

Laskuharjoitus 1. Nopeuksia ja kasvuja

Tehtävä 1.1.

Kysymys

Tehtävänäsi on siirtää ilmastomalliin liittyvää dataa 50 TB Otaniemestä Kumpulaan. Käytettävissäsi on yksi siirrettävä 8 TB:n kovalevy sekä polkupyörä. Matkaa on 10 km - mutta onneksi on kesäkuu. Oletetaan että kovalevyllä sekä kirjoitus- että lukunopeus on 200 MB/s.

Toinen vaihtoehto on datayhteys Otaniemen ja Kumpulan välillä.

- A. Kuinka nopea datayhteyden tulee vähintään olla, jotta koko datamäärän siirto tapahtuisi nopeammin datayhteydellä kuin polkupyörällä?
- B. Entä jos käytettävissä on auto ja kaksi kappaletta 30 TB:n puolijohdelevyjä (solid-state drive, SSD)? Arvioi realistinen luku- ja kirjoitusnopeus puolijohdelevylle.
- C. Entä jos data pitäisikin siirtää Macquarie-yliopistoon Sydneyn esikaupunkialueella? Käytettävissä on edelleen B-kohdan puolijohdelevyt. Arvioi realistinen matka-aika, kun oletetaan, että aikaan lasketaan myös paluumatka.
- D. Arvioi paljonko datan esittämiseen tarvittaisiin tiheään painettuja kaksipuoleisia A4-arkkeja, jos data esitettäisiin 8 bitin merkeillä, fonttikoolla 8.

Ratkaisu

Tehdään seuraavat lisäoletukset (tenttivastauksessa ja muutenkin on syytä selkeästi ilmaista mitä mahdollisia lisäoletuksia on tehty):

- Reipas opiskelija polkee polkupyörällä 20 km/h
- Autolla kaupunkioissa pääsee keskimäärin 30 km/h
- Jos kovalevyjä on käytössä useampia, niitä voidaan lukea ja kirjoittaa rinnakkain (mutta tämä on siis oletus eikä välttämättä tarkkaan päde reaali maailmassa).
- Terat (T) on esitetty 10-järjestelmän muodossa.

Huomatkaa lisäksi, että tehtävänannossa tiedonsiirtonopeus on MB/s eli megatavua sekunnissa (ei siis bittiiä sekunnissa).

A-kohdassa matka-ajaksi tulee siten $(13 \cdot 10 \text{ km} = 150 \text{ km}) / (20 \text{ km/h}) = 390$ minuuttia. Kovalevyille kirjoittamiseen ja lukemiseen menee yhteensä aikaa $2 \cdot 50 \text{ TB} / 200 \text{ MB/s} = 8333$ min. Yhteensä aikaa menee 8723 min. Tämä vastaa keskinopeutta

$$(8 \cdot 50\,000\,000 \text{ Mbit}) / (60 \cdot 8723 \text{ s}) = \underline{766 \text{ Mbit/s.}}$$

Kerroin kahdeksan (8) kaavan alussa johtuu siis siitä, että tässä täytyy tehdä muunnos tavuista biteiksi.

B-kohdassa

Kirjoitus ja luku: $100 \text{ TB} / 2 \text{ GB/s} = 13 \text{ h } 53 \text{ min}$

1 automatka: $10 \text{ km} / 30 \text{ km/h} = 20 \text{ min}$

Yhteensä $14 \text{ h } 13 \text{ min}$

Siirtonopeus $50\,000\,000 \text{ MB} / 51180 \text{ s} = 976,94 \text{ MB/s} = 7816 \text{ Mbit/s}$

C-kohdassa

Kirjoitus ja luku: $100 \text{ TB} / 2 \text{ GB/s} = 13 \text{ h } 53 \text{ min}$

Edestakainen matka: $2 * (1 \text{ h metro+juna} + 23 \text{ h lento} + 1 \text{ h juna+bussi}) = 50 \text{ h}$

Yhteensä $63 \text{ h } 53 \text{ min}$

Siirtonopeus $50\,000\,000 \text{ MB} / 229980 \text{ s} = 217,41 \text{ MB/s} = 1739 \text{ Mbit/s}$

D-kohdassa

$2 \text{ puolta} * 86 \text{ riviä} * 112 \text{ merkkiä} = 19264 \text{ merkkiä arkia kohti}$

$50\,000\,000\,000\,000 \text{ B} / 19264 \text{ B} = 2\,595\,514\,951 \text{ arkia}$

Siis melkein $13\,000$ tonnia paperia tai paperiarkit päällekkäin pinossa melkein 260 km

Tehtävä 1.2.

Kysymys

Vuonna 2010 siirretyn mobiilidatan määrä Suomessa oli n. 350 MB/kk liittymää kohti. Vuonna 2020 vastaava lukema oli noin $22,4 \text{ GB/kk}$. Milloin mobiilidatan määrä ylittää 100 GB/kk , jos oletetaan että

A. kasvu on lineaarista ajan suhteen eli kasvaa saman määrän vuosittain, tai

B. kasvu on eksponentiaalista ajan suhteen eli kasvaa yhtä monta prosenttia joka vuosi?

Laske lisäksi

C. Mikä on liittymän keskimääräinen siirtonopeus (yksikkönä Mbit/s) koko kuukauden yli, kun siirretyn datan määrä on 100 GB/kk .

