



IT KOLLEDŽ
TALLINNA TEHNIKAÜLIKOOL

ISS, liidesed, DAS

Andmesalvestustehnoloogiad I395

Edmund Laugasson

edmund.laugasson@itcollege.ee

https://wiki.itcollege.ee/index.php/User:Edmund#eesti_keeles

Käesoleva dokumendi paljundamine, edasiandmine ja/või muutmine on sätestatud ühega järgnevatest litsentsidest kasutaja valikul:

* GNU Vaba Dokumentatsiooni Litsentsi versioon 1.2 või uuem

* Creative Commons Autorile viitamine + Jagamine samadel tingimustel 4.0 litsents (CC BY-SA)

Loenguteemad

- Sissejuhatus intelligentsetesse salvestussüsteemidesse (ISS)
- Liideste kordamine
- *Direct Attached Storage* ehk *DAS*



Intelligentised salvestussüsteemid (ISS)



- Samm edasi RAID ja LVM tehnoloogiast:
 - Hea > Parem > Kõige parem
 - HDD > RAID/LVM > **ISS**
- **Intelligentne salvestussüsteem (ISS)** on võimalusterohke RAID-massiiv, mis pakub kõrgelt optimeeritud sisendi/väljundi (I/O) töötlemise võimalust
- Pakub suuremahulist puhvrit (*cache*) ja mitmeid I/O liideseid suurendamaks jõudlust (*performance*)
- Sisaldab korraldussüsteemi, mis võimaldab:
 - intelligentset puhvrihaldust
 - massiivide ressursside intelligentset haldust, tagatud keerukate algoritmidega
 - ühenduvus erinevate masinatega
 - toetab pooljuhtkettaid, virtuaalset jagamist (*virtual provisioning*), automaatset salvestusmahu järjestamist (*automated storage tiering*) – lisab uue dimensiooni skaleeruvuse, saadavuse osas



Intelligentised salvestussüsteemid (ISS)

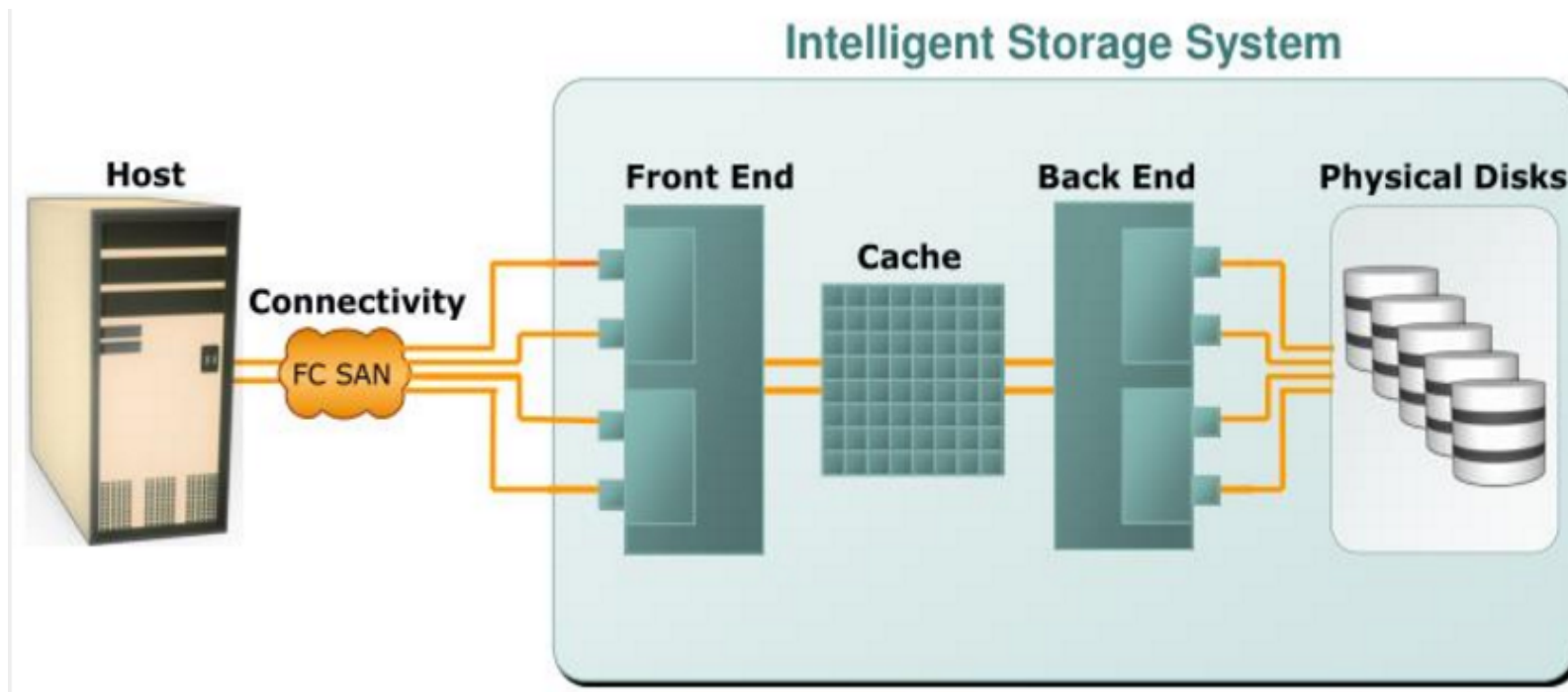
- Samm edasi RAID ja LVM tehnoloogiast:
 - Hea > Parem > Kõige parem
 - HDD > RAID/LVM > ISS
- Eelised:
 - Suurem mahutavus
 - Parem jõudlus, aitab kompenseerida võrgust ja kettapöördumistest tulenevaid viivitusi
 - Kergem andmete hallatavus
 - Parem andmete kaitstus
 - Tõhusam turvalisus ja ligipääsupiirangud



Intelligentesed salvestussüsteemid II

- *Front end* ehk esiliides (kasutajapoolne)
- *Cache* ehk vahemälu
- *Back end* ehk tagaliides (serveripoolne)
- *Physical Disks* ehk füüsilised kettad
- *Watermark* ehk vesimärgistus

Esiliides, vahemälu, tagaliides on sageli ühe trükiplaadi peal koos.



