

Link Aggregation

A link aggregation vagy IEEE 802.1AX egy szabvány, mely leírja, hogy használjunk szimultán több portot egy adatkapcsolatra. Ezzel a szabvánnyal túlléphetjük az egy kábelben használható sebességet, vagy redundánssá tehetünk egy kapcsolatot a magasabb rendelkezésre állás érdekében.

A szabvány közismertebb neve mára már az IEEE 802.3ad, mely magába foglal több olyan technológiát mint: Ethernet Trunk, NIC teaming (Windows), bonding (Linux), EtherChannel (Cisco) stb.

A probléma az 1990-es évek elején kezdődött, mikor rájöttek, hogy egy hálózat sebessége nem lineárisan nő, nem úgy mint a sávszélesség-igény. Ekkor kezdtek el használni két fizikai interfészt egy logikai csatornával helyettesíteni. Hozzá kell tenni, az effajta technológia mint kliens, mint kiszolgáló oldalon manuális konfigurációt igényel.

A másik ok, amiért alkalmazzák az a egy csatorna meghibásodása (single point of failure) mely tipikus hibája a port – cable – port felépítésű hálózatoknak. A szokásos struktúráknak nagy hátránya (számítógép – switch, switch – switch), hogy ha a kábel be van dugva, akkor is megszakadhat a probléma. Ezt orvosolhatja, ha a kábelezést is redundánsan oldjuk meg, hisz a OSI modell magasabb rétegeiben helyet foglaló protokollok nagy többsége nincs felkészítve a hálózat egyes részeinek kiesésére.

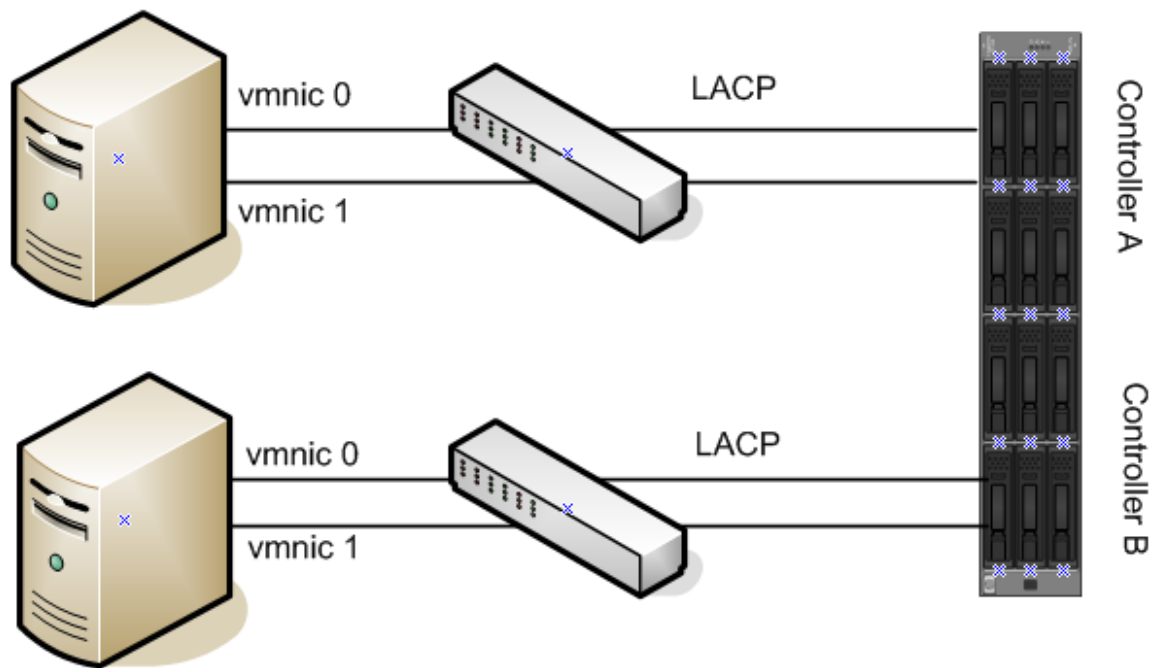
Link Aggregation Control Protocol

Az IEEE leírásában a Link Aggregation Control Protocol (LACP) egyezteti le a különböző portok közötti kommunikációt a végponttal, amennyiben a végponton található eszköz is képes LACP kommunikációra. Ezzel lehetőség nyílik néhány fizikai interfész egy logikai interfészé alakítása.

a manuális(statikus) konfigurálás hátrányai:

- Amennyiben egy eszköz alsóbb rétegek-béli feladatokat lát el (media-converter), nem tudja érzékelni ha egy összeköttetés nem él, így egy „halott” linken küldi az adatokat, mely veszteséggel jár
- egy statikusan konfigurált hálózatban, több kábel áll rendelkezésre, melyből egyet kihúzva is működőképes hálózat marad, mely később a hálózat rendellenes működéséhez vezethet (hurok)

Az LACP csomagokat (LACPDU) küld azokon a portokon melyeken a protokoll be van kapcsolva, és amennyiben van LACP képes eszköz a detektálhatóvá válik az eszközök közötti több csatorna melyből egy logikai csatornát hoznak létre.



