



## Kyselyjen käsittely ja optimointi

- Kyselyn käsittelyn vaiheet:
  - TKHJ ottaa vastaan kyselyn asiakasohjelmalta
  - Kysely selataan ja jäsennetään
    - tarkistetaan kyselyn rakenteellinen oikeellisuus
  - Jäsennetty kysely muunnetaan relaatiolausekkeiksi ja tallennetaan sisäisessä esitysmuodossa kyselypuuna
  - Laaditaan kyselyn toteutus suunnitelma
    - suunnitelman laadintaan liittyy valintaa eri toteutusvaihtoehtojen välillä 'optimointia'
  - Suoritetaan kysely suunnitelman mukaisesti
  - Toimitetaan tulokset asiakasohjelmalle

1



## Kyselyjen käsittely ja optimointi

- Miksi optimoidaan
  - SQL-kysely määrittää halutun tuloksen, ei sitä miten tulos muodostetaan (deklaraatiivinen kyselykieli)
    - useimmissa oliokannoissa kysely etenee navigoimalla proseduraalisesti – eli ohjelmoija kontrolloi kyselyn toteutustapaa, silloin ohjelmoijan pitää osata valita oikea tapa
  - Valitut tiedostorakenteet vaikuttavat siihen, miten operaatio kannattaa suorittaa
  - Taulujen koko ja arvojen jakautumat muuttuvat jatkuvasti, tämä saattaa vaikuttaa siihen, miten kysely kannattaa toteuttaa

2



## Kyselyjen käsittely ja optimointi

- Kyselyn optimointi ei ole todellista matemaattista optimointia
  - tarkoitus ei ole löytää parasta vaan riittävän hyvä
  - parhaan etsiminen on laskennallisesti raskasta
  - taustatietoja ei yleensä ole riittävästi parhaan valintaan
- Myös SQL:ssä käyttäjä voi jonkin verran vaikuttaa kyselyn suoritustapaan
  - monissa SQL-murteissa käyttäjä voi liittää kyselyyn vihjeitä siitä, miten kysely pitäisi käsitellä
  - kyselyn esittämistapa voi myös vaikuttaa,
    - esim. jos taululla on indeksi sarakkeen Nimi perusteella Nimi like 'Vir%' käyttää indeksia, mutta upper(Nimi) like 'VIR%' ei käytä

3



## Kyselyjen käsittely ja optimointi

- Kyselyn optimoinnin periaatteet
  - Heuristinen optimointi (sääntöpohjainen optimointi)
    - käytössä on joukko sääntöjä, joiden mukaisesti operaatiot järjestetään tai suoritustavat valitaan
      - 'valinnat ennen liitoksia'
      - 'käytä indeksia aina kun mahdollista'
      - jne
    - säännöt ovat yleiskäyttöisiä
    - suoritustapa ei välttämättä ole optimaalinen

4



## Kyselyjen käsittely ja optimointi

- Kyselyn optimoinnin periaatteet
  - Kustannuslaskentaoptimointi (tilastollinen optimointi)
    - laaditaan vaihtoehtoisia suunnitelmia, esim. sääntöpohjalta
    - lasketaan vaihtoehdoille kustannukset
    - valitaan edullisin
    - edellyttää tilastoaineistoa ja sen uudistamista ajoittain

5



## Kyselyjen käsittely ja optimointi

- SQL-kyselyn voi ajatella muodostuvat joukosta kyselylohkoja (query block) (select –from – where lohko)

```

select lname, fname
from employee
where salary >=
(select max(salary)
from employee
where dno=5);

```

kaksi kyselylohkoa

6



## Kyselyjen käsittely ja optimointi

- Kyselylohkot muunnetaan relaatioalgebran lausekkeiksi
- Lohko 1:  $\pi_{\text{name, fname}}(\sigma_{\text{salary} > C}(\text{employee}))$ ,
- Lohko 2:  $F_{\text{max salary}}(\sigma_{\text{dno}=5}(\text{employee}))$
- $F$  on laajennetun relaatioalgebran funktio-operaatio, tarvitaan esim. yhteenvetofunktioissa, ryhmittelyssä jne.