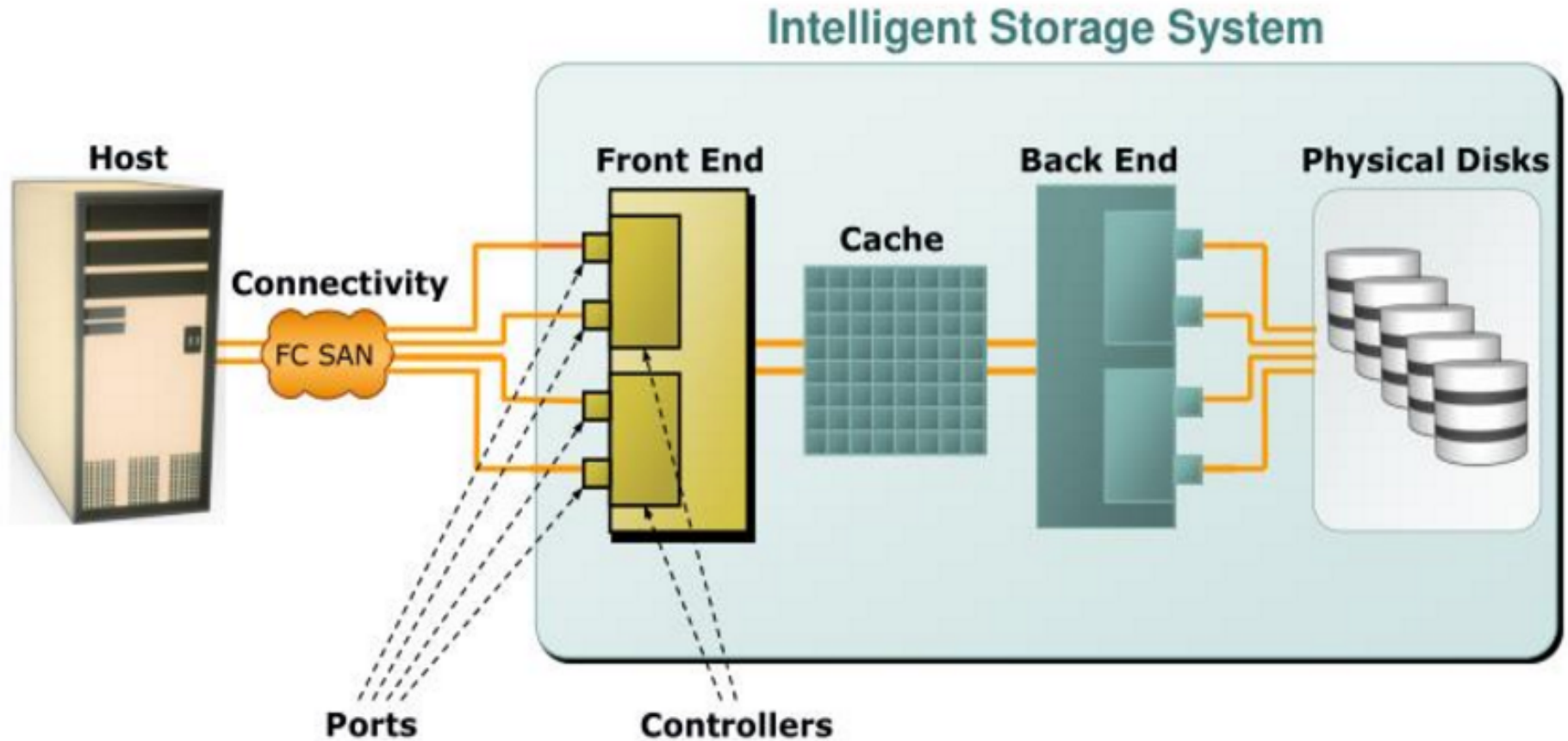


Intelligentsed salvestussüsteemid ?

Mis oli kettalt otsimise aeg (*seek time*) ja mis kettal ligipääsuaeg (*access time*)?



ISS: *front-end*



ISS: *front-end*

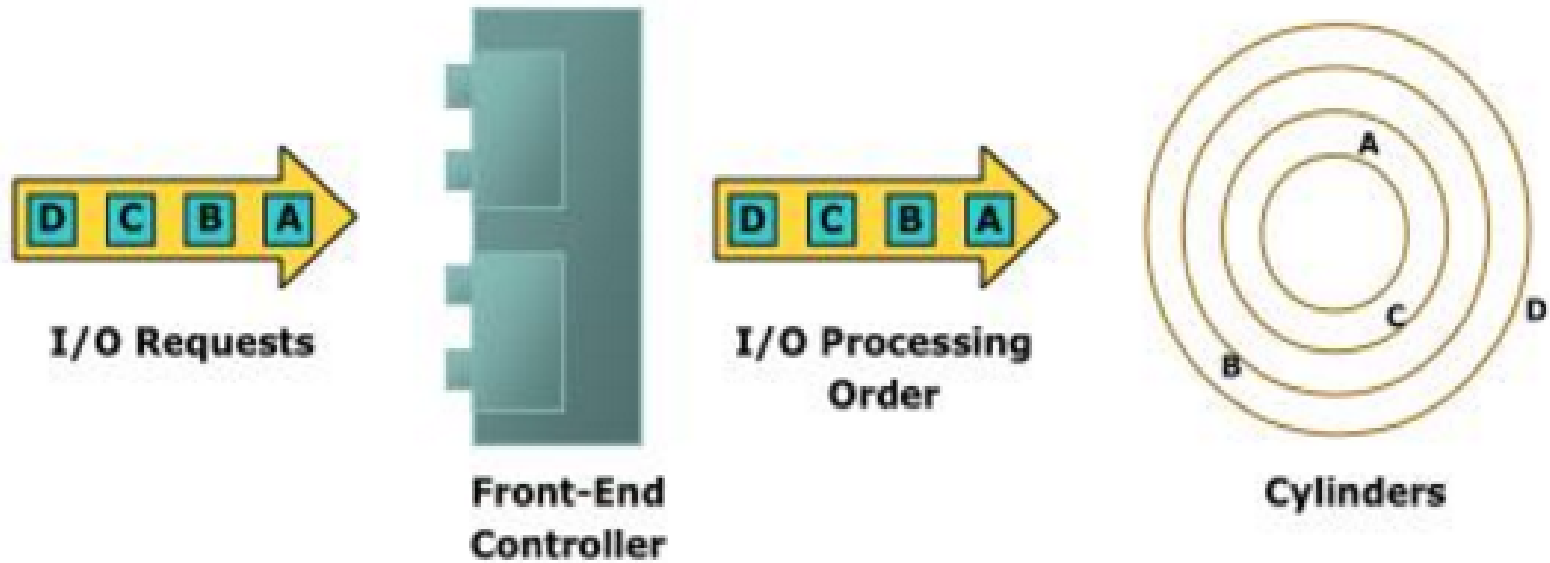


- *Front-end* ehk liides serveri ning kettakasti vahel
 - Pordid - ühendusliidesed serveri ja salvestussüsteemi vahel
 - Kontroller(id) - suunavad andmeid puhvri ja serveri vahel, kui puhver kirjutab andmeid siis saadetakse serverile vastav signaal (teavitus) kontrolleri poolt
- Käsujärjestusalgoritmid
 - **FirstInFirstOut** (FIFO)
 - *Seek time* – andmete järjekord lugemispera asukoha järgi
 - *Access time* – andmete järjekord andmete asukoha ning kosteaja järgi

ISS: *front-end*



- front-end command queuing:

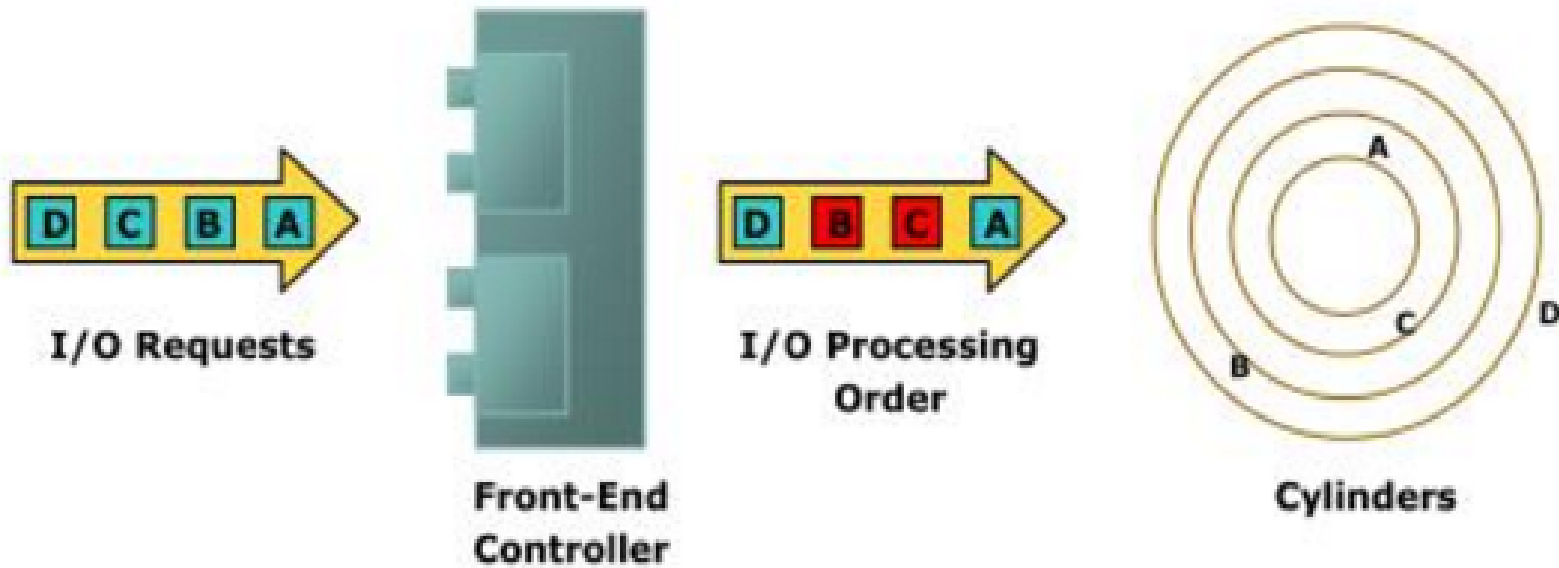


Without Optimization (FIFO)