1. ábra Tipikus LACP felépítés

Az LACP-t két féleképpen lehet konfigurálni:

- Aktív – mikor mindenképpen küld csomagokat a felkonfigurált portokon
- Passzív – mikor csak akkor kezd ez egy host LACP csomagokat küldeni ha ő is kap ("speak when spoken to")

A passzív mód alkalmas a hálózat hibáinak kiküszöbölésére is. Például amennyiben egy switch passzív, míg egy szerver aktív módban van felkonfigurálva amennyiben a szerver meghibásodik, vagy leáll, úgy a switch is abbahagyja a LACP beszélgetéseket, amennyiben a szerver elindul, úgy ismét felépíti a kapcsolatot több csatornán keresztül és létrejön a logikai csatorna.

Használat

A Link aggregation segítségével nagysebességű és olcsó gerinchálózatokat lehet létrehozni mely nagyobb sávszélességet, kínál mint egy fizikai port.

A LACP segítségével egy hálózati összeköttetés sebessége drasztikusan megnőhet, ha az igény úgy hozza, anélkül, hogy új hardware-t kellene vennünk, valamint különböző load-balancing (erőforrás elosztás) algoritmusok is segíthetnek e sebesség elérésében.

A legtöbb gerinchálózat építésénél több kábelt fektetnek le (még ha nem is kell rögtön), hisz egyszer nő meg a munkadíj, és nem kell újra felásni a fél országot, ha a hálózaton változtatnak. LACP képes ezeket a redundáns vonalakat használni a hálózati sebesség növekedésére.

Hálózati kártyák

A hálózati kártyák „összetrömkölése” segítségével létre tudunk hozni két 1Gb/s interfészből 1 2Gb/s kapcsolatot.

A Microsoft nem támogatja natívan a Link Aggregation-t (Windows 2008-ig), de a gyártók a Windowshoz adott driverek mellett olyan programokat is mellékelnek melyek lehetővé teszik az IEEE 802.3ad használatát Windows rendszerekben.

- ANS - Advanced Networking Services (INTEL) az Intel Fast ethernet és Gigabit kártyáihoz
- Nvidia Network Access Manager/Firewall Tool .

megjegyzés: a Windowsban a kártyák összekapcsolását teaming-nek nevezik

A Linux/UNIX rendszerekben a bonding (itt így hívják) natívan megtalálható, amíg a NIC támogatva van a kernelben.

Bonding beállítása Debian Etch alatt

A 3com Switch 2948-SFP Plus beállítása.

Először is szükségünk van egy VLAN létrehozására, ahol beállítjuk a LACP-t.

Baseline Switch 2948-SFP Plus
Device Summary [Device View]

Device View | Color Key

Device Summary
Save Configuration

Administration ▶
Device ▶
Port ▶
Security ▶
Monitoring ▶
Help

Logout

Copyright © 2007
3Com Corporation.
All Rights Reserved.

Device Summary Information

Product Description:	3Com Baseline Switch 2948 Plus
System Name:	Baseline Switch 2948-SFP Plus
System Location:	
System Contact:	
Serial Number:	2TGGAUNEDF940
Product 3C Number:	3CBLSG48
System Object ID:	1.3.6.1.4.1.43.1.8.62
MAC Address:	00:22:57:ed:f9:40
System Up Time:	91 days, 4 hours, 58 minutes, 44 seconds
Software Version:	01.00.09
Boot Version:	00.00.15
Hardware Version:	1.0.1

Poll Now The default polling interval is 60 sec

http://192.168.100.8/index.htm

2. ábra 3Com switch kezelőfelülete

Látható, hogy a switchben az utolsó 8 port aktív melyből 4-4 a két összekötendő gép 4 portjára van rákötve.

Baseline Switch 2948-SFP Plus
Device > VLAN [Modify VLAN]

Setup **Modify VLAN** Modify Port Rename Remove Port Detail VLAN Detail

Select a VLAN to modify: 5

Select membership type: ☒ Untagged ☐ Tagged ☐ Not A Member ☐ Not available for selection

Select All Select None

NOTE:

- You may set different membership types on multiple ports before applying ;
- To remove a port in a Link-aggregation, please select all the ports in the group;

Untagged Member(s) Tagged Member(s)

Port 41-48

Logout

Copyright © 2007
3Com Corporation.
All Rights Reserved.

3. ábra VLAN-ok Létrehozása

Ezen az ábrán létrehozunk egy VLAN-t melyhez hozzáadjuk azokat a portokat melyek az átvitelben részt vesznek.

Baseline Switch 2948-SFP Plus
Port > Link Aggregation [Modify]

Summary **Create** **Modify** Remove

Select Aggregation to modify: 1

Select ports to add to aggregation or De-select ports to remove from aggregation:

Selected Ports: ☒ Member of the modified aggregation. Deselected Ports: ☐ Not a member of any aggregation. ☐ Member of an existing aggregation or VLAN.

