

Vertaistietojenkäsittelyn (peer-to-peer-computing) sovelluksia Internetissä

Reima Suomi

Turun kauppakorkeakoulu

Tallenneteollisuus vaikeuksissa

Helsinki (27.08.2002) - Internetissä tapahtuva musiikin online-levitys syö ääniteteollisuuden pöydästä, kertoo PricewaterhouseCoopersin tutkimus: cd-levyjen myynti laski vuoden ensi puoliskolla seitsemän prosenttia. Recording Industry Association of American, eli RIAAn mielestä tämä merkitsi alan teollisuudelle 284 miljoonan dollarin myynnin menetyksiä.

Vuotta aiemmin cd-levyjen myynti laski 5,3 prosenttia.

Maanantaina julkistetun cd-myyntitutkimuksen lisäksi RIAA julkisti erillisen tutkimuksen Internet-käyttäjien musiikkitottumuksista, jonka mukaan useimmat 12-54-vuotiaat kuluttajat ostivat aiempaa vähemmän äänitteitä, koska he latasivat musiikkia koneilleen internetistä.

Aiempien ääniteteollisuudesta riippumattomien tutkimusten tulos on ollut se, että pääsy ilmaiseen nettimusiikkiin itse asiassa rohkaisee kuuntelijoita kokeilemaan uusien kappaleiden kuuntelemista ja ostamaan enemmän musiikkitalenteita.

Keskeiset sovellukset

- Sisällön jakaminen
- Resurssien jakaminen
- Ryhmätyösovellukset

Määritelmiä

- Vertaistietojenkäsittely, vertaisverkot

– *peer or peer-to-peer ...*

... *computing*

network

processing

entity

P2P

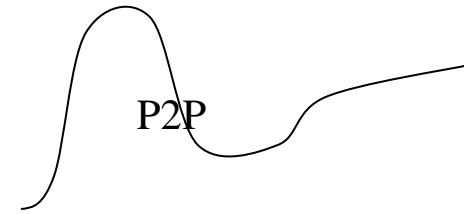
Ei asiakas-palvelin –jakoa

Jokainen voi olla sekä asiakas että palvelija

Määritelmiä jatkuu...

- *A type of network in which each workstation has equivalent capabilities and responsibilities* (www.webopedia.com)
- an umbrella covering five areas:
 - file sharing
 - distributed computing
 - web services
 - messaging
 - gaming
- P2P is nota a technology, it is a mindset (Jon Orwant)
- *the ability to share computing resources with other users or systems* (Gartner)
- *a class or applications that takes advantage of resources at the edges of the Internet* (Clay Shirky)

Vain yksi uusi kupla?



The peer-to-peer backlash has begun. On the same day, the Wall St. Journal ran an article by Lee Gomes entitled "Is P2P plunging off the deep end?", while Slashdot's resident commentator, Jon Katz, ran a review of O'Reilly's *Peer to Peer* book under the title "Does peer-to-peer suck?"

It's tempting to write this off as part of the Great Wheel of Hype we've been living with for years:

New Thing happens; someone thinks up catchy label for New Thing; press picks up on New Thing story; pundits line up to declare New Thing "Greatest Since Sliced Bread." Whole world not transformed in matter of months; press investigates further; New Thing turns out to be only best thing since soda in cans; pundits (often the same ones) line up to say they never believed it anyway.

This quick reversal is certainly part of the story here. The Journal quoted entrepreneurs and investors recently associated with peer-to-peer who are now distancing themselves from the phrase in order to avoid getting caught in the backlash. There is more to these critiques than business people simply repositioning themselves when the story crescendos, however, because each of the articles captures something important and true about peer-to-peer.

Gartner Group: Niche technology

- Gartner Groupille täsmämarkkina, jonka yrityspenetraatio max. 10 % vuoteen 2010 mennessä

*Häviää käsitteenä,
koska muodostuu valtavirraksi....?*

Ei niin uutta

- Unix-käyttöjärjestelmässä tyypillinen ajattelutapa
- lähiverkoissa vakiintunut toimintatapa
 - kirjoittimien jakaminen
 - tiedostojen jakaminen

