

**Tietokannan hallinta kertaus**

- TKHJ:n pääkomponentit
  - metadata TKHJ:ssä
- Tiedostojen käsittely
  - puskurien rooli tiedostokäsittelyssä
  - levymuistin rakenne ja käsittely
  - mistä tekijöistä haku aika muodostuu
    - jonotus – jos useita samanaikaisia levyn käyttäjiä – vaikeasti arvioitava
    - hakuvarren kohdistus – luokkaa 2-10 ms
    - pyörähdysviive – luokkaa 2- 10 ms
    - siirtoaika – luokkaa 0.1-0.5 ms/sivu, riippuu sivukoosta ja pyörimisnopeudesta
  - satunnaisen sivun hakemisen ja peräkkäishaun ero
  - RAID levyjen idea ja hyödyt

**Tietokannan hallinta kertaus**

- Talletusrakenteet
  - sivun rakenne
    - mistä ylimäärä koostuu -tietuehakemisto
  - tietueen sisäinen rakenne – kiinteä/vaihteleva
- Tiedostorakenteet
  - rakenteiden arviointi
    - levyhaut / todellinen suoritus aika
  - kasarakenne
  - järjestetty peräkkäisrakenne
  - hajautusrakenne (soluhakemistolla/ ilman soluhakemistoa)

**Tietokannan hallinta kertaus**

- Tiedostorakenteet
  - hakemistoratkaisut
    - tiheä/harva – mikä on ero
    - järjestävä (primary) / oheishakemisto (secondary)
    - hakemistotietueen sisältö
      - osoite : tietueosoite
      - osoite: tietueen pääavain (epäsuora osoittaminen, käytössä, jos tietueen sijainti voi vaihtua)
    - harvat järjestävät hakemistot
      - ISAM
      - B+puu
        - » tasapainotus

**Tietokannan hallinta kertaus**

Mikä ero on sillä, että taulu toteutetaan B+ -puuta tai taulun indeksi toteutetaan B+ -puuna.

**Tietokannan hallinta kertaus**

- B+ -puu
  - lehtisolmut sisältävät datatietueita järjestettynä peräkkäisrakenteena (solmut linkitetty)
  - indeksisolmut sisältävät erottimia ja osoittimia alipuihin
  - kaikki solmut juurta lukuun ottamatta ovat aina vähintään puoliksi täynnä
  - rakenne kasvaa dynaamisesti: kun solmuun ei enää mahdu uutta tietueita, solmun sisältö jaetaan kahtia ja nostetaan solmujen välinen erotin seuraavalle tasolle ylöspäin sillä hakureitillä, jota pitkin alkuperäiseen solmuun päädyttiin

**Tietokannan hallinta kertaus**

**Tietokannan hallinta kertaus**

- Tiedostorakenteet
  - dynaaminen hajautus
- Kyselyn käsittely
  - vaiheet
  - optimoijan tarkoitus ja optimoijatyytit
  - relaatioalgebran operaatioiden toteutustavat
    - ulkoinen järjestäminen
    - valinta
    - projekio
    - liitos

**Tietokannan hallinta kertaus**

- Kyselyn käsittely
  - optimoijan toimintaperiaate
    - kyselypuun muokkauksen idea
    - kustannuslaskentaoptimoinnin idea
      - tilastoaineisto
      - lokaali optimointi

**Tietokannan hallinta kertaus**

kurssi(kkoodi, knimi, opintoviikot...) 2000 a' n.300B, 10/sivu => 200 sivua  
 opiskelija(opnro, nimi, paa\_aine, aloitusvuosi...) 32000 a' n 400B, 8/sivu => 4000 sivua  
 suoritus(opnro->opiskelija, kkoodi->kurssi, vuosi, lukukausi, arvosana...) 1 500 000 a' 200B, 15/sivu => 100 000 sivua  
 suorituksilla indeksi kurssikoodin ja opiskelijanumeron mukaan erikseen, muilla pääavaimen