ISS: *front-end*

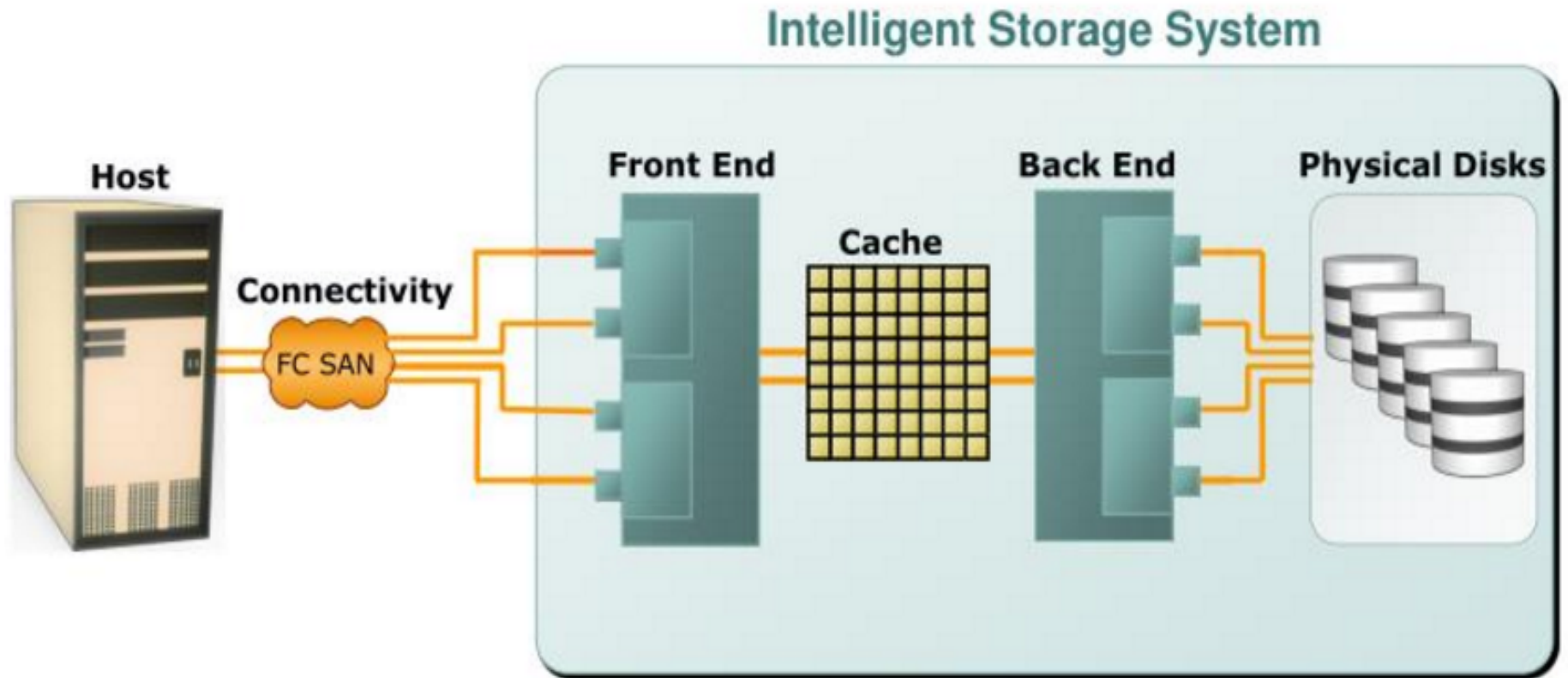


- front-end command queuing:



With command queuing

ISS: *cache*



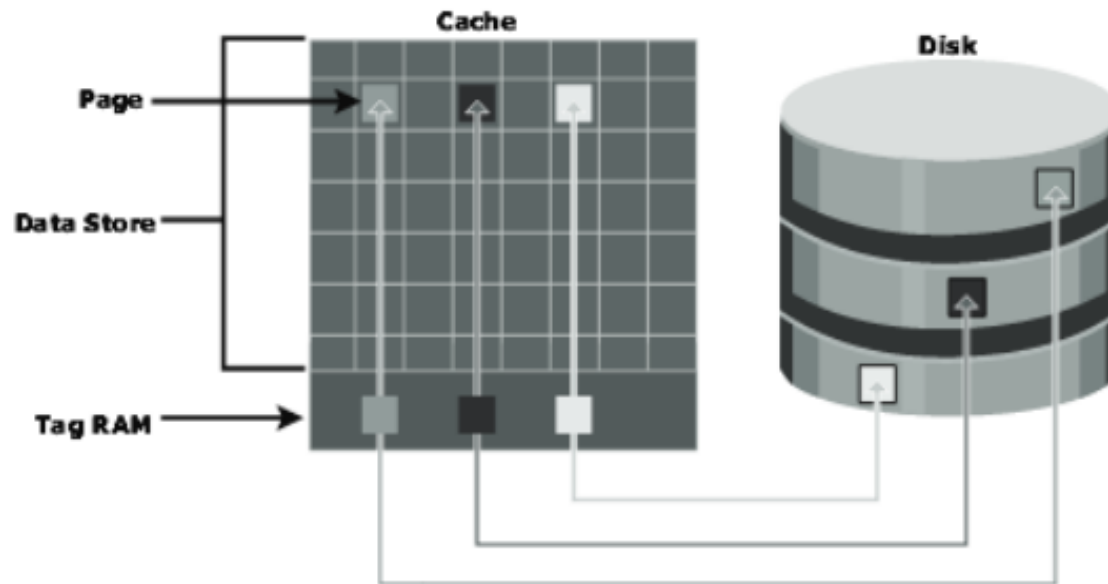
ISS: *cache*

- *Cache* ehk vahemälu esi- ning tagaliidese vahel – pooljuhtmälu kus andmed on paigutatud ajutiselt vähendamaks I/O taotlustele kuluvat aega
- oluline komponent I/O jõudluse suurendamisel
- Vahemälu kasutamine kirjutamisel
 - Läbi vahemälu kirjutamine (3 sammu)
 - Otse tagasi kirjutamine (4 sammu)
- Vahemälu kasutamine lugemisel
 - Cache hit ehk vahemälust lugemine
 - Cache miss ehk kettalt lugemine
- **Mis on *Cache Hit Ratio*?**