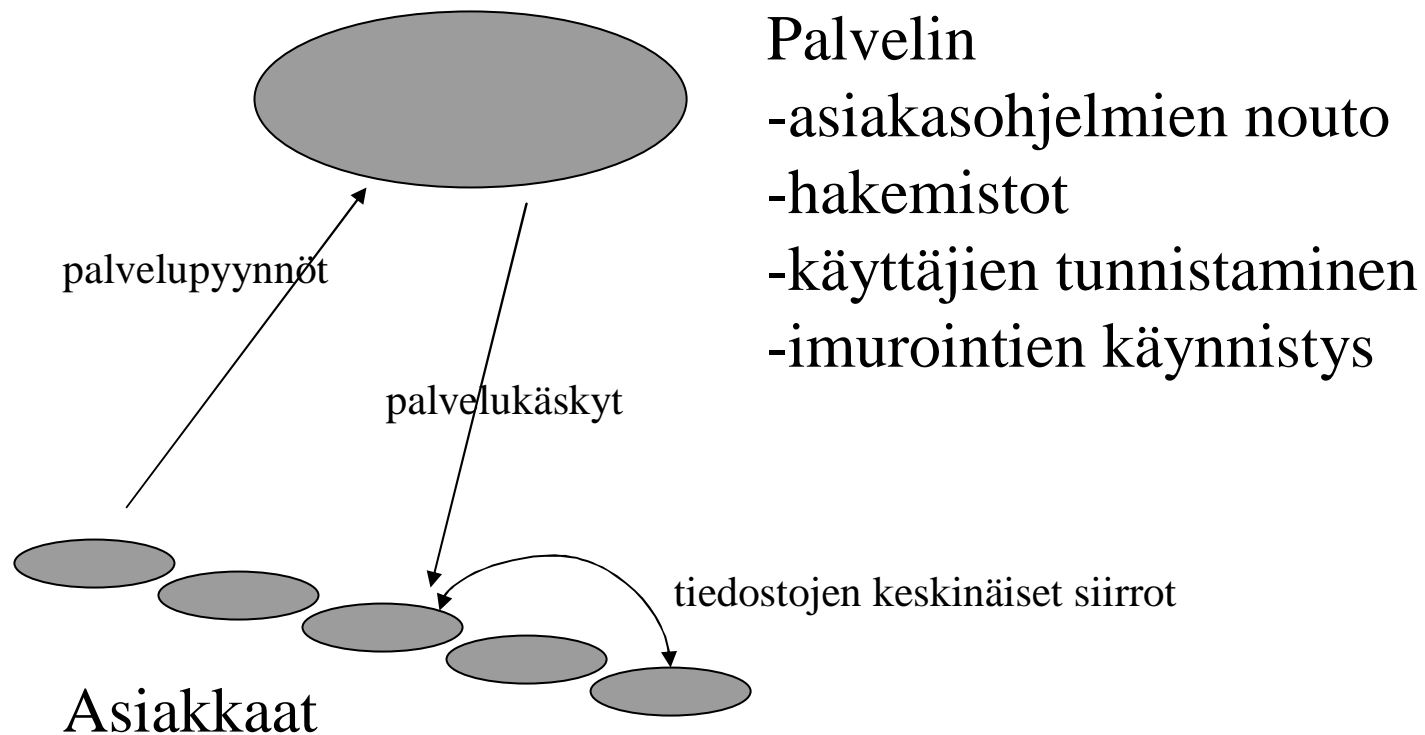
Lähikontaktointi...

- Bluetooth
- Jini
 - Software from Sun Microsystems that seeks to simplify the connection and sharing of devices, such as printers and disk drives, on a network. Currently adding such devices to a computer or network requires installation and boot-up, but a device that incorporates Jini will announce itself to the network, provide some details about its capabilities, and immediately become accessible to other devices on the network.

Läpilyöntisovellus sisällön jakamisessa - Napster

- Sovellus MP3 –musiikkitiedostojen jakamiseksi
- Toukokuussa 1999, Shawn Fanning (18v) Kaliforniassa teki ensimmäisen sovelluksen
- Parhaimmillaan
 - 30 miljoonaa käyttäjää
 - 3 miljardia imurointia kuukaudessa
- Lokakuu 2000: Bertelsman ostaa oikeudet
- Nyt suljettu USA:n oikeuden päätöksellä

