Configuring Shared Ethernet Adapter Failover

This document describes some general concepts related to Shared Ethernet Adapter (SEA), the procedures to configure SEA Failover, as well as migrating from an existing configuration to SEA Failover.

Shared Ethernet Adapter:

A Shared Ethernet Adapter can be used to connect a physical network to a virtual Ethernet network. It provides the ability for several client partitions to share one physical adapter. SEA can only be configured on the Virtual I/O server (VIOS) and requires the POWER Hypervisor and Advanced POWER virtualization feature. The SEA, hosted on the Virtual I/O server, acts as a Layer-2 bridge between the internal and external network.

Restrictions with configuring SEA Failover:

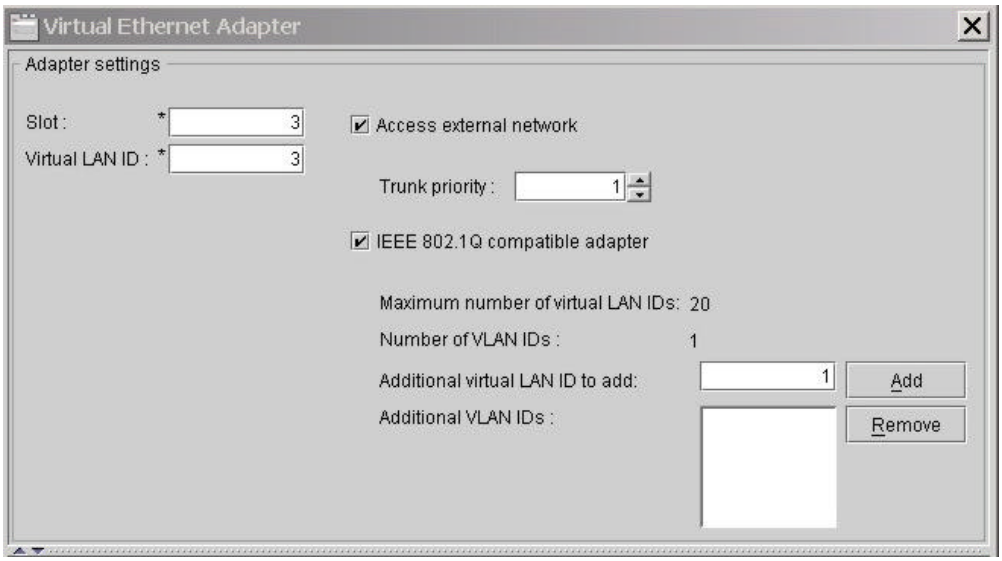
- It can only be hosted on the VIOS and not on the client partition.
- The VIOS running Integrated Virtualization Manager (IVM) cannot implement SEA Failover because only one single VIOS can be configured on the P5 with IVM.
- SEA Failover was introduced with FixPack 7, i.e., in Virtual I/O server version 1.2. so both the Virtual I/O servers need to be at this minimum level.

Requirements for configuring SEA Failover:

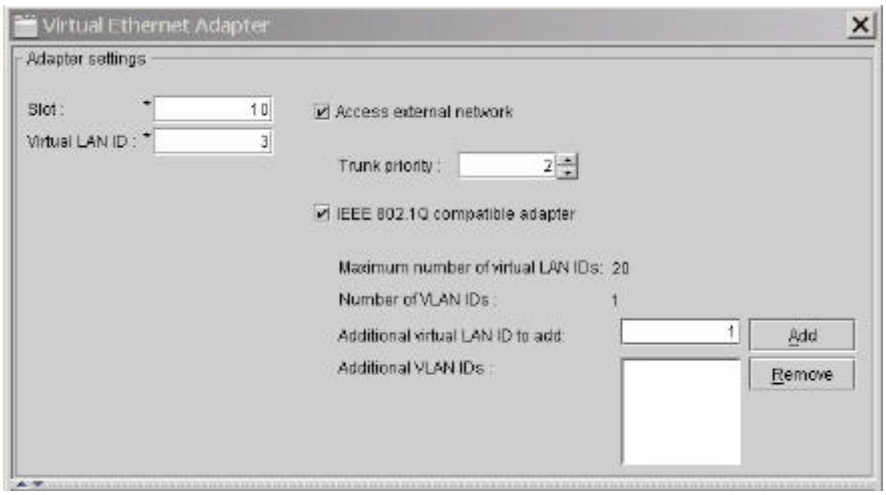
- Two Virtual I/O servers, both atleast at FixPack 7, with bridging functionality of Shared Ethernet Adapters configured on them.
- One SEA on one VIOS acts as the primary (active) adapter and the second SEA on the second VIOS acts as a backup (standby) adapter.
- Each SEA will have at least one virtual Ethernet adapter with the "Access external network" flag (previously known as "trunk" flag) checked.
- This adapter on both the SEAs has the same PVID, but will have a different priority value.
- A SEA in ha_mode (Failover mode) might have more than one trunk adapters, in which case all should have the same priority value.
- The priority value defines which of the two SEAs will be the primary and which will be the backup. The lower the priority value, the higher the priority, e.g. an adapter with priority 1 will have the highest priority.
- An additional virtual Ethernet adapter , which belongs to a unique VLAN on the system, is used to create the control channel between the SEAs, and must be specified in each SEA when configured in ha_mode.
- The purpose of this control channel is to communicate between the two SEA adapters to determine when a fail over should take place.

Steps to configure SEA Failover:

- Check that the P5 is attached to the HMC.
- Both the VIO servers should be at least at version 1.2. The latest FixPack has some SEA fixes, so we recommend upgrading to the latest FixPack level.
- Each VIO server should have at least one available physical Ethernet adapter assigned to it.
- Using the HMC GUI, create a virtual Ethernet adapter on each of the VIO servers, configure the adapter as the trunk adapter, each with a different trunk priority value.
- The SEA configured with the trunk adapter with the lowest priority value will be the active (primary) adapter. Set the priority=1 on the adapter with "Access external network" flag checked on the VIOS that you have chosen to configure as the primary SEA.
- Set the priority=2 or higher on the adapter with "Access external network" flag checked on the second VIO server. This VIOS will have the backup SEA adapter.
- Make sure that the PVID and the VID values of the trunk adapter for VIOS 1 and VIOS 2 are the same.

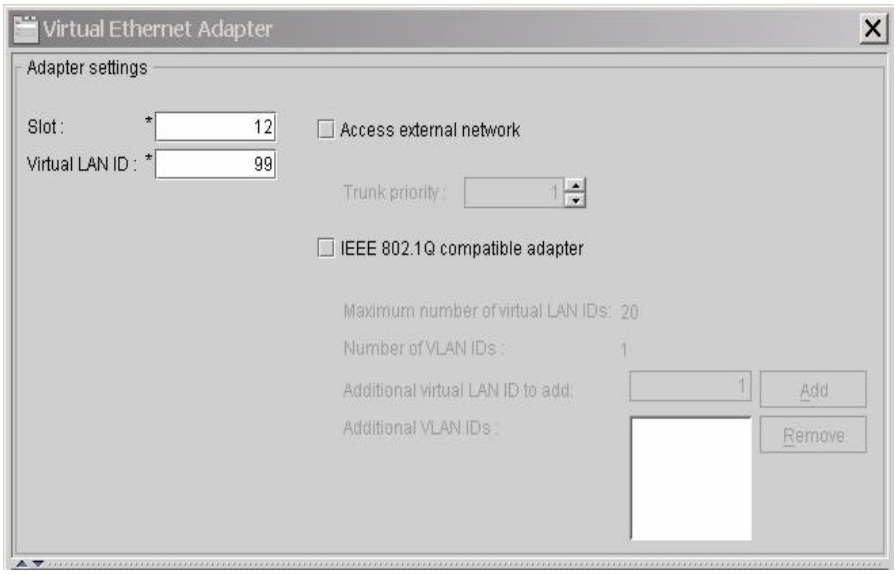


VIOS 1 – configure virtual Ethernet with "Access external network"



VIOS 2 – configure virtual Ethernet with "Access external network"

- Configure another virtual Ethernet adapter (to be used for the control channel) on each of the VIO servers. These adapters must have the same PVID. Double check that this PVID is unique to the system, i.e., that no other virtual Ethernet adapter on the P5 uses this PVID. The following screen shot shows the configuration of control channel of VIOS 1. The screen shot for configuration of control channel of VIOS 2 would look the same (however, the slot number may differ).