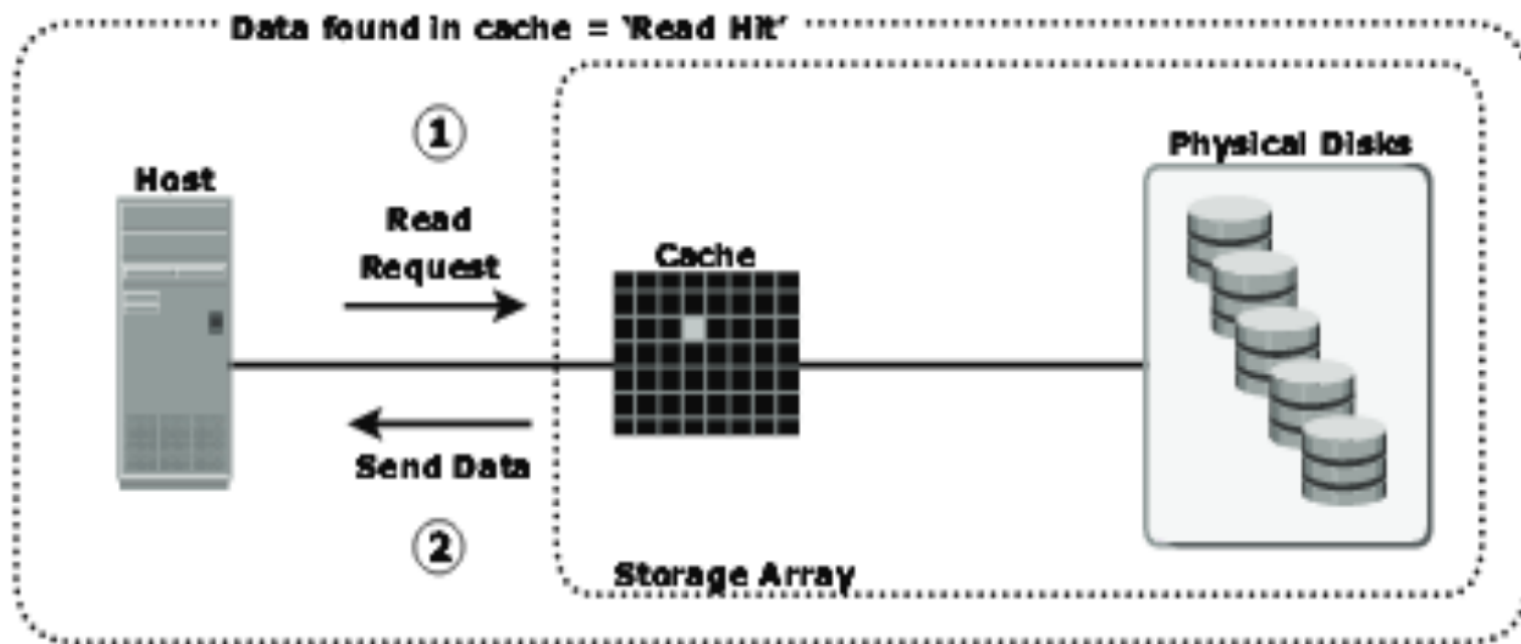
ISS: *cache*

- Vahemälu struktuur – on organiseeritud *page (slot)* üksusteks, mille suurus on määratud rakenduse I/O suurusega. Koosneb andmehoidlast (*Data Store, DS*) ja muutmälu märgendist (*TagRAM, TR*). DS hoiab andmeid, TR infot andmete asukoha kohta DS'is ja füüsilisel kettal (*Disk*)



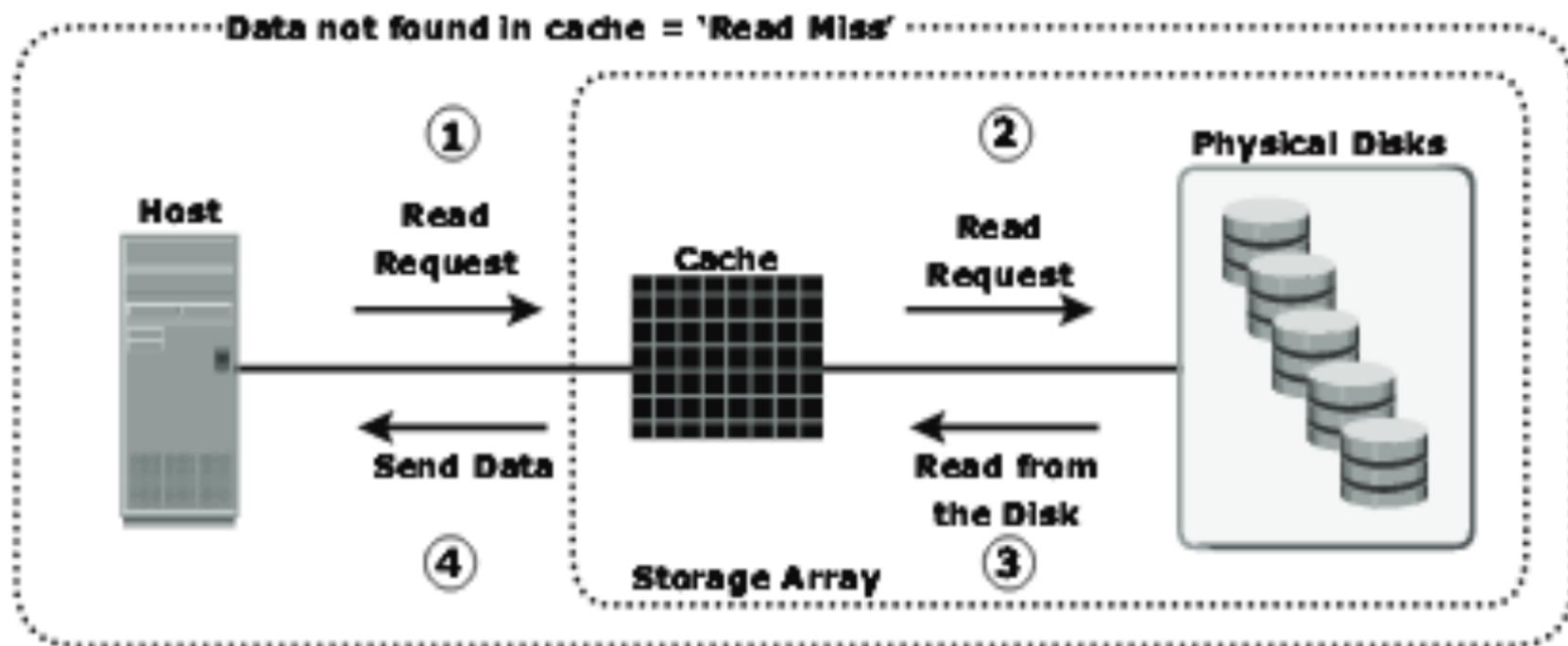
ISS: vahemälu kasutamine lugemisel

- andmete lugemisel vahemälust, *front-end* kontrolleri poole pöördub *TagRAM*'i poole, tegemaks kindlaks kas andmed on vahemälus olemas
- kui andmed on vahemälus olemas siis andmed saadetakse otse vahemälust serverisse, ketaste poole pöördumata (2 operatsiooni), nimetatakse "*Read Cache Hit*" või ka "*Read Hit*"