Napsterin toimintaperiaate



...mutta elämä jatkuu

- Uusi, kuukausimaksuun perustuva malli kehitteillä
- Monia muita seuraajia, jotka entistä hajautetumpia
 - Gnutella / Bear Share
 - Morpheus
 - Grokster
 - KaZaA
- Tiedostomuodolle ei rajoituksia

www.musiccity.com

gnutella.wego.com

www.listen.com



Freenet on laaja vertaisverkko (peer-to-peer network), joka kokoaa siihen liitettyjen tietokoneiden voiman maailmanlaajuisesti tietovarastoksi, jossa kuka tahansa voi vapaasti julkaista tai katsella minkä tahansa kaltaista tietoa. Freenet on:

- Kestävä: Kaikki sisäiset prosessit ovat täysin anonyymejä ja keskittämättömiä koko verkossa, jolloin on lähes mahdotonta hyökkääjän poistaa jotain tietoa tai ottaa järjestelmää haltuun.
- Yksityinen: Freenet tekee äärimmäisen vaikeaksi kekenkään vakoilla, mitä tietoa katselet, julkaiset tai tallennat.
- Turvallinen: Freenettiin tallennettava tieto suojataan vahvalla salauksella ilkeämielistä sormeilua ja väärentämistä vastaan.
- Tehokas: Freenet kopioi ja siirtää tietoa dynaamisesti kysynnän mukaan tarjotakseen tehokkaan palvelun ja minimaalisen kaistan käytön riippumatta kuormituksesta. Merkittävää on, että Freenet yleisesti ottaen tarvitsee vain $\log(n)$ aikaa hakeakseen tiedon verkosta, jonka koko on n .

<http://freenetproject.org/cgi-bin/twiki/view/FI/WhatIs>

Historia

Freenet on laajennettu avoimen lähdekoodin (Open Source) toteutus järjestelmästä, joka kuvataan Ian Clarken 1999 julkaisemasta tutkielmasta 'A distributed decentralized information storage and retrieval system' (Katso [täältä](#)). Ian Clarke ja pieni ryhmä vapaaehtoisia aloittivat työn Freenetin luomiseksi pian tutkielman julkaisun jälkeen heinäkuussa 1999. Maaliskuussa 2000 versio 0.1 Freetistä oli valmis julkaisua varten. Maaliskuusta 2000 lähtien Freetistä on kirjoitettu laajalti lehdistössä, vaikkakin pääasiassa sen vaikutuksista tekijänoikeuksiin, eikä sen laajemmasta kohteesta, nimittäin viestintävapaudesta.

Freenetin kehitys jatkuu kiihtyvällä tahdilla, uusien versioiden ilmestyessä lähes kahden viikon välein, jokaisen version tuodessa merkittäviä tehokkuus- ja käyttöparannuksia. Tätä kirjoitettaessa 0.4-julkaisu mittavine rakenteellisine, tehokkuutta ja turvallisuutta parantavine muutoksineen, on lähestymässä valmistumistaan.

Sovellukset

Järjestelmä tarjoaa joustavan ja tehokkaan rakenteen, jota voidaan hyödyntää monenlaisin sovelluksin, mukaanlukien:

- Kyseenalaisen tiedon levittäminen sensuroimattomana: Freenet suojelee sananvapautta mahdollistamalla materiaalin ruohonjuuritason vaihtoehtoisesta journalismista kiellettyjen julkaisujen, kuten Peter (Spycatcher) Wrightin ja David Shaylerin paljastuksien MI5:stä sensuroimattoman julkaisemisen anonyymisti.
- Tehokkaan suuren kaistanleveyden vaatiman sisällön levittäminen: Freenetin mukautuvaa välimuistia ja peilausta käytetään levittämään Debian Linuxin ohjelmistopäivityksiä ja taistelemiseen Slashdot-efektiä vastaan.
- Maailmanlaajuinen henkilökohtainen julkaiseminen: Freenet tarjoaa kaikille mahdollisuuden saada oma verkkisivu ilman tilarajoituksia tai pakollisia mainoksia, vaikkei omistaisi edes tietokonetta.

Freenet on avoin, demokraattinen järjestelmä, mitä ei voi kukaan ihminen, ei edes sen tekijät, kontrolloida. Se on alunperin suunnitellut Ian Clarke ja sitä tehdään avoimen lähdekoodin mallin mukaan vapaaehtoisjoukoin.