Vuonna 2002 aloittaneiden tietojenkäsittelytieteen opiskelijoiden opintosuoritusten kokonaismäärä syksyllä 2004

```

select sum(opintoviikot)
from kurssi, opiskelija, suoritus
where kurssi.kkoodi=suoritus.kkoodi and
suoritus.opnro=opiskelija.opnro and
opiskelija.aloitusvuosi=2002 and opiskelija.paa_aine='TKT' and
suoritus.vuosi=2004
    
```

**Tietokannan hallinta kertaus**

Lähtökohta 'suoraan' relaatioalgebraksi muunneltuna.

```

sum
├── s(kurssi.kkoodi=suoritus.kkoodi and suoritus.opnro=opiskelija.opnro and
│   opiskelija.aloitusvuosi=2002 and opiskelija.paa_aine='TKT' and
│   suoritus.vuosi=2004)
└── x
    ├── x
    │   ├── kurssi
    │   └── opiskelija
    └── suoritus
    
```

**Tietokannan hallinta kertaus**

valinnat alas, ristitulot liittöksi.

```

sum 400
├── j(kurssi.kkoodi=suoritus.kkoodi) 400
│   ├── j(suoritus.opnro=opiskelija.opnro) 400
│   │   ├── s(paa_aine='TKT') 200
│   │   └── s(aloitusvuosi=2002) 3000
│   └── suoritus 60000
│       └── s(suoritus.vuosi=2004)
    └── kurssi
    
```

arvio

ei ehkä paras mahdollinen, edellyttää koko tiedoston läpikäyntiä

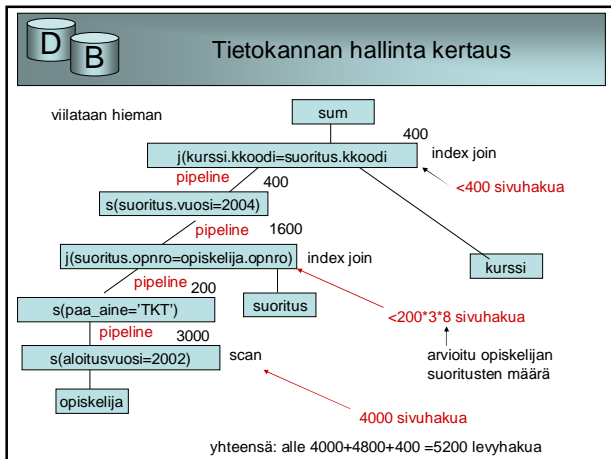
**Tietokannan hallinta kertaus**

viilataan hieman

```

sum 400
├── j(kurssi.kkoodi=suoritus.kkoodi) 400
│   ├── s(suoritus.vuosi=2004) 1600
│   │   ├── j(suoritus.opnro=opiskelija.opnro) 200
│   │   │   ├── s(paa_aine='TKT') 200
│   │   │   └── suoritus
│   │   └── s(aloitusvuosi=2002) 3000
│   └── suoritus
    └── kurssi
    
```

pienempi välitulos ja tehokkaammin tehtävissä



**Tietokannan hallinta kertaus**

- Transaktionhallinta
  - mikä on transaktio, vaatimukset transaktiolle (ACID)
  - transaktion tilat
  - ajoitusjärjestys
  - loki

**Tietokannan hallinta kertaus**

- Transaktionhallinta
  - sitoutumiskäytäntö ja loki
  - lokiperustainen elvytys
    - tarkistuspiste
    - peruutettavat ja uudelleensuoritettavat
    - pageLSN
  - samanaikaisuuden hallinta
    - sarjallinen suoritusjärjestys
    - sarjallisuus
    - eristyvyysanomaliat
    - SQL:n eristyvyystasot

**Tietokannan hallinta kertaus**

- samanaikaisuuden hallinta
- lukot
  - tyypit
  - 2PL (ankara)
  - lukkiutuma
  - Lukkiutumien hallinnan periaatteita