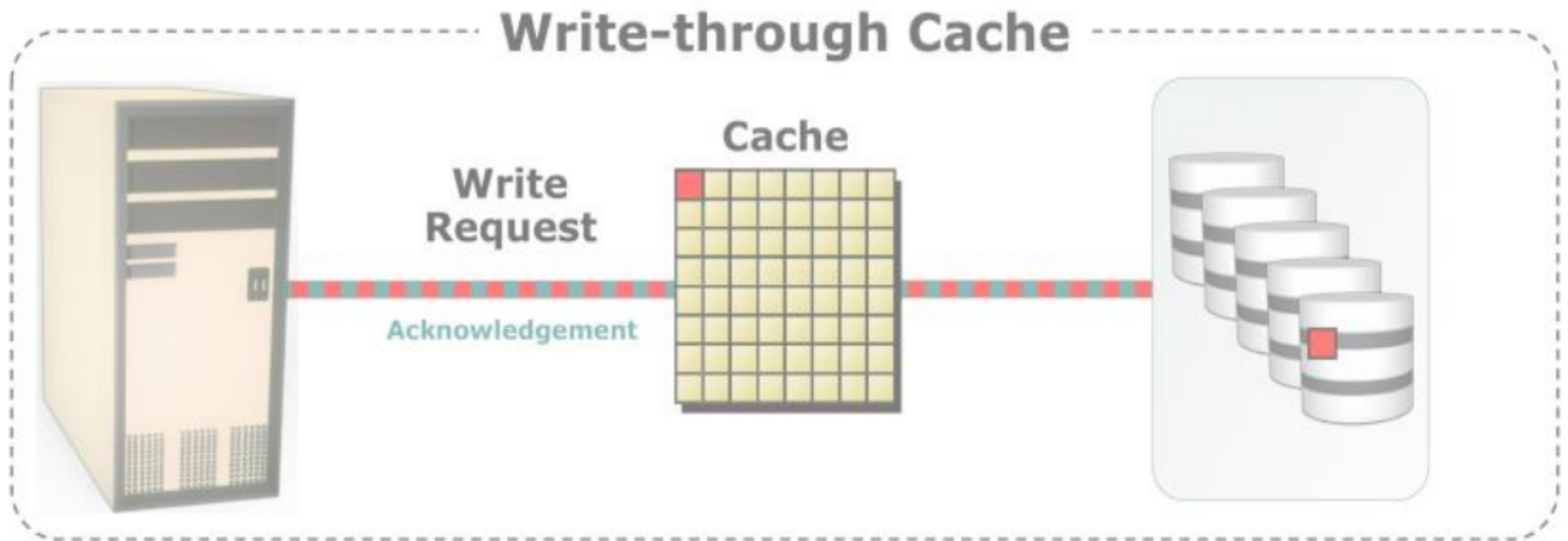
ISS: vahemälu kasutamine lugemisel

- kui andmed ei ole vahemälus olemas siis tuleb need lugeda ketastelt (võtab rohkem I/O aega – 4 operatsiooni), nimetatakse “*Read Cache Miss*” või ka “*Read Miss*”



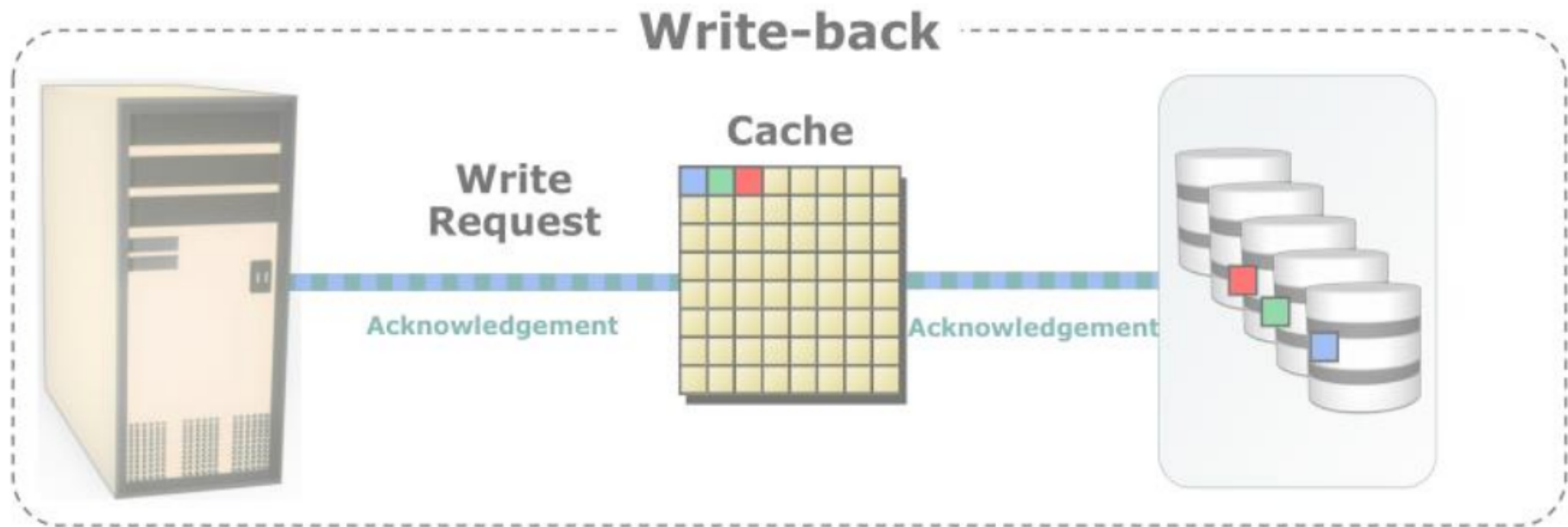
ISS: vahemälu kasutamine kirjutamisel

- Läbi vahemälu kirjutamine (3 sammu) – andmed paigutatakse vahemällu ja kirjutatakse koheselt kettale, teade saadetakse serverile. Kuna andmed on kirjutatud kettale kohe kui need saabuivad siis risk andmete kadumiseks on väike kuid **kirjutamise aeg suurem** tänu kettaoperatsioonidele



ISS: vahemälu kasutamine kirjutamisel

- Otse tagasi kirjutamine (4 sammu) – andmed kirjutatakse vahemällu ja teade serverile saadetakse kohe. **Kirjutamise kiirus on palju suurem** kuna kirjutamise operatsioonid on isoleeritud mehaanilistest ketta viivitustest. Siiski vahemälu tõrke korral kinnitamata andmed lähevad kaduma.





ISS: *cache*

- Vahemälu (*cache*) saab lahendada:
 - spetsiaalne (*dedicated*) – eraldi mäluaadressid on reserveeritud lugemiseks/kirjutamiseks
 - globaalne (*global*) – lugemine/kirjutamine saab kasutada suvalist mäluaadressi ja vahemälu haldus seetõttu tõhusam kuna vaid üks globaalne aadress on vaja hallata