Esimerkki: Real One

- 9,95 \$ kuukaudessa
- 200 latausta: puolet vain suoraan kuultaviksi on-line, puolet 30 päivää kestävässä formaatissa, jota ei voi kopioida (ei MP3 eikä CD)
- vain USA

Lemm, Karsten (2001)

”Wie ungeschützer Sex” Stern 51/2001, 246

Naspter ja laki

- Monimutkainen oikeudenkäyntivyyhti
- Heinäkuusta 2001 alkaen suljettu, kunnes laillisuus on selvitetty – valitus
- Napster maksoi 26 miljoonaa dollaria musiikin tuottajille riidan osittaiseksi sopimiseksi
 - 1/3 artisteille
 - 2/3 levy-yhtiöille

Oikeusperiaatteita Napsterille

Musiikkiteollisuus

- P2P laitonta tekijänoikeuksien rikkomista
- *Secure Digital Music Initiative*
- *Digital Millennium Copyright Act*
 - ... jopa tietokoneiden valmistajien pitäisi maksaa tekijänoikeusmaksuja

Oikeusperiaatteita Napsterille – vasta-argumentteja

- 70 vuoden iässä tekijänoikeus vanhenee
- tekijänoikeuslait: vapaa kopiointi kotikäyttöön, myös rajatulle määrälle ystäviä (15-20?)
 - digitaaliset suojat loukkaavat lain henkeä
- P2P ok niin kauan kuin ei kaupallista hyötyä kenellekään
 - keskitetyt palvelut rahoitetaan mainosrahoin

EU:n uusi tekijänoikeusluonnos

Euroopan komission luonnos uudesta tekijänoikeusdirektiivistä on saanut vihaisen vastaanoton äänite- ja elokuva-alalla. Komision tämän hetkisen kannan mukaan **vain kaupalliset tekijänoikeusrikkomukset ovat rangaistavia.**

Europemedia kertoo uuden tekijänoikeusdirektiivin luonnoksesta, jossa otetaan kantaa muun muassa vertaisverkkoihin perustuvan tiedostonvaihdon laillisuuteen. Komissio on näillä näkyvin ehdottamassa parlamentille ja ministerineuvostolle esitystä, joka määrittelisi tekijänoikeuslain vastaiseksi vain kaupallista hyötyä tavoittelevan piratismiin.

Äänite- ja elokuvateollisuuden edustajat ovat reagoineet vihaisesti komission kantaan, joka **selvästi jättäisi vertaisverkkojen käytön rangaistavien tekojen ulkopuolelle.** Ohjelmia levittävien yhtiöiden, kuten Kazaan takana olevan Sharman Networksinkin toiminnan, komissio näkee laittomana.

Euroopan unionin parlamentin ja ministerineuvoston on vielä hyväksyttävä komission ehdotus uudeksi direktiiviksi ennen sen voimaan tuloa.

Musiikkia ja muita teoksia kopioidaan yksityiseen käyttöön merkittävässä määrin. Tämä käy ilmi Kasettimaksutoimiston juuri valmistuneesta kotikopiointitutkimuksesta. Joka viides suomalainen on tallentanut musiikkia digitaaliseen muotoon, kuten cd-r-levyille, tietokoneen kovalevyille ja Minidiscille. Ahkerimpia kopioijia ovat 15-19 -vuotiaat miehet. Digikopiointi on kuitenkin levinnyt jo koko väestön harrastukseksi. Pääosin musiikki kopioidaan alkuperäisiltä cd-levyiltä (68 %) ja Internetistä (23 %). Keskimäärin tallennetaan kerralla koko cd-levyllinen musiikkia eli noin 12 teosta.

Kolmanneksella suomalaisista on kotonaan itsepoltetuja cd-r-levyjä. Kaikesta cd-r-levyille tallennetusta aineistosta 46 prosenttia on musiikkia ja 18 prosenttia kuvia. Tämän lisäksi cd-r-levyille kopioitiin muun muassa tekstitiedostoja ja taulukoita.

Tutkimuksen mukaan tietokoneen kovalevyille tallennetaan kotikäyttöön keskimäärin noin 50 prosenttia tekstiä, 15 prosenttia kuvaa ja 10 prosenttia musiikkia. Loppu on pelejä, liikkuvaa kuvaa, tietokoneohjelmia, taulukoita ja muuta informaatiota.

