

Feladatátvételi fürtök (Failover clustering, High availability clustering)

Windows Server Failover Clustering

Linux High Availability Clustering

Oracle Clusterware

IBM High Availability Cluster Multiprocessing

Sun Solaris Cluster

Service Availability Forum AIS Availability Management Framework

Hardver architektúra

- **Egyszeres hibapont (SPOF)** kiküszöbölése elfogadható áron
 - Szerver: Tartalék (átveszi az elsődleges szerepét)
 - Diszk: RAID, SAN: blokkos (iSCSI), NAS: fájl szintű adattárhálózat
 - Hálózat: Redundáns kapcsolat kliensekkel és a fürtön belül is
 - Környezet: Áramellátás, légkondicionálás, ...
- **Megosztás: Állapottal rendelkező szolgáltatások**
 - **Shared disk** (shared SCSI, serial attached SCSI, Fibre Channel)
 - Közös használat: fizikai szintű sorosítás, globális zármenedzser szükséges
 - **Shared nothing** (csak hálózati kapcsolat)
 - Kizárólagos használat: logikai szinten biztosított a kizárólagos hozzáférés
 - **Replicated disks** (tükrözött lemezek, aktív és tartalék)
 - Bináris (blokkos), fájl szintű (cluster file system), vagy alkalmazás szintű replikáció (pl. adatbázis szintű változások); többféle teljesítmény opció
- **Topológia:**
 - Pár (aktív-aktív, aktív-passzív), N+1, N+I, gyűrű topológiák

Szoftver architektúra

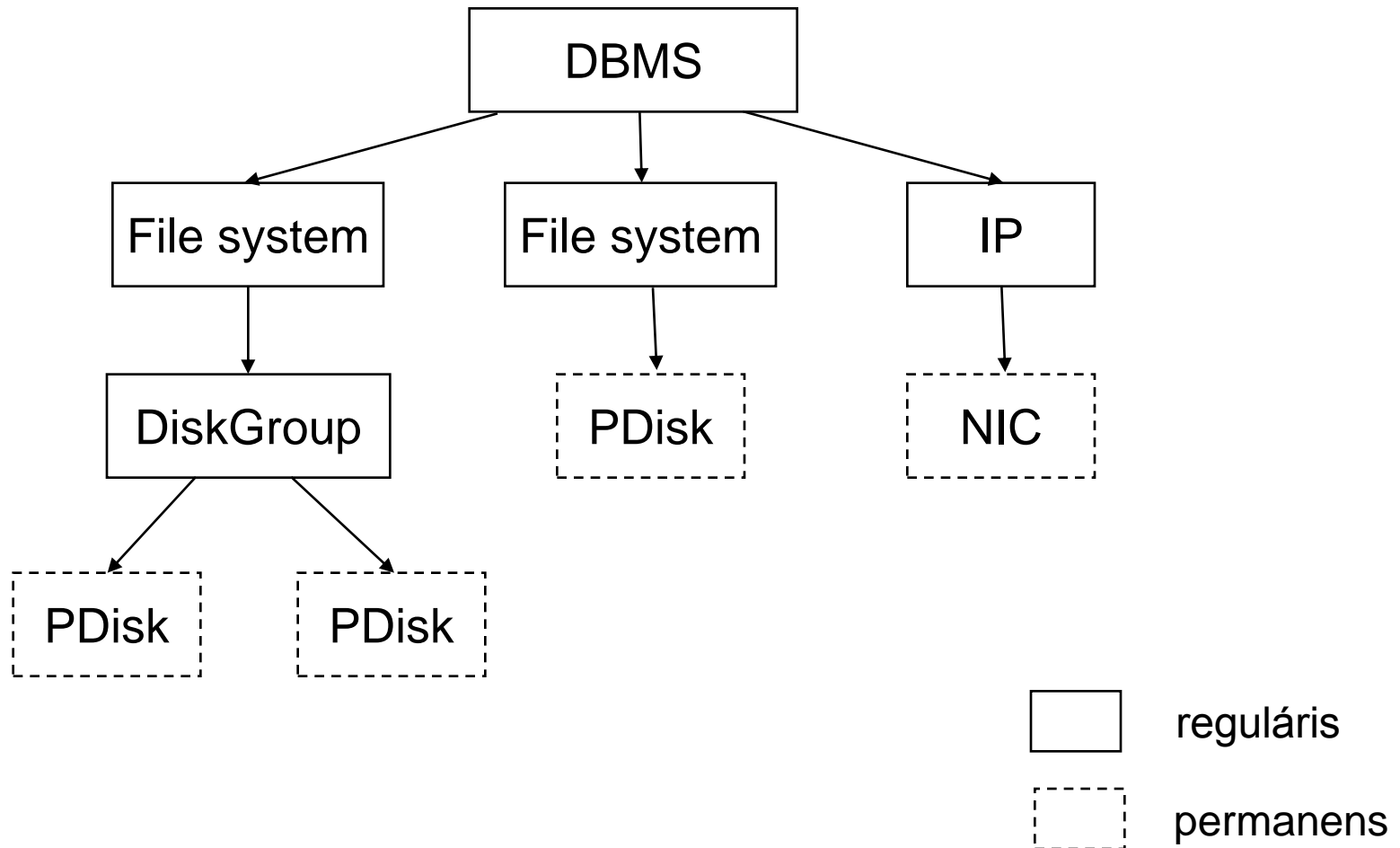
- Operációs rendszer
 - Általában nem fürt specifikus
 - Keretprogrammal integrálható
- **Fürt keretprogram** (HA keretprogram): Fürt menedzselése
 - **Hibadetektálás**
 - Heartbeat, challenge-response
 - **Hibakezelés:**
 - Újraindítás (tranziens hiba esetén, szerverre STOMITH)
 - Feladatátvétel (**failover**), feladat-visszavétel (**failback**)
 - **Értesítések** (újabb hiba már szolgáltatás kimaradást okozhat)
 - **Újrakonfigurálás** (pl. terheléselosztás)
 - Átkapcsolás (**switchover**), visszakapcsolás (**switchback**), ki/beléptetés
- Alkalmazói keretprogram
 - Alkalmazás-specifikus hibadetektálás
 - Alkalmazás szintű állapotmentés, helyreállítás, átkonfigurálás