Reklaamipaus

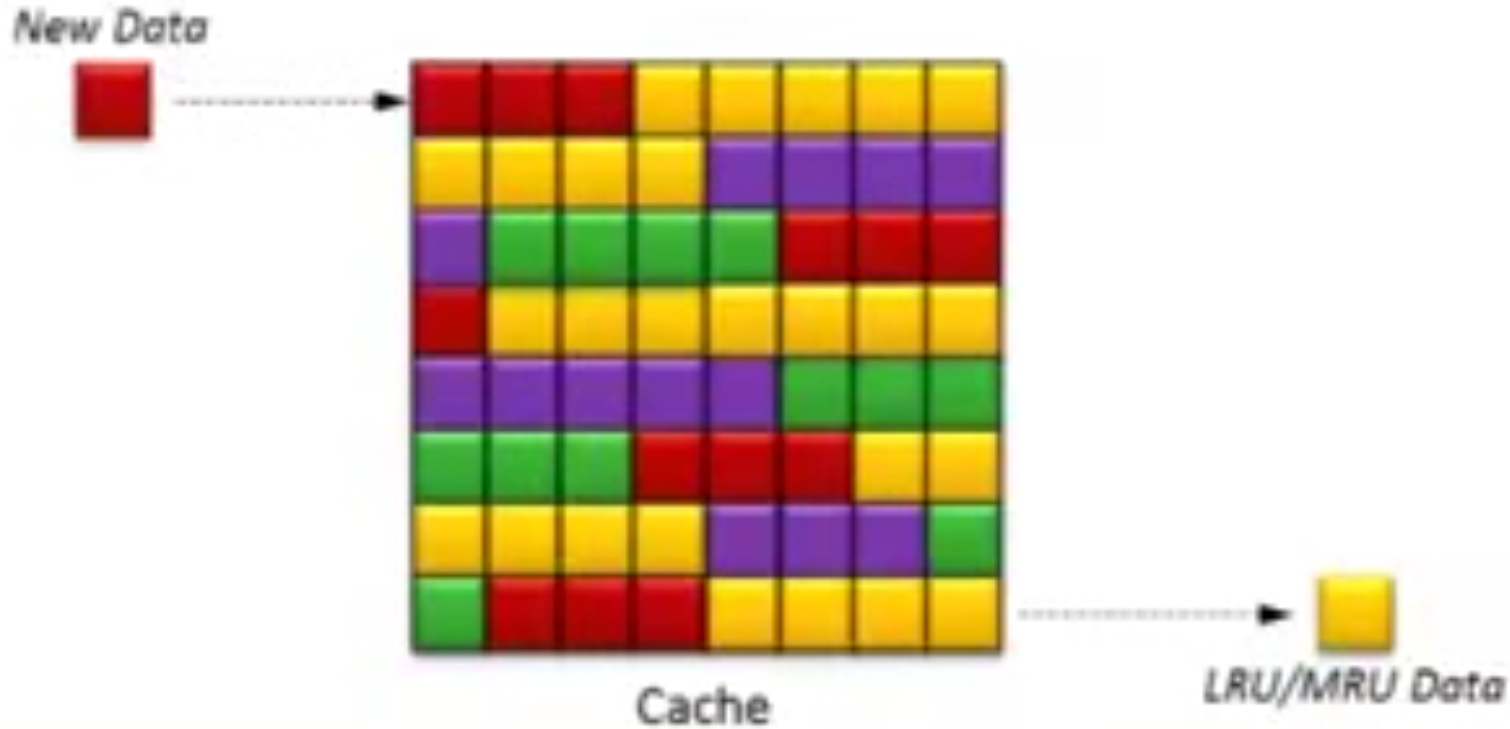


ISS: *cache* – LRU/MRU



- Vahemälus olevate andmete haldus
 - *Least Recently Used (LRU)* – algoritm, mis pidevalt jälgib andmete ligipääsu vahemälus ja teeb kindlaks mälupeasad, mille poole pole pikka aega pöördutud. LRU vabastab või märgib taaskasutuseks sellised mälupeasad. See algoritm põhineb ennustusel, et andmed mida pole pikka aega vajatud – ei hakata ka küsima serveri poolt
 - *Most Recently Used (MRU)* – harvem. On LRU vastand ja siin kõige tihedamini kasutatavad mälupeasad märgitakse vabaks või uuesti kasutatavaks. MRU algoritm põhineb ennustusel, et hiljuti kasutatud andmeid ei vajata mõnda aega.

ISS: *cache* – LRU/MRU



ISS: *cache watermarking*



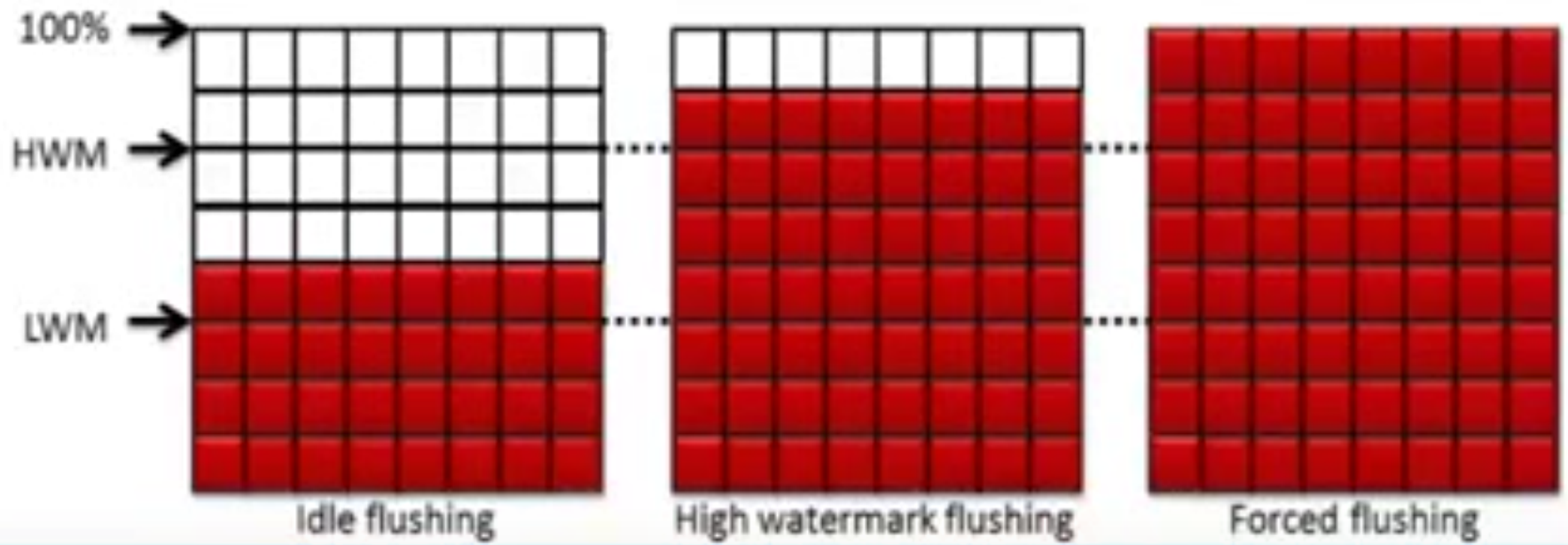
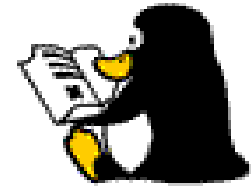
- Vesimärgistus (*watermarking*)
 - Täituvusmõõdik ning kustutuse käivitaja
 - *Idle, High ja Force andmete kustutus*
- Andmete kustutamine (*flushing*) on protsess kus andmed kirjutatakse vahemälust kettale. I/O kasutamise määr ja mustrid – kõrged ja madalad tasemed nimetatakse vesimärkideks, määratakse vahemälus haldamaks kustutamise protsessi.
- HWM (*High Watermark*) on vahemälu kasutamise tase kus salvestussüsteem alustab kõrgkiirusel vahemälu andmete kustutamist. Aktiveeritakse kui vastav tase on saavutatud. Salvestussüsteem rakendab lisaresursid kustutamiseks. Serveri poolelt on siin I/O tase madalaim.

ISS: *cache watermarking*



- LWM (*Low Watermark*) on punkt kus maalt salvestussüsteem peatab kõrgkiirusel või jõuga andmete kustutamise vahemälust ja jääb jõudeolekusse (*idle*)
- Jõudeolekus kustutamine (*idle flushing*) esineb pidevalt tagasihoidlikul määral kui vahemälu kasutamise tase on HWM (*High Watermark*) ja LWM (*Low Watermark*) vahel
- Jõuga kustutamine (*forced flushing*) ilmneb kõrge I/O korral kui vahemälu saavutab 100% oma mahutavusest, mis omakorda mõjutab I/O reageerimisaega

ISS: *cache watermarking*





ISS: *cache*



- Vahemälu kaitsemehhanismid:
 - Vahemälu akud – kuna vahemälu on muutuv (hoiab andmeid vaid toite olemasolul) siis igasugune toite kadumine põhjustab andmekadu kui andmeid ei ole jõutud kettale kirjutada
 - Peegeldus (*cache mirroring*) – iga vahemällu kirjutus hoitakse kahes eraldi mälupesas või isegi eraldi mälumoodulis. Kui tekib vahemälu viga siis peegeldatud asukohas on andmed veel olemas
 - Vahemälu paigutus ketastele (*cache vaulting*) – kasutatakse toite kadumisel spetsiaalsetele ketastele. Kui toide taastub siis andmed nendelt ketastelt kirjutatakse tagasi vahemällu ja sealt edasi neile ketastele kuhu need pidid jõudma.