VIOS 1 – configure virtual Ethernet adapter for control channel

- Then on CLI, first on one VIOS and then on the second, run the following command:

mkvdev -sea *physical_adapter* -vadapter *virtual_adapter* -default *virtual_adapter*\
-defaultid *PVID_of_virtual_adapter* –attr ha_mode=auto \
ctl_chan= *control_channel_adapter*

```
$ lsdev |grep ent
ent0      Available  Virtual I/O Ethernet Adapter (l-lan)
ent1      Available  2-Port 10/100/1000 Base-TX PCI-X Adapter (1410890
ent2      Available  2-Port 10/100/1000 Base-TX PCI-X Adapter (1410890
ent3      Available  Virtual I/O Ethernet Adapter (l-lan)
ent4      Available  Virtual I/O Ethernet Adapter (l-lan)

$ mkvdev -sea ent1 -vadapter ent4 -default ent4 -defaultid 3 -attr
ha_mode=auto ctl_chan=ent3
ent5 Available
en5
et5

$ lsdev -dev ent5 -attr
attribute     value description                                user_settable

ctl_chan      ent3  Control Channel adapter for SEA Failover   True
ha_mode       auto  High Availability Mode                     True
…
```

Screen shot from VIOS 2 while configuring the SEA

Migrating From Existing Configuration to SEA Failover:

It is possible to migrate an existing SEA configuration on the servers, with NIB on the client, to SEA Failover with only a short, planned downtime on client LPARs.
A prerequisite is a working DLPAR environment, as some virtual Ethernet adapters will be added using DLPAR. An adapter added via DLPAR must also be added to the profile, in order to maintain the configuration after a reboot.
The NIB on the client LPAR can be removed at a later time and can be scheduled at a different time for each client LPAR. This implies a short downtime on a client LPAR.

Overview of working configuration:

VIOS1	adapter	type	bridged	priority	pvid	slot
	ent0	physical	-	-	-	int
	ent2	virtual	yes	1	2	8
	ent3	SEA	-	-	2	

VIOS2	adapter	type	bridged	priority	pvid	slot
	ent0	physical	-	-	-	int
	ent1	virtual	yes	1	3	3
	ent2	SEA	-	-	3	

```
Client     adapter     type          bridged     priority     pvid  slot
           ent0        virtual       -           -            2     2
           ent1        virtual       -           -            3     3
           ent2        NIB (primary=ent0)      -            -     -
```

Switch configuration from NIB & SEA to SEA Failover:

- Add a virtual Ethernet adapter to each VIOS which will be used as the control channel for SEA Failover. After the DLPAR operation, add the adapter to the profile.

- Activate SEA Failover feature on VIOS 1 by running the chdev command to switch on control channel and ha_mode:

```
$ chdev -dev ent3 -attr ha_mode=auto ctl_chan=ent4
ent3 changed
$ lsdev -dev ent3 –attr
…
ctl_chan      ent4  Control Channel adapter for SEA Failover    True
ha_mode       auto  High Availability Mode                      True
…
```

- Add a virtual Ethernet adapter to VIOS 2 which has the same PVID as virtual Ethernet on VIOS 1, check the "Access external network" box and set the priority to a lower value (higher number) than the one configured on VIOS 1.

- Remove SEA on VIOS 2 as it uses a different PVID than SEA on VIOS 1:

```
$ chdev -dev en2 -attr state=detach
$ rmdev -dev en2
en2 deleted
$ rmdev -dev et2
et2 deleted
$ rmdev -dev ent2
ent2 deleted
```

The virtual Ethernet adapter used with this SEA is of no use for SEA Failover and may be removed.