7



## Kyselyjen käsittely ja optimointi

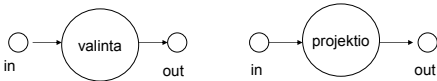
- Relaatioalgebran operaatioiden suoritukseen on useita vaihtoehtoja
- Vaihtoehtojen tehokkuus vaihtelee
- Tehokkuuden ensisijaisena kriteerinä tarvittavien levyoperaatioiden lukumäärä
  - luettaessa relaatioiden osia levyllä
  - kirjoitettaessa operaation tulosta levyllä
    - tulos on joko välitulos tai lopputulos

8



## Kyselyjen käsittely ja optimointi

- Välitulosten kirjoittamisen (ja edelleen lukemisen) kustannus voidaan usein jättää huomiotta, jos tkhj käyttää putkitekniikkaa: (pipelining) operaation tulosjono ohjataan toisen syöttöjonoiksi, jolloin levyllä kirjoitusta ei tapahdu

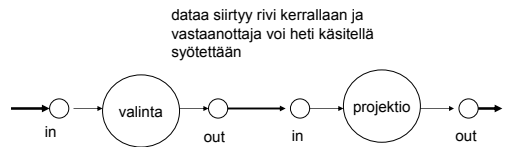


9



## Kyselyjen käsittely ja optimointi

- Operaation tulosjono ohjataan toisen syöttöjonoiksi, jolloin levyllä kirjoitusta ei tapahdu



10



## Tiedoston järjestäminen

- Monet tietokantaoperaatiot (muutkin kuin order by) edellyttävät tietueiden järjestämistä
  - group by,
  - relaatiomallin mukainen projektio (select distinct)
  - yhdiste
  - erotus
- järjestäminen voi luovuttaa ensimmäisen tulosrivin eteenpäin vastaa vastaanotettuaan viimeisen
- Jos käsiteltävät rivijoukot ovat isoja tarvitaan ulkoista järjestämistä (external sorting)

11



## Tiedoston järjestäminen

- Ulkoisessa järjestämisessä käytetään yleisesti ns. lomituserjestämistä (merge sort):
  - Tiedosto luetaan osina keskusmuistiin (osien koon määrää järjestämiseen käytettävissä olevan keskusmuistin määrä)
  - Järjestetään kukin osatiedosto keskusmuistissa (esim. quicksort)
  - Kirjoitetaan järjestetty osatiedosto (run, jono) levyllä.
  - Tehdään lomituserroksia tarpeellinen määrä, jotta jäljellä on vain yksi järjestetty osatiedosto

12



## Tiedoston järjestäminen

- Lomituskierros:
  - Luetaan tietueita  $m$  ( $m \geq 2$ ) järjestetystä jonosta (osatieidostosta) - lukua  $m$  kutsutaan lomitusasteeksi
  - Lomitetaan tiedostot eli kirjoitetaan yhtä järjestettyä tulostiedostoa
    - kirjoitetaan avaimeltaan pienin jonojen vuorossa olevista tietueista tulosjonoon ja edetään tämän jonossa
  - Järjestettyjen osatieidostojen pituudet kasvavat kierroksella  $m$ -kertaiseksi.
  - Jos jonoja (osatieidostoja) oli alunperin  $n$ , tarvitaan lomituskierroksia  $\text{roof}(\log_m n)$

13



## Tiedoston järjestäminen

- Oletetaan, että järjestämiseen on käytettävissä puskuritilaa  $128 \times 4\text{KB}$  (vähän yli 0.5MB). Tiedoston koko olkoon 2048 sivua = 8MB
  - Alkuperäiseen järjestämiseen voidaan käyttää 128 syöttöpuskuri ja 1 tulospuskuri, joten saadaan 128 sivun pituisia järjestettyjä jonoja 16 kappaletta.
  - Näiden järjestäminen hoituu 1 lomituskierroksella - 16 asteinen lomitus.
- Yhdellä lomituskierroksella selvittäisiin tällä puskurimäärällä  $128 \times 128 \times 4\text{KB}$  (64MB) kokoisesta tiedostosta

14



## Tiedoston järjestäminen

- Kullakin kierroksella kaikki sivut luetaan ja kirjoitetaan.
- Olkoon tiedoston koko  $N$  sivua ja muistitilaa käytettävissä  $B$  sivua. Levyoperaatioita tarvitaan tällöin  $2N + 2N \cdot \text{roof}(\log_{B-1} \text{roof}(N/(B-1)))$ 
  - levyoperaatiot peräkkäiskäsittelyä
  - lomituksesta on olemassa erilaisia muunnelmia esim. muistitilan käytön suhteen,
    - esim. kaksoispuskuroinnilla voidaan vähentää jonotusaikoja (toisiin puskuureihin luetaan kun toisia käsitellään), mutta  $B$  puolittuu ja levyoperaatioiden määrä kasvaa vaikka kokonaisaika väheneekin