Tietokoneen kovalevyille kopioidusta kotikäyttöön tarkoitettua materiaalista 28 prosenttia on julkaistua aineistoa. Noin 70 prosenttia kovalevyille kopioidusta julkaistusta aineistosta on lähtöisin Internetistä. Hieman alle 20 prosentissa tapauksissa lähde on alkuperäinen cd.

Tutkimukseen haastateltiin noin tuhat 15-74-vuotiasta suomalaista sekä helmi- että lokakuussa 2002.

Cd-r-levyjen määrä kasvaa

Vuonna 2001 cd-r-levyjä tuotiin maahan Kasettimaksutoimiston tilastojen mukaan 14,5 miljoonaa kappaletta. Vastaava luku edellisvuonna oli 12,5 miljoonaa. Vuonna 2002 myynnin kasvu jatkuu edelleen.

Suomen ääni- ja kuvatallennetuottajat ÄKT:n tilastojen mukaan vuonna 2001 Suomessa myytiin 11,7 miljoonaa alkuperäistä äänitallennetta. Vuoden 2002 yhdeksän ensimmäisen kuukauden kappalemääräinen myynti on pienentynyt noin 5 prosenttia.

Hallituksen esitys tekijänoikeuslain muuttamiseksi annettiin eduskunnalle lokakuussa. EU:n tietoyhteiskuntadirektiivin edellyttämät muutokset tulee saattaa kansallisesti voimaan 22.12.2002 mennessä. Esityksen mukaan myös digitaalinen kopiointi yksityiseen käyttöön on jatkossakin Suomessa vapaata. Direktiivi edellyttää, että oikeudenhaltijat saavat sopivan hyvityksen laajamittaisesta yksityisestä kopioinnista. Sopiva hyvitys ehdotetaan Suomessa toteutettavaksi nykyistä kasettimaksusäännöstöä kehittämällä. Kasettimaksun perinnän ja käytännön hallinnoinnin hoitaa Teoston Kasettimaksutoimisto, jonka opetusministeriö on nimennyt tehtävänsä.

Digital Rights Management

- laiton kopiointi estetään, mahdollistetaan omaan käyttöön kopiointi
- työn alla, piirteitä Windows XP-järjestelmässä
- Palvelun vaihtoehdot:
 - kiinteä tilausmaksu
 - maksu jokaisesta imuroinnista

Resurssien jakaminen

- 98% tietokoneresursseista on käyttämättä, erityisesti tietoliikennekapasiteetissa ja prosessointitehossa
- Uutta:
 - jokaisella osaprosessorilla vain hyvin pieni työmäärä
 - jokainen noodi voi poistua ja tulla mukaan vapaasti

SETI – tarkoituksellisten signaalien etsimistä

Search for Extraterrestrial Intelligence

- markkinoitu näytönsäästäjänä
- noin 3 miljoonaa käyttäjää, 30 terafloppia kapasiteettia



ISOROKON LÄÄKITYS

IBM on käynnistänyt yhdessä tietotekniikkayritys United Devicesin ja ohjelmistovalmistaja Accelrysin kanssa projektin, jolla etsitään lääkettä isorokkoinfektioon. Projektia varten luodaan miljoonien tietokoneiden muodostama maailmanlaajuinen grid-tietokoneverkko, joka yhdistää yksittäisten koneiden kapasiteetin ja hyödyntää niiden ylimääräistä tehoa potentiaalisten isorokkolääkkeiden etsimiseen. IBM:n mukaan tutkimuksen tulokset annetaan Yhdysvaltain puolustusministeriön käyttöön.

Projektiin osallistuu tutkijoita muun muassa brittiläisistä Oxfordin ja Essexin yliopistoista sekä isorokkoasiantuntijoita monista tutkimuslaitoksista. Projekti perustuu IBM:n, United Devicesin, Accelrysin sekä lääkkeitä kehittävän Evotec OAI -yrityksen kaupallisessa käytössä oleviin teknologioihin ja palveluihin, joita muun muassa monet lääke- ja bioteknologiayritykset käyttävät tutkimustyössään. Tutkimuksen runkona toimiva grid-verkko yhdistää jopa kahden miljoonan tietokoneen kapasiteetin, mikä voi antaa tutkijoiden käyttöön suurimmillaan 30 kertaa maailman tehokkaimpien supertietokoneiden laskentatehon.

