HOWTO: Link Aggregation - As Essential As Disk Mirroring

Contributed by John R Peck

As LVM mirroring is best practice for disks, so 802.3a Link Aggregation should be regarded as the thing to do where there are multiple NIC ports and a TCP/IP service to be run. It's a variation on Etherchannel which gives HACMP-like resilience more simply.

Configuring Ether Channel As A Link Aggregation In Backup Mode

"Etherchannel" has been around for some time, there is a "smit etherchannel" in AIX 4.3.3 for example. It's a proprietary Cisco implementation which requires a capable switch and for all involved connections to be to that switch with "link polling" activated. By linking NIC ports as an "Etherchannel" it is possible to achieve greater combined throughput or load balancing - that would be using the "standard" or "round robin" modes.

Link Aggregation, as in "802.3ad", is an Open standard for essentially the same thing. In AIX there are some added control features. Importantly, as shown here it does not require a single special switch to provide resilient fail-over between NIC ports in "backup" mode. The backup adapter is not normally in use, it is a hot standby and will fail-back automatically after a fail-over - much like an HACMP controlled pair of adapters. The 802.3ad features came in with AIX V5.1, and have been modified a bit in AIX 5.2 and 5.3. The screens here reflect AIX 5.3.

802.3ad ether channel link aggregation in backup adapter mode provides a resilient twin-path connection for a single service IP address by creating a new virtual interface for TCP/IP configuration to use. Only identical adapter types are supported in such a pairing, although mixed types of ports on different cards may well work. The MAC address of the primary adapter in the pairing (ent0 say) will be used as the MAC address for the combined interface (of say ent0 and ent1 to create a new ent2).

Before you start, work out which logical enX interface is associated with which physical NIC port,

so that you know which ports you want to conjoin. And of course decide what IP addresses etc you will use.

The pairing must be done before any configuration of IP addresses, but after making the underlying interface tuning. So use "smitty chgenet", "smitty chif" and "chdev -l enX –a tcp_nodelay=1" first. Set "jumbo frames" on each adapter if also required in the later aggregation, make sure link polling is activated.

Once established as an aggregation, the individual component interfaces cannot be tuned.

A pairing cannot be made if the individual interfaces have IP addresses assigned.

So let's consider en0 and en1 to be paired...

“smit etherchannel” - select the "Add An EtherChannel / Link Aggregation" option, and then "en0" say as the primary of the pair:

EtherChannel / IEEE 802.3ad Link Aggregation

Move cursor to desired item and press Enter.

```
 List All EtherChannels / Link Aggregations
 Add An EtherChannel / Link Aggregation
 Change / Show Characteristics of an EtherChannel / Link Aggregation
 Remove An EtherChannel / Link Aggregation
 Force A Failover In An EtherChannel / Link Aggregation
 +------------------------------------------------------------------+
 |                  Available Network Interfaces              |
 |                                                   |
 | Move cursor to desired item and press F7.                  |
 |    ONE OR MORE items can be selected.                    |
 | Press Enter AFTER making all selections.             |
 |                                                   |
 |   ent0                                        |
 |   ent1                                        |
 |   ent2                                        |
 |                                                   |
 | F1=Help          F2=Refresh          F3=Cancel      |
 | F7=Select        F8=Image            F10=Exit       |
 F1| Enter=Do         /=Find             n=Find Next     |
 F9+------------------------------------------------------------------+
```

 Only selecting one adapter there, as when more than one adapter is selected there it requires the Etherchannel proprietary switch configuration.

 In the next screen, tab-toggle to 802.3ad mode and select "ent1" as the "Backup Aadapter" to "en0" from an "F4" list:

```
            Add An EtherChannel / Link Aggregation
 Type or select values in entry fields.
 Press Enter AFTER making all desired changes.

                              [Entry Fields]
  EtherChannel / Link Aggregation Adapters        ent0            +
  Enable Alternate Address                no              +
  Alternate Address                 []              +
  Enable Gigabit Ethernet Jumbo Frames        no              +
  Mode                          8023ad          +
  Hash Mode                       default         +
```

```
Backup Adapter                           ent1          +
   Automatically Recover to Main Channel    yes              +
Internet Address to Ping              []
Number of Retries                     []            +#
Retry Timeout (sec)                   []            +#


F1=Help       F2=Refresh      F3=Cancel       F4=List
Esc+5=Reset      F6=Command      F7=Edit         F8=Image
F9=Shell      F10=Exit      Enter=Do
```

The “8023ad” mode requires no special switch configuration to link ports in this manner. The AIX implementation of this standard allows the administrator to choose which adapters are linked at the pSeries end – with other implementations it is usually negotiated. The backup adapter will not be used until the live one fails-over to it, fail-back is then automatic when the original connection is restored.

The final options relating to an IP address to ping, enable the connection to be constantly doing a heartbeat check from the higher protocol layers, although that is not absolutely necessary for the fail-over to work because of the underlying link polling.

Finally configure TCP/IP as usual with the new “en2” (the next available number) interface that has been configured here.