15



## Valintaoperaation toteutus

- Valinta:
  - $\text{select } * \text{ from } R \text{ where } A \text{ op value}$
  - $\sigma_{A \text{ op value}}(R)$
- Valinnan vaihtoehdot riippuvat
  - valintaehdosta ja
  - käytössä olevista tiedostorakenteista

16



## Valintaoperaation toteutus

- Tiedoston läpikäynti (scan)
  - soveltuu aina
  - käytettävä jos
    - sarake  $A$  ei ole mukana missään indeksissä
    - eikä tiedosto ole sarakkeen  $A$ :n perusteella järjestetty
    - sarake  $A$  on jonkin funktion argumenttina, esim.  $\text{upper}(A)$ , myös  $\text{substring}(A, 1, x)$ ?
    - voi hyödyntää peräkkäiskäsittelyä

17



## Valintaoperaation toteutus

- Binäärihaku (binary search)
  - soveltuu, jos tiedosto on järjestetty sarakkeen  $A$  perusteella
  - etsitään ensimmäinen ehdon täyttävä, josta voidaan jatkaa eteenpäin
  - voi hyödyntää peräkkäishakua vasta ensimmäisen ehdon täyttävän löytymisen jälkeen

18



## Valintaoperaation toteutus

- **Indeksihaku (index scan)**
  - soveltuu, jos käytössä on hakusarakeeseen perustuva indeksi
    - pelkästään hakusarake
    - hakusarakealkuinen indeksointiperusta
    - hakusarake mukana indeksissä
      - tämäkin saattaa olla käyttökelpoinen
      - edellyttää koko hakemiston läpikäyntiä, mutta toimii jos hakemisto on pieni suhteessa tiedostoon ja hakuehto on hyvin rajaava (selective) = jäljittää pienen tietuejoukon,

19



## Valintaoperaation toteutus

- Valintaehto muodostuu 'ja' ( $\wedge$ ) ja 'tai' ( $\vee$ ) operaatioilla yhdistetyistä alkeisehdoista.
- Ehdon konjunkttiivinen normaalimuoto (conjunctive normal form) tarkoittaa, että ehto esitetään tai operaatioilla yhdistettyinä konjunktioina:
 
$$(e_1 \wedge e_2 \wedge \dots \wedge e_n) \vee (e_m \wedge \dots \wedge e_p) \vee \dots$$

20



## Valintaoperaation toteutus

- Konjunktioita ( $e_1 \wedge e_2 \wedge \dots \wedge e_n$ ) evaluoitaessa ei riitä, että jokin ehdoista täyttyy vaan kaikkien pitää täyttyä. Jos joillakin ehdoissa olevista sarakeista on indeksit näitä pyritään käyttämään:
  - valitsemalla eniten rajaava indeksi, ja tutkimalla sen jälkeen tietueista muut ehdot, tai
  - muodostamalla hakemistojen leikkaus ja tutkimalla leikkausjoukon tietueista ne ehdot, joita ei voida ratkaista indeksien perusteella
    - tämä ratkaisu edellyttää tietueosoitteita
    - leikkausjoukon muodostamiseksi jäljitetyt osoitelistat järjestetään osoitteen perusteella ja lomitetaan
      - Huom: tietueosoitteessa alussa sivuosoite.

21



## Valintaoperaation toteutus

- Olkoon taululle työntekijä määritelty indeksi sekä sukunimen että kotiosoitteen perusteella. Tällöin
 

```
select * from työntekijä
  where sukunimi='Löpönen' and
        kotiosoite='Teollisuuskatu 23 B446, Helsinki'
```
- voitaisiin toteuttaa
  - valitsemalla rajaavampi indeksi (kotiosoite), ja tutkimalla tietueet
    - keskimäärin vähän yli 1 (olisi tässä parempi vaihtoehto koska osoite on hyvin rajaava), tai
    - hakemalla kummankin indeksin perusteella (keskimäärin 30 ja 1+) ja muodostamalla näiden leikkaus. On tässä tapauksessa huonompi, mutta jos kumpikaan indeksi ei ole kovin tiukasti rajaava (vaikkapa kurssikoodi ja opiskelijanumero ilmoittautumistiedoissa), niin niiden leikkaus voi olla silti hyvin pieni)