Failover és failback

- Failover és failback alapja: **Erőforrás függőségi fa**
 - Erőforrások: hardver, szoftver, vagy alkalmazás
 - Erőforrás **csoportok**: Hierarchikus egymásra épülés (ld. példa)
 - Erőforrás **típusok**:
 - **Reguláris**: ki/be kapcsolható (pl. mount/umount)
 - **Permanens**: nem kapcsolható ki (pl. hálókártya)
 - Erőforrások **kezelése**:
 - Szkriptek: bekapcsolás, kikapcsolás (regulárishoz), monitorozás
- Alkalmazás jellemzői:
 - Használt erőforrás csoportok (függőségi fa)
 - Hol indítható, automatikusan indítható-e egy-egy erőforrás
- Alkalmazás **failover** és **failback**:
 - **Leállítás**: Erőforrások kikapcsolása – felülről lefelé
 - **Elindítás**: Erőforrások bekapcsolása – alulról felfelé
- Erőforrás szintű hiba detektálása esetén:
 - Alkalmazás failover, ahol a függőségi fában szerepel

Erőforrás függőségi fa

- Egy példa:



Jellegzetes problémák fürtökben

- **Tudathasadás** (split brain) – térbeli particionálás
 - A fürt **felbomlása** (partíciók)
 - Quorum (többség) képzése szükséges
 - Szerverek többsége (páratlan számú szerver esetén)
 - Tanúlemez (witness disk) birtoklása
 - Tanú fájl (witness file) birtoklása (kijelölt fájlmegosztás, távoli lehet)
 - Általános: Szavazatok hozzárendelése szerverekhez, tanúlemezhez, tanú fájlhoz
- **Amnézia** (amnesia) – időbeli particionálás
 - Meghibásodott, majd javított szerver visszaléptetése, közben az aktív is meghibásodik: aktuális **konfiguráció elveszhet**
 - Megoldás: Fürt konfigurációt közös adattárolóra írni
 - **Fürt adatbázis**: Konfiguráció, változás napló (quorum logging)
- **Szoftverfrissítés**
 - Gördülő frissítés (rolling upgrade): Kiléptetés, frissítés, visszaléptetés sorozat