Verkon avulla tutkijat analysoivat yhteensä 35 miljoonan lääkkeeseen mahdollisesti soveltuvan molekyylin keskinäistä suhdetta ja kykyä toimia isorokkoviruksen proteiineja vastaan. Tutkimukseen voi osallistua lataamalla tietokoneeseensa näytönsäästäjän osoitteesta www.grid.org.

- Kannustamme yhdessä yhteistyökumppaniemme kanssa tietokoneen ja Internet-yhteyden omistajia kaikkialla maailmassa liittymään verkkoon ja antamaan siten tietokoneidensa ylimääräisen kapasiteetin projektin käyttöön. Tulevaisuudessa grid-verkot avaavat esimerkiksi lääketeollisuudelle aivan uusia mahdollisuuksia kehittää nopeasti ja tehokkaasti uusia tuotteita, sanoo IBM:n grid-tiedonkäsittelystä vastaava johtaja Tom Hawk.



Resurssien jakamisen kysymyksiä

- Superdistribution: resurssin imurointi samanaikaisesti useasta paikasta: ei pullonkauloja
- Tietoturva ja –suoja molempiin suuntiin
- Suurten keskitettyjen palvelimien tarve vähenee

Ryhmätyösovellukset

- ICQ
- AOL Instant messaging
- Microsoft Messenger
- Lotus Sametime

Suosituin versio: tekstiviestit ja seurannaiset

”non-intrusive but efficient and quick”

Kehittyneempiä ryhmätyösovelluksia

- On-line konferenssit
Microsoft netMeeting
- Hajautettu sisällön hallinta
NextPage NP3

Jaetut virtuaalityötilat

Groove Networks

- kehittäjä Ray Ozzie, Lotus Notesin ” isä ”
- mm. kalenteri, VoIP, chat, dynaaminen käyttäjien hallinta
- valmiita alustoja sovelluksille: projektinhallinta, huutokaupat, yms...