22



## Valintaoperaation toteutus

- Disjunkttiivisen ehdon  $e_1 \vee \dots \vee e_n$  tulos vastaa erillisillä ehdoilla  $e_1 \dots e_n$  saatujen rivijoukkojen yhdistettä.
- Koko tiedosto on käytävä läpi, jos yksikin ehdoista perustuu sarakeeseen, jolle ei ole indeksia

23



## Projektion toteutus

- Projektioilla on kaksi tehtävää
  - sarakkeiden poiminta ja
  - toistuvien vastausrivien karsinta
    - SQL-kyselyissä toistuvia rivejä ei välttämättä haluta karsia (kyselyssä ei DISTINCT-määrettä)
  - Sarakkeiden poiminta on suoraviivaista
    - otetaan tietueesta vain ne sarakkeet, jotka on lueteltu projisoitaviksi
    - jos kaikki luetellut sarakkeet löytyvät hakemistosta ei varsinaisia datarivejä tarvitse lukea lainkaan.

24



## Projektion toteutus

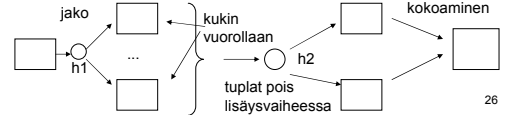
- Toistuvien vastausrivien (duplikaattien) karsinta on hieman hankalampaa.
  - Tietueet voidaan järjestää käyttäen järjestysavaimena koko tietuetta (jos tulokselle on määritelty järjestys käytetään järjestysperustaa avaimen alussa). Näin duplikaatit saadaan peräkkäin ja ne voidaan helposti karsia yhdellä läpikäynnillä (tai lomituksen viimeisellä kierroksella)
    - jos vastaus on pieni, riittää keskusmuisti-järjestäminen, mutta ensimmäinen tulosrivi voidaan toimittaa eteenpäin vasta kun viimeinen syöttörivi on saatu.
    - isot vastaukset edellyttävät ulkoista järjestämistä.

25



## Projektion toteutus

- Toistuvien rivien karsinta
  - Tietueet voidaan myös hajauttaa koko tietueesta lasketun hajautusavaimen perusteella. Tällöin duplikaatit sijoittuvat samaan soluun ja ne voidaan lisäysvaiheessa karsia. Jos osoiteavaruus on riittävän iso, tulee samaan soluun vähän tietueita ja vähän tutkittavaa.
  - Tyypillisesti käytetään kahta peräkkäistä hajautusta



26



## Projektion toteutus

- Toistuvien rivien karsinta
  - Jos projektiio sisältää taulun avaimen ei tulosrivejä tarvitse lajitella tai hajauttaa, tiedetään että toistuvia rivejä ei ole.

27



## Liitoksen toteutus

- Liitos on tärkeimpiä ja yleisimpiä relaatioalgebran operaatioita
- Yleisin liitos on taulun pääavaimen ja siihen viittaavan viiteavaimen yhtäsuuruuteen perustuva liitos.
- Seuraavassa käytetään esimerkkinä taulujen opiskelija (500 sivua, kussakin 20 tietuetta) ja suoritus (2000 sivua, kussakin 50 tietuetta) liitosta opiskelijanumeron perusteella.

28



## Liitoksen toteutus

- Liitos voidaan määrittellä ristitulon ja valinnan yhdistelmäksi. Yksinkertaista olisi muodostaa ristitulo ja tehdä sen jälkeen valinta. Ratkaisu on kuitenkin tehoton.
- Liitoksen toteutukseen on tarjolla useita tekniikoita. Päättyypit:
  - sisäkkäiset silmukat (nested loop join)
  - hakemistoliiitos (index nested loop join)
  - lomitusliitos (sort merge join)
  - hajautusliitos (hash join)

29



## Liitoksen toteutus

- Tarkastellaan taulujen R ja S liitosta ehdolla  $R.x=S.y$
- Sisäkkäisten silmukoiden tekniikassa perusidea on seuraava:
 

```
foreach tuple r in R do
  foreach tuple s in S do
    if r.x==s.y then add {r,s} to result
```
- eli käydään läpi ulompi taulu ja jokaisen rivin kohdalla käydään läpi sisempi taulu