ISS: *back-end*



- *Back-end* ehk tagaliides
 - Pordid
 - Kontroller(id)
- Tagab andmevahetuse vahemälu ja füüsiliste ketaste vahel
- Pordid on suuremas süsteemis parema käideldavuse ning kättesaadavuse tarbeks dubleeritud
- Füüsilised kettad on ühendatud tagaliidese portide külge
- Tagaliidese kontroller suhtleb füüsiliste ketastega lugemisel, kirjutamisel ja pakub ka piiratud võimalustega ajutist andmesäilitamist



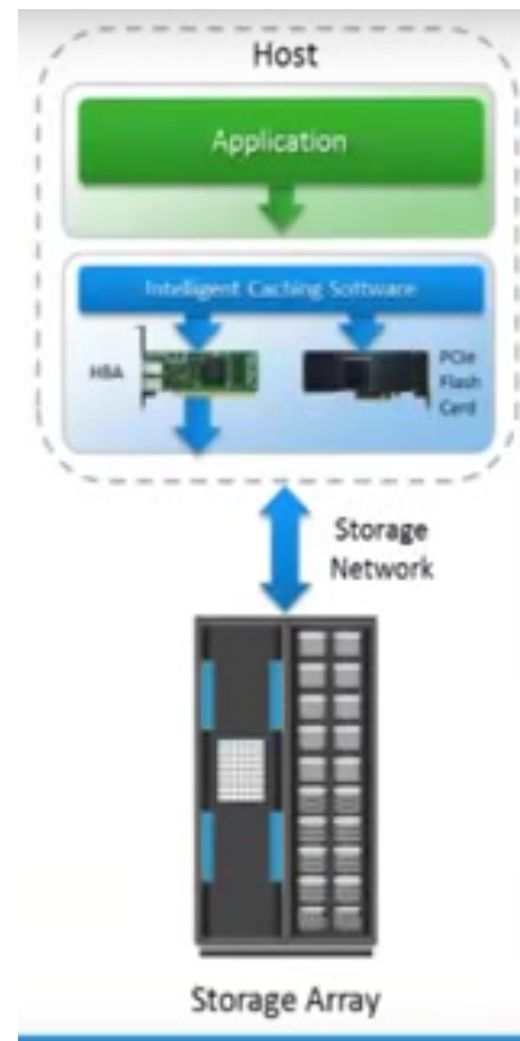
ISS: *physical disk*



- *Physical disk* ehk füüsiline kõvaketas
 - hoiavad andmeid püsivalt ehk siis ka toite puudumisel
 - on ühendatud tagaliidese portidega SCSI, FC (*Fibre Channel*) kaudu
 - intelligentne salvestussüsteem lubab kasutada läbisegi SCSI, FC, IDE/ATA kettaid

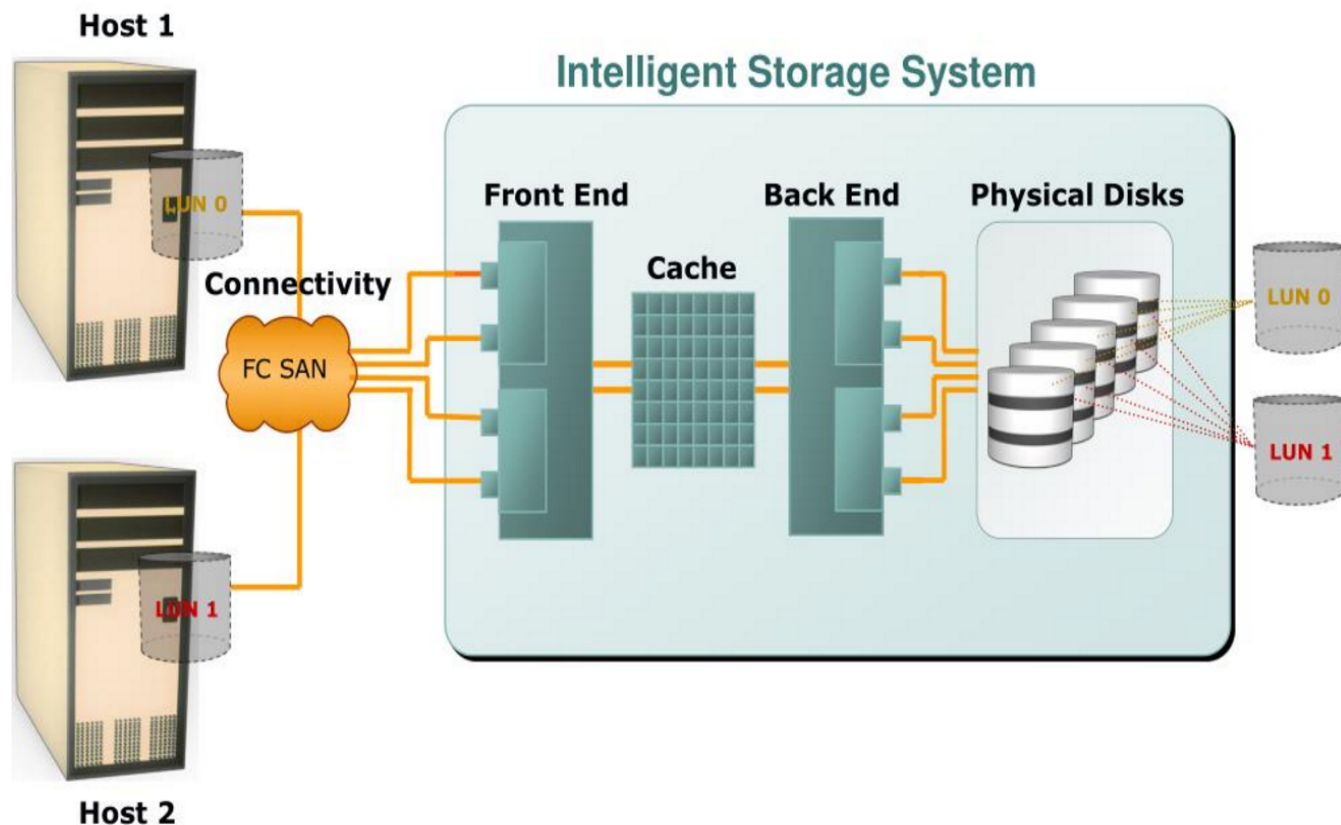
ISS: *server flash-caching* tehnoloogia

- kasutab intelligentset puhverdusmälu tarkvara ja PCIe kaarti serveris
- tõstab olulisel määral rakenduste jõudlust
 - pakub kiirendamist andmete intensiivsel lugemisel, kirjutamisel
 - väldib võrgust ja kettapöördumistest tulenevaid I/O viivitusi
- tuvastab intelligentselt konkreetsed andmed, mida võidakse vajada välk-vahemälus (*flash-cache*)
- kasutab minimaalselt CPU ja RAM ressursse serveris kuna välkmälu haldus on rajatud PCIe kaardile



ISS: *server flash-caching* tehnoloogia

- serveri poolelt on näha RAID-massiiv ja LUN (*Logical Unit Number*)