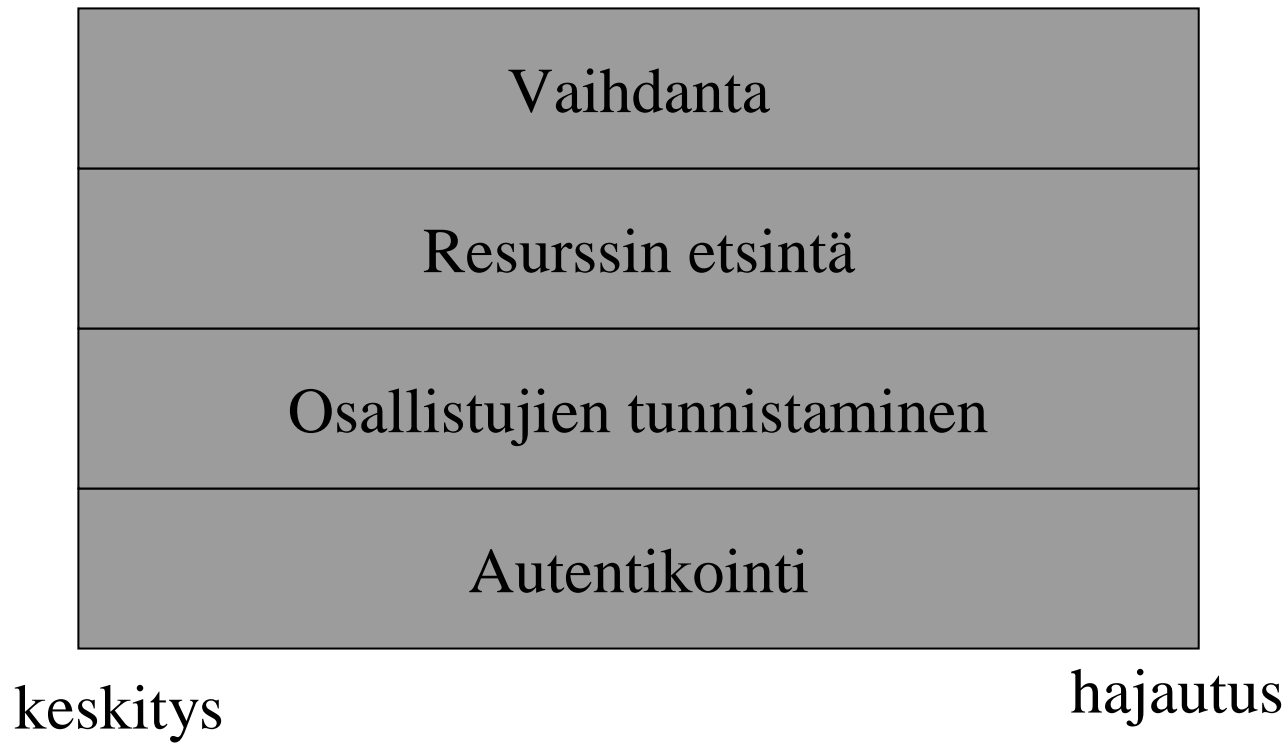
Ryhmätyösovellukset - Mahdollisuudet

- Ei raskaita keskitettyjä arkkitehtuureja, sopii nykyajan verkottuvaan maailmaan
- Joustavuus ja vikasietoisuus
- Suoraviestintä jo vakiintunut sovellus

Ryhmätyösovellukset - Heikkoudet

- Käyttäjätuki vaikeaa
- Tietosuoja- ja –turvaongelmat
- Käyttäjien tunnistaminen
- Vaatii työasemilta lisää tehoa

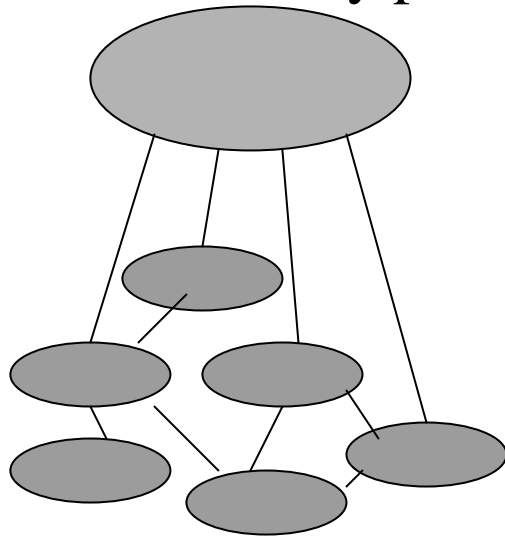
Tasottainen malli



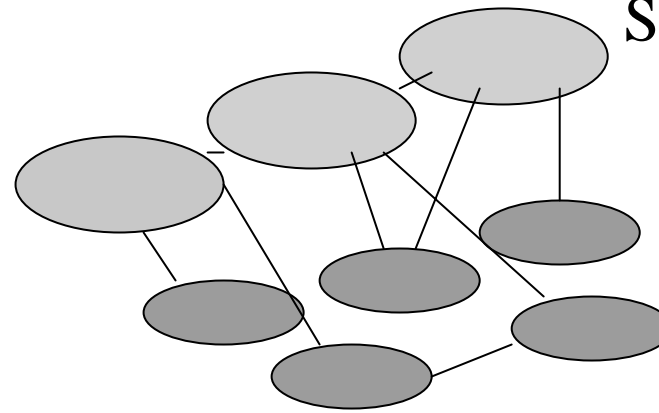
Jokaisen tason toiminnat voivat olla tosimailmassa keskitettyjä tai hajautettuja

Erilaisia kokonaisarkkitehtuureja

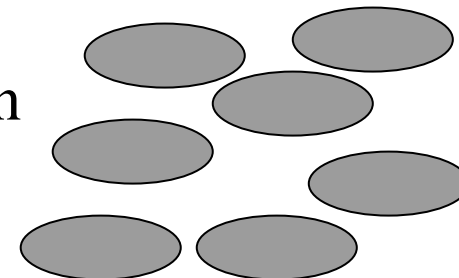
Keskistetty palvelin



Supernoodeja



Täydellinen
hajautus



Standardiehdotus JXTA

- JXTA (Juxtapose)
- SUN Microsystems
- Java toteutusväline
- Kuusi perusprotokollaa
 - peer discovery (autentikointi)
 - peer resolver (etsintä)
 - peer information
 - peer membership
 - pipe binding
 - peer endpoint routing

Muita standardiehdotuksia

PtPTL Peer-to-peer Trusted Library

Intel

Microsoft.NET

MS Passport (Hotmail, yli 100 miljoonaa käyttäjää)

.NET My Services (läsnöolo, personointi)