30



## Liitoksen toteutus

- Olkoon R:n sivumäärä pages(R) ja S:n sivumäärä pages(S) ja vastaavasti sivukoot tietueina psize(R) ja psize(S). Tällöin edellä kuvatun yksinkertaisen algoritmin kustannus levyhakuina olisi
  - $op = pages(R) + psize(R) * pages(R) * pages(S) +$  tuloksen kirjoitus
- Esim:
  - jos R= opiskelija ja S= suoritus, niin
    - operaatioita =  $500 + 500 * 20 * 2000 = 20\ 000\ 500 +$  tulos.
  - jos R= suoritus ja S= opiskelija, niin
    - operaatioita =  $2000 + 2000 * 50 * 500 = 50\ 002\ 000 +$  tulos
- Hidasta

31



## Liitoksen toteutus

- Parannetaan algoritmia hieman ja käsitellään koko sivu kerralla
  - Siis jokaista R:n sivua kohti käydään S kertaalleen läpi ja verrataan kaikkia R:n sivulla olevia tietueita vuorossa olevan S:n sivun tietueisiin.
  - operaatioita  $pages(R) + pages(R) * pages(S) +$  tuloksen kirjoitus
  - R= opiskelija, S= suoritus:  $500 + 500 * 2000 = 1\ 000\ 500$
  - R= suoritus, S= opiskelija:  $2000 + 2000 * 500 = 1\ 002\ 000$
  - satunnaishakuajalla 10ms laskettuna tämä veisi 10 005 s
- Mitä enemmän R:stä käsitellään kerralla sitä vähemmän kierroksia tarvitaan, niinpä parannetaan vielä

32



## Liitoksen toteutus

- Jaksottaiset sisäkkäiset silmukat (Block nested loops)
  - versio joka yleensä tarjolla tkhj:ssä
  - Oletetaan, että liitoksen suorittamiseen on käytettävissä B syöttöpuskuria. Varataan näistä B-1 ulomalle taululle R ja 1 sisemmälle taululle S.
  - Käydään yhdellä vertailukierroksella läpi kaikki ulomman taulun puskurit ja verrataan tietueita sisemmän taulun vuorossa olevan sivun tietueisiin.
  - Sisempi taulu S joudutaan käymään läpi  $roof(pages(R)/(B-1))$  kertaa
  - Vaikka levyoperaatioiden määrä on tässä mallissa pieni, paras käsittelyteho saavutetaan, jos puskurikapasiteetti jaetaan tasan kummankin liitostiedoston kesken – levyhakuja tulee enemmän, mutta voidaan hyödyntää peräkkäisyyttä

33



## Liitoksen toteutus

- Esimerkki:
  - Varataan liitoksen toteutukseen puskuritilaa 101 syöttöpuskurille (404KB)
  - R= opiskelija, S= suoritus
    - operaatioita =  $500 + (500/100) * 2000 +$  tulos = 10500 + tulos
    - satunnaishakuajalla 10 ms laskettuna tämä olisi 105 s
      - R voidaan kuitenkin lukea 100 sivun pätkinä, jolloin aikaa oikeastaan kuluisi  $5 * (satunnaishaku aika + 100 sivun siirto) + 5 * 2000 * (satunnaishaku aika + 1 sivun siirto) =$
      - $5 * (10ms + 100 * 0.2 ms) + 10000 * (10 ms + 0.2 ms) = 150 ms + 102000ms =$  noin 102 s
  - R= suoritus, S= opiskelija
    - operaatioita =  $2000 + (2000/100) * 500 = 12000$

34



## Liitoksen toteutus

- Jos jaettaisiin puskuritila puoliksi R:lle ja S:lle ja R = opiskelija ja S= suoritus niin
- haku aika:
  - $10 * (10ms + 50 * 0.2 ms) +$
  - $10 * 2000 / 50 * (10 ms + 50 * 0.2 ms) =$
  - $10 * 11ms + 10 * 40 * (11ms) = 110 ms + 4400ms = 4.51 s$
- odotusajat lisäävät kokonaisaikaa

35



## Liitoksen toteutus

- Hakemistoliiotus (index nested loop)
  - Jos toisella taululla on hakemisto liitossarakkeeseen perustuen tehdään tästä sisempi taulu

```
foreach tuple r in R do
  foreach tuple s in index_search(S,r.x) do
    add {r,s} to result
```

- Jos hakemistona on rakenteeltaan B+ -puu rakenteinen oheishakemisto, tarvitaan indeksin kautta hakuun tyypillisesti 1-4 levyhakua isollakin tiedostolla.

