and allow for more efficient processing of large streaming workloads with lower CPU overhead. IP networks perform path MTU discovery and auto-negotiate the MTU size between the client (VMs), switches and database-server. Client interfaces using a smaller MTU can still communicate with switches and servers that support jumbo frames.



Optimal settings for bondeth1 are shown in this example:

Client Access Network Configuration

All IP addresses and corresponding host names should be registered in DNS

Starting IP Address : **10.80.108.101**
Subnet Mask : **255.255.252.0**
Gateway : **10.80.108.1**
Client Network Type : **SFP28 PCI Dual Port Card Optical Bonded**
Client Network Interface : **eth3,eth4**
Network Speed : **25000**

LACP : **Enabled, please ensure switches support LACP and configured for LACP and aggregation**
BONDING_OPTS="mode=802.3ad miimon=100 downdelay=200 updelay=200 lacp_rate=1 xmit_hash_policy=layer3+4"

```

$ cat /proc/net/bonding/bondeth1
Bonding Mode: IEEE 802.3ad Dynamic link aggregation
Transmit Hash Policy: layer3+4 (1)
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 200
Down Delay (ms): 200

```

```

802.3ad info
LACP rate: fast
Min links: 0
Aggregator selection policy (ad_select): stable

```

```

Slave Interface: eth3
MII Status: up
Speed: 25000 Mbps
Duplex: full
<...>
Slave Interface: eth4
MII Status: up
Speed: 25000 Mbps
Duplex: full
<...>

```

Table 2. Transceivers and cable options

TRANSCEIVERS AND CABLES	X-OPT MKTG PN
QSFP28 100Gb Fiber Short-range Transceiver	7119728
QSFP+ 40Gb Fiber Short-range Transceiver	X2124A
SFP28 25Gb Fiber Short-range Transceiver	7341433
SFP+ 10Gb/1Gb Dual-Speed Fiber Short-range Transceiver	X2129A, 7358257
MPO to 4 LC Optical Splitter OM3 10m/20m/50m	X2127A-10M, X2127A-20M, X2127A-50M
CBL,1M,CU SPLITTER,QSFP28 to 4xSFP28	7118363
CBL,2M,CU SPLITTER,QSFP28 to 4xSFP28	7118364
CBL,3M,CU SPLITTER,QSFP28 to 4xSFP28	7118365
CBL,5M,CU SPLITTER,QSFP28 to 4xSFP28	7118366
QSFP+ to 4xSFP+ Splitter Cable 1m/3m/5m	X2125A-1M, X2125A-3M, X2125A-5M

Components and Connections – Cables and Transceivers

Copper Breakout Cable



Optical Breakout Cable



LC-LC Coupler



LC-LC Cable



QSFP28 Transceiver



SFP28 Transceiver



LC-LC OM3 MMF Patch Panel



FREQUENTLY ASKED QUESTIONS

1. Can older Exadata racks be connected to PCA directly?

Yes, the steps shown above will be the same for Exadata-X7, X8 which have the Dual-Port-25Gb-Adapter. Same process can be used for older Exadata racks if they have 10Gb SFP+ adapters in the database servers. In such case, the breakout mode should be set to 10g-4x when using the `pca-create-uplink` command.

2. The default-gateway, DNS and NTP servers are accessed over the Client-network on the dbnode. What options are available if the Client-Network connects to the PCA spine switches?

There are several ways to solve this, depending on the type of deployment.

- Set the default route on dbnodes via any other network which is connected to the customer's network. This option is best when all the applications accessing the database are running on PCA VMs.
- Setup a **service-VM** inside PCA to **provide DNS service** for the VMs and the dbnodes in Exadata
[Oracle® Linux 7 Setting Up Networking - Chapter 3 Configuring the Name Service:](#)

- Setup a **service-VM** inside PCA to **forward the DNS queries**, received from dbnodes over the `direct_connect` network, to customer's DNS server via `default_external` network.

In this case, the DNS service is configured as **forward-only**. In this mode, the DNS service forwards all queries to another name server and caches the results, which reduces local processing.

Adding the "**forwarders**" parameter to the options section of the `/etc/named.conf` file allows the DNS to forward any unresolved names to alternative DNS servers.

```
allow-query { any; };
forward only;
forwarders { 201.172.4.100; 201.172.8.100; }; #Customer's DNS servers reachable via default_network
recursion yes;
# Open ports for DNS
iptables -A INPUT -p tcp --dport 53 -j ACCEPT
iptables -A INPUT -p udp --dport 53 -j ACCEPT
```

- Only if the above changes cannot be done, a pair of `direct_connect` ports (for example `po174`, i.e. 1/7/4 or `po184` i.e. 1/8/4 from both switches) can be used to connect to customer's data-center switch.
 - Both ports **MUST** go to the same customer switch
 - Customer switch ports should be in a LACP port-channel
 - The `vlan/network` should be different than the `default_network` used by `po151` (ports 1/5/1-2)

3. Are VLANs supported?

Yes, both Exadata and PCA support traffic isolation using VLANs. Please refer to the product documentation.

<https://docs.oracle.com/en/engineered-systems/exadata-database-machine/dbmin/exadata-network-requirements.html#GUID-7D772F28-3D43-4530-8A48-14F90EBE3926>

<https://docs.oracle.com/en/engineered-systems/private-cloud-appliance/2.4/admin-2.4.3/admin-managing-vms.html#admin-managing-ovmm-networking>

<https://docs.oracle.com/en/virtualization/oracle-vm/3.4/user/vmusg-networking-vlans.html#vmusg-networking-vlans-create>

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com.

Outside North America, find your local office at oracle.com/contact.

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

Copyright © 2021, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

This device has not been authorized as required by the rules of the Federal Communications Commission. This device is not, and may not be, offered for sale or lease, or sold or leased, until authorization is obtained.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 0120

DIRECT CONNECT EXADATA-X8 TO PCA
March 2121

Author: Lalit Bhola