Summary

Group ID	Type	Member Ports
1	LACP	41, 42, 43, 44
2	LACP	45, 46, 47, 48

Logout

4. ábra LACP engedélyezése

Ezek után a 4-4 porton engedélyezzük a LACP és ezzel a switch konfigurációjával végeztünk.

Debian beállítása

Fontos megjegyezni, hogy ezen beállítások csak Debian Etch Disztribúció alatt működnek.

Etch alatt adjuk ki a következő parancsot:

```
apt-get install ifenslave
```

A következő parancssal listázhatjuk ki a rendelkezésre álló interfészeket:

```
#mii-tool
eth0: negotiated 100baseTX-FD, link ok
eth1: negotiated 100baseTX-FD, link ok
eth2: negotiated 100baseTX-FD, link ok
eth3: negotiated 100baseTX-FD, link ok
eth4: negotiated 100baseTX-FD, link ok
```

Látható, hogy 5 rendelkezésre álló interfész van a gépben, melyből az eth0 az alaplapi interfész, melyet hagyjunk a jelenlegi beállításokkal.

Módok (Bonding Types):

A Linux kernel bonding modulja a következő módokon működhet:

1. Round Robin

A csomagokat Round Robin algoritmussal küldi a rendelkezésre álló interfészeken, így valósítva meg az erőforrás-eloszlást, és a hibatűrést.

2. Active Backup

Ha az elsődleges interfész kiesik, a szolga interfész automatikusan átveszi a helyét, és leklónozza a MAC címét is. Csak hibatűrést eredményez. Nincs szükség a Switch konfigurálásra.

3. Balance XOR

Minden egyes szolgálhoz hozzárendel egy cél MAC címet és csak annak küld adatot ((forrás MAC) XOR (cél MAC)) így valósítva meg az erőforrás-eloszlást, és a hibatűrést.

4. Broadcast

Minden szolga kiküldi az összes adatot, hibatűrést eredményez.

5. 802.3ad

A klasszikus IEEE 802.3ad dinamikus link aggregation. Szükség van arra, hogy a Switch képes legyen LACP protokoll használatára. Majd a protokoll segítségével lekommunikálja a adatküldés lehetőségeit.

Módosítsuk a `/etc/modprobe.d/arch/i386`, adjuk hozzá a következő két sort:

```
alias bond0 bonding
options bond0 mode=5 miimon=200
```

Látható, hogy mi az 5-ös módot választjuk, a `miimon=200` pedig az interfészek monitorozására szolgál melynek időközzeit 200ms-ra állítjuk.

Majd szerkesztjük a `/etc/network/interfaces` fájlt:

```
auto bond0
iface bond0 inet static
address 10.0.0.1
netmask 255.255.255.0
network 10.0.0.0
up ifenslave bond0 eth1 eth2 eth3 eth4
```

Majd következik a `/etc/init.d/networking restart`

Ezzel létrehoztunk egy `bond0` interfészt, melynek szolgálja az `eth1-eth4`. Vegyük észre az `ifconfig` kimenetén az összes interfész MAC címe azonos lett.

```
# ifconfig | grep Ethernet
bond0      Link encap:Ethernet  HWaddr 00:15:17:54:D7:09
eth0       Link encap:Ethernet  HWaddr 00:11:11:BB:D8:01
eth1       Link encap:Ethernet  HWaddr 00:15:17:54:D7:09
eth2       Link encap:Ethernet  HWaddr 00:15:17:54:D7:09
eth3       Link encap:Ethernet  HWaddr 00:15:17:54:D7:09
eth4       Link encap:Ethernet  HWaddr 00:15:17:54:D7:09
```

Végezzük el ezt egy másik számítógépen, csak az IP címet változtassuk meg `10.0.0.2`-re.

```
auto bond0
iface bond0 inet static
address 10.0.0.2
netmask 255.255.255.0
network 10.0.0.0
up ifenslave bond0 eth1 eth2 eth3 eth4
```

Ezzel készen is vagyunk, kezdjük el pingelni a két számítógép között, majd az aktív interfészt kihúzva nem szakad meg a kapcsolat. A syslogot megnézve láthatjuk mi történik a háttérben:

```
#tail -f /var/log/syslog
Nov 16 12:34:41 fekete5 kernel: e1000: eth1: e1000_watchdog: NIC
Link is Down
Nov 16 12:34:41 fekete5 kernel: bonding: bond0: link status
definitely down for interface eth1, disabling it
Nov 16 12:34:41 fekete5 kernel: bonding: bond0: making interface
eth2 the new active one.
Nov 16 12:34:45 fekete5 kernel: e1000: eth1: e1000_watchdog: NIC
Link is Up 1000 Mbps Full Duplex
Nov 16 12:34:45 fekete5 kernel: bonding: bond0: link status
definitely up for interface eth1.
```

Kovács Ákos
2009-11-16