- Create SEA with Failover feature on VIOS 2

```
$ mkvdev -sea ent0 -vadapter ent4 -default ent4 -defaultid 2 -attr
      ha_mode=auto  ctl_chan=ent3
ent2 Available
en2
et2
```

- The interface over the newly-configured SEA on VIOS 2 may now be configured with an IP address in order to have network access to VIOS 2 again.

The client LPAR will run on the primary channel, even if VIOS 1 is down, as a failover would now take place at the VIOS/SEA level. The NIB adapter may be removed on the client LPAR at any time now.

Cases Initiating Failover:

- When the standby SEA detects that the keep-alive (heartbeat) messages are no longer received over the control channel
- When the active SEA detects that a loss of physical link is reported by physical Ethernet adapter's device driver
- On VIOS with primary adapter, when ha_mode is manually set to standby
- When the active SEA detects it is no longer receiving replies from the IP address it is pinging (if the "Internet Address to Ping" feature is enabled).

Test Cases for SEA Failover:

Here are the five scenarios where SEA Failover can be tested and should work:

- Manual Failover
  - ✓ Set ha_mode to standby on primary with chdev command:
    chdev -dev <SEA device> -attr ha_mode=standby
  - ✓ Reset it back to auto and the SEA should fail back to the primary:
    chdev -dev <SEA device> -attr ha_mode=auto
  - ✓
- Primary VIOS shutdown
  - ✓ Reboot the VIOS on the primary for fail over to backup adapter
  - ✓ When the primary VIOS is up again, should fail back to the primary adapter.
- Primary VIOS error
  - ✓ Deactivate primary VIOS from the HMC for fail over to backup adapter
  - ✓ Activate the primary VIOS for the fail back to the primary again
- Physical link failure
  - ✓ unplug the link of the physical adapter on primary for fail over to the backup adapter
  - ✓ and replug of physical link of physical Ethernet adapter on primary for the fail back to the primary
- Reverse boot sequence
  - ✓ Shut down both the VIO servers
  - ✓ Reboot the VIOS with backup SEA and the adapter becomes active

- ✓ Reboot the VIOS with primary SEA and fail back happens to the primary SEA

Advantages of using SEA Failover over Network Interface Backup (NIB) on the client:

Network Interface Backup was used before the new failover feature was available - pre-FixPack 7. This is implemented on the client for high availability in case one VIOS went down. This involved at least a Shared Ethernet Adapter configured on each of the two VIO servers with a different PVID on the trunk adapter on each of them. A control channel is not required to be configured on the VIO servers. On the client partition, a NIB is configured using two virtual Ethernet adapters, one with PVID from the SEA on the VIOS1 and a second one with the PVID from the SEA on the VIOS 2.

- In NIB, the client partition needs 2 virtual Ethernet adapters, implemented on the client. In SEA Failover, only one virtual Ethernet adapter needs to be configured on the client.
- The client partition configuration is simpler with SEA Failover: there is no logic implemented for failover in the client as opposed to NIB.
- Redundancy and High Availability-related functions implemented on the VIOS itself in SEA Failover.
- SEA Failover can be used with IEEE 802.1Q VLAN-tagging, while NIB cannot be.

Troubleshooting:

- mkvdev to create SEA in ha_mode needs to be run in single step on the second VIOS, not in two steps
- Please check that an IP address is not configured on either the virtual or the physical Ethernet adapters that you are going to use for the SEA, otherwise the configuration will fail.
- If you have multiple SEAs configured on each of the VIO servers, then for each SEA pair, you need to configure a separate control channel with a unique PVID on the system.
- If loading a new "seadd" driver, remove and recreate all the SEA adapters on the VIOS to have the kernel extension active, a reboot is not required.
- The cable pull on the primary SEA adapter fails over to backup adapter but the cable plug in back does not have the fail back to the primary if the SEA's physical adapter is an EtherChannel device --- there is an APAR already assigned to the problem: IY82579.