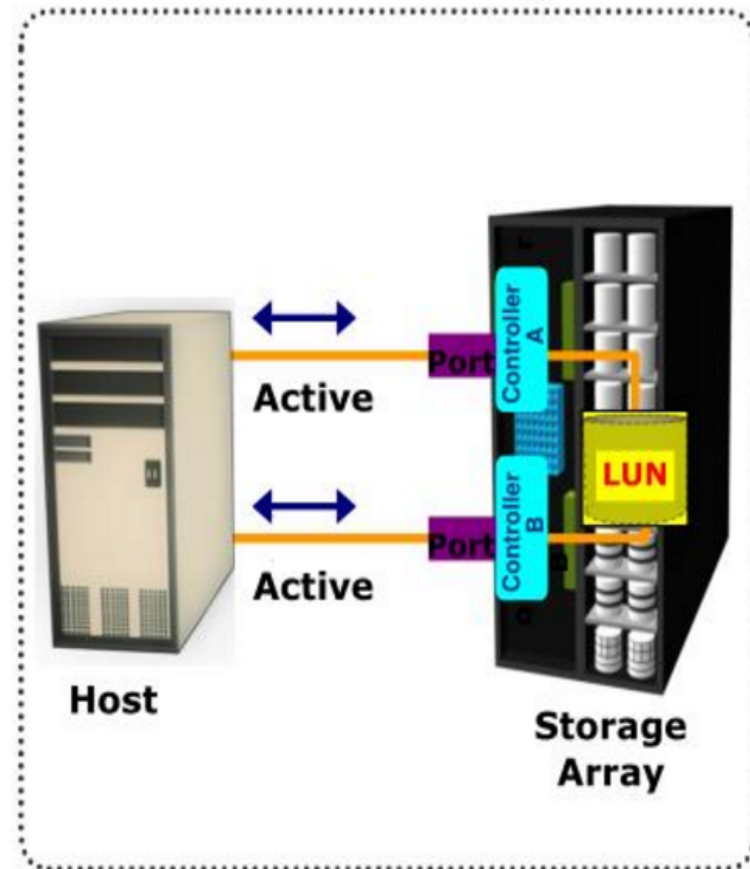
ISS: LUN (*Logical Unit Number*)

- ligipääsu kontrolli süsteem
- väldib volitamata ligipääsu
- kasutatakse andmesalvestustehnoloogiates
- salvestusgrupi loogiline üksus, mis sisaldab ühte või rohkem LUN'i ning ühte serverit

ISS: high-end storage systems

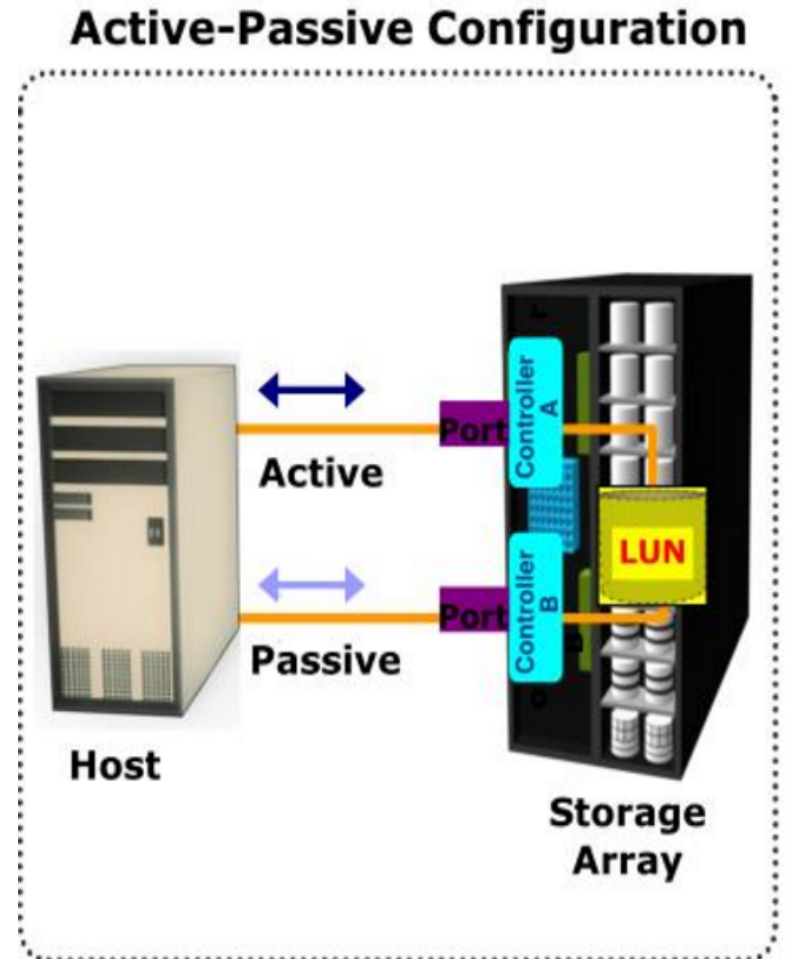
- tuntud ka kui aktiivne-aktiivne massiiv
- I/O serveeritakse läbi kõikide kanalite
- suured salvestusmahud
- suur vahemälu
- kõrge tõrkekindlus
- palju eesliidese porte
- tugi erinevatele ühendusliidestele
- kõrge käideldavus
- hea paralleelne I/O töotlusvõime
- disainitud suurettevõtetele

Active-Active Configuration

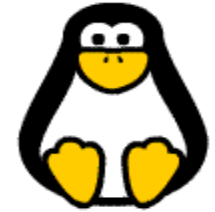


ISS: mid-range storage systems

- tuntud ka kui aktiivne-passiivne massiiv
- I/O serveeritakse läbi aktiivsete kanalite
- teised kanalid jäävad passiivseks kui aktiivne ebaõnnestub
- kaks kontrolleriit, igaüks vahemäluga, RAID-kontrolleriga, kettaühendusliidestega
- väike ja keskmise suurusega ettevõtetele



Liideste kordamine



- IDE/ATA
 - Ühendab omavahel emaplaadi ning salvestusvahendiseadme
 - Võimaldab master-slave-i (määratakse sildkontakti ehk *jumper*'i abil) kaudu ühendada 2 seadet ühe kaabli külge
- SATA
 - “Kuum-ühendatavad” ehk seadme saab lisada serveri külge masina töötamise ajal
 - Võimaldab ühendada ainult ühe seadme korruga

Liideste kordamine II



- SCSI
 - Ühendab salvestusseadme serveriga ning vahendab infot nende vahel
 - Näiteks kõvakettad; lindiseadmed:printerid jms;
 - Enimlevinud serverite juures
- SAS
 - Edasiarendus SCSI-st
 - Kiirem andmevahetus, paindlikum, rohkem eriseadmeid ühendav



Direct Attached Storage



- Tehnoloogia, kus salvestusvahendid ühenduvad otse serveri külge
- Sisemine DAS
 - Kasutatavad liidesed: IDE; SATA; paralleelne SCSI;
 - Näiteks Serial ATA liidesega kõvaketas personaalarvuti küljes
- Väline DAS
 - Kasutatavad liidesed: Serial SCSI; Fiiberoptiline kaabel
 - Näiteks FBO kaabliga ühendatud kettad serveri küljes



Direct Attached Storage II

- Plussid:
 - Ideaalne kohalikuks andmetöötluks/salvestuseks
 - Kerge seadistada
 - Lihtne konfiguratsioon
 - Madal üldhind
 - “Lollikindel” ehk usaldusväärne



Direct Attached Storage II

- Miinused:
 - Füüsilised piirangud
 - Liideste arv
 - Distsants
 - Kettaste arv
 - Piiratud kettajagamise võimalus
 - Kettapinna jagamine
 - Üle- või alarakendatud kettapinnad
 - Hooldustööde põhjustatud seisakud



Küsimused? Tänan tähelepanu eest!



IT KOLLEDŽ
TALLINNA TEHNIKAÜLIKOOL



TTÜ IT KOLLEDŽ

Raja 4C, 12616 Tallinn

tel +372 628 5800

info@itcollege.ee

<http://www.itcollege.ee/>