Ratkaisu

A.

Kasvua on $22400 \text{ MB/kk} - 350 \text{ MB/kk} = 22\,050 \text{ MB/kk}$

eli kasvu on $2,205 \text{ GB/kk}$ vuotta kohti

Kasvua tarvitaan $100 \text{ GB/kk} - 22,4 \text{ GB/kk} = 78,6 \text{ GB/kk}$

eli tähän kuluu $78,6 / 2,205 = 35,2$ vuotta

joten 100 GB/kk lukemaan päädyttäisiin noin vuonna 2056.

B.

Kasvua on $(22400/350)^{1/10} = 1,516$ - kertaistasi vuodessa eli 52 % kasvu vuodessa.

$100000 / 22400 = 3,60$

Näistä kahdesta luvusta saadaan yhtälö, jossa x on aika vuosissa:

$$1,516^x = 3,60$$

josta saadaan: $x = \ln(3,60) / \ln(1,516) = 3,60$ vuotta

eli yli 100 GB/kk arvoon päästään vuonna 2024.

C.

$100 \text{ GB} = 100\,000 \text{ MB} = 800\,000 \text{ Mbit}$

$1 \text{ kk} = 30 \text{ p} = 720 \text{ h} = 43\,200 \text{ min} = 2\,592\,000 \text{ s}$

joten keskimääräinen nopeus = $800\,000 \text{ Mbit} / 2\,592\,000 \text{ s} = 0,308642 \text{ Mbit/s}$

Tehtävä 1.3.

Kysymys

Oletetaan, että signaali sisältää 8 erilaista merkkiä (a, b, c, d, e, f, g, h) siten että merkkien esiintymistodennäköisyydet ovat 55, 20, 10, 5, 4, 3, 2 ja 1 prosenttia.

- A. Kuinka monta bittiä tarvitaan keskimäärin 1000 merkin esittämiseen, kun jokainen merkki koodataan samalla määrällä bittejä?
- B. Kuinka monta bittiä tarvitaan keskimäärin 1000 merkin esittämiseen, kun koodit ovat a:sta h:n seuraavasti: 1, 01, 001, 0001, 00001, 000001, 0000001 ja 00000001. Huomaa, että uusi merkki alkaa tässä koodauksessa aina 1:sen jälkeen!

Ratkaisu

A.

Koska erilaisia merkkejä on 8, niiden esittämiseen tarvitaan 3 bittiä ($2^3 = 8$). Tuhannen merkin esittämiseen tarvitaan siten 3000 bittiä.

B.

Keskimäärin yhden merkin esittämiseen tarvitaan

$$1 \cdot 0,55 + 2 \cdot 0,20 + 3 \cdot 0,10 + 4 \cdot 0,05 + 5 \cdot 0,04 + 6 \cdot 0,03 + 7 \cdot 0,02 + 8 \cdot 0,01 = 2,05 \text{ bittiä.}$$

1000 merkin esittämiseen tarvitaan siis keskimäärin 2050 bittiä.

Huomioita

Tehtävä 1.1. Käytännössä kumpikaan malli tuskin toteutuu sellaisenaan (tosin Mooren laki, jota käsitellään kurssilla myöhemmin, on toiminut erittäin hyvin 50 vuoden ajan). Eksponentiaalinen kasvu on tyypillistä kasvun alkuvaiheessa, mutta jossain vaiheessa kasvu väistämättä hidastuu. Käännekohtia on hyvin vaikea ennustaa menneisyyttä kuvaavan datan perusteella. Lineaarinen kasvu on mahdollista väliaikaisesti silloin, kun kasvu alkaa tasaantua, mutta reaali maailmassa pitkäaikainen lineaarinen kasvu on harvinaista, kun kyseessä on jonkinlainen populaation koon määrän muutos.

C-kohdan tavoitteena on ensisijaisesti osoittaa, että 100 GB/kk ei vaadi mitään äärimmäisen suurta jatkuvaa datanopeutta. Arvoon 100 GB/kk pääsee, kun katsoo tavanomaista videokuvaa noin kaksi ja puoli tuntia päivässä.

Tehtävässä 1.3 koodausta voisi vielä hieman parantaa määrittelemällä h:n koodiksi "0000000", koska tässä tapauksessa tiedetään, että seitsemän nollan jälkeen ei voi tulla kuin 1! Eli tällöin 1000 merkkiä voitaisiin esittää keskimäärin 2040 bitillä. Tässä hyödynnetään siis vain yksittäisten merkkien erilaisia esiintymistiheyksiä, lisäksi tiedon pakkaamisessa voidaan hyödyntää erilaisia merkkiyhdistelmien esiintymistiheyksiä, jos tiheyksissä on eroja verrattuna satunnaisesti syntyviin merkkiyhdistelmiin.