36



## Liitoksen toteutus

- Hakujen kokonaismäärä:  
 $\text{pages}(R) + \text{pages}(R) * \text{psize}(R) * (\text{height}(S) + \text{matches}(S))$
- $\text{height}(S)$ :
  - levyhakujen määrä hakemiston kautta haettaessa, mukaan mahdolliset ylivuotoketjut ja oletukset siitä, mitä säilyy keskusmuistissa
- $\text{matches}(S)$ :
  - montako keskimäärin löytyy yhtä olumpaa kohti. Jos kyseessä on B+ -puuna toteutettu oheishakemisto, on jokaista tietuetta kohti oma hakemistotietue (jos tietueosoite on mukana avaimessa nämä ovat osoitejärjestyksessä), yksi haku / osuma riittää

37



## Liitoksen toteutus

- Esimerkki:
  - R= opiskelija S=suoritus
    - levyhakuja  $500 + 500 * 20 * (2 + 10) = 120\ 500$
    - oletetaan tehollinen hakukorkeus 2, opiskelijalla 10 suoritusta
    - kaikki S:n haut satunnaishakuja -> noin 1200 s
  - R=suoritus, S=opiskelija
    - levyhakuja  $2000 + 2000 * 50 * (2 + 1) = 302\ 000$
- Huomaamme, että kokonaan läpikäytäväksi (ulommasi) kannattaa valita pienempi tauluista

38



## Liitoksen toteutus

- Lomitusliitos (sort merge join)
  - soveltuu yhtäsuuruusliitokseen
  - Lomitusliitoksessa taulut järjestetään liitossarakkeen perusteella
  - Kun rivit ovat järjestyksessä on liitos muodostettavissa yhdellä läpikäynnillä
    - Rivit ovat järjestyksessä esim., jos tiedoston toteutusrakenteena on ISAM tai B+ -puu
    - Jos tiedostolla on B+ -puurakenteinen oheishakemisto, tietueet saadaan myös järjestyksessä (tosin hitaammin)

39



## Liitoksen toteutus

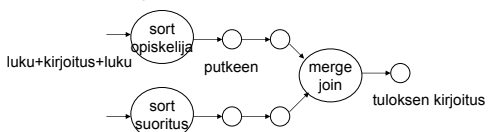
- Elleivät rivit ole järjestyksessä, ne pitää ensin järjestää.
- Järjestämisen levyoperaatioiden määrään oli  $2N + 2N * \text{roof}(\log_{B-1} \text{roof}(N/(B-1)))$
- tuloksen muodostus hoituu yhdellä läpikäynnillä.

40



## Liitoksen toteutus

- Esim.
  - Olkoon B=101, Kumpikin esimerkkitaulu vaatii jonojen muodostuksen ja yhden lomituskierron sekä yhdistämisen liitoksen aikaansaamiseksi,
  - Oletetaan että lomituksen tulos putkitetaan liitokseen, tällöin operaatioita on 3\* sivulukumäärä kummallakin



41



## Liitoksen toteutus

- eli esimerkkitapauksessa levyoperaatioita
  - $3 * 500 + 3 * 2000 + \text{tulos} = 7500 + \text{tulos}$
  - jonojen muodostus sujuu peräkkäiskäsitelyinä
- lomituksessa voidaan joutua odottamaan,
- Oletetaan, että jonoja luetaan lomituksessa 10 ja 5 sivun jaksoissa (käytetään kaikki puskurit)
  - opiskelija jonojen teko:  $2 * 5 * (10 \text{ms} + 100 * 0.2 \text{ms}) +$
  - suoritus jonojen teko:  $2 * 20 * (10 \text{ms} + 100 * 0.2 \text{ms}) +$
  - opiskelija jonojen luku:  $500 / 10 * (10 \text{ms} + 10 * 0.2 \text{ms}) +$
  - suoritus jonojen luku:  $2000 / 5 * (10 \text{ms} + 5 * 0.2 \text{ms}) =$
  - $300 \text{ms} + 1200 \text{ms} + 1500 \text{ms} + 8000 \text{ms} = 11 \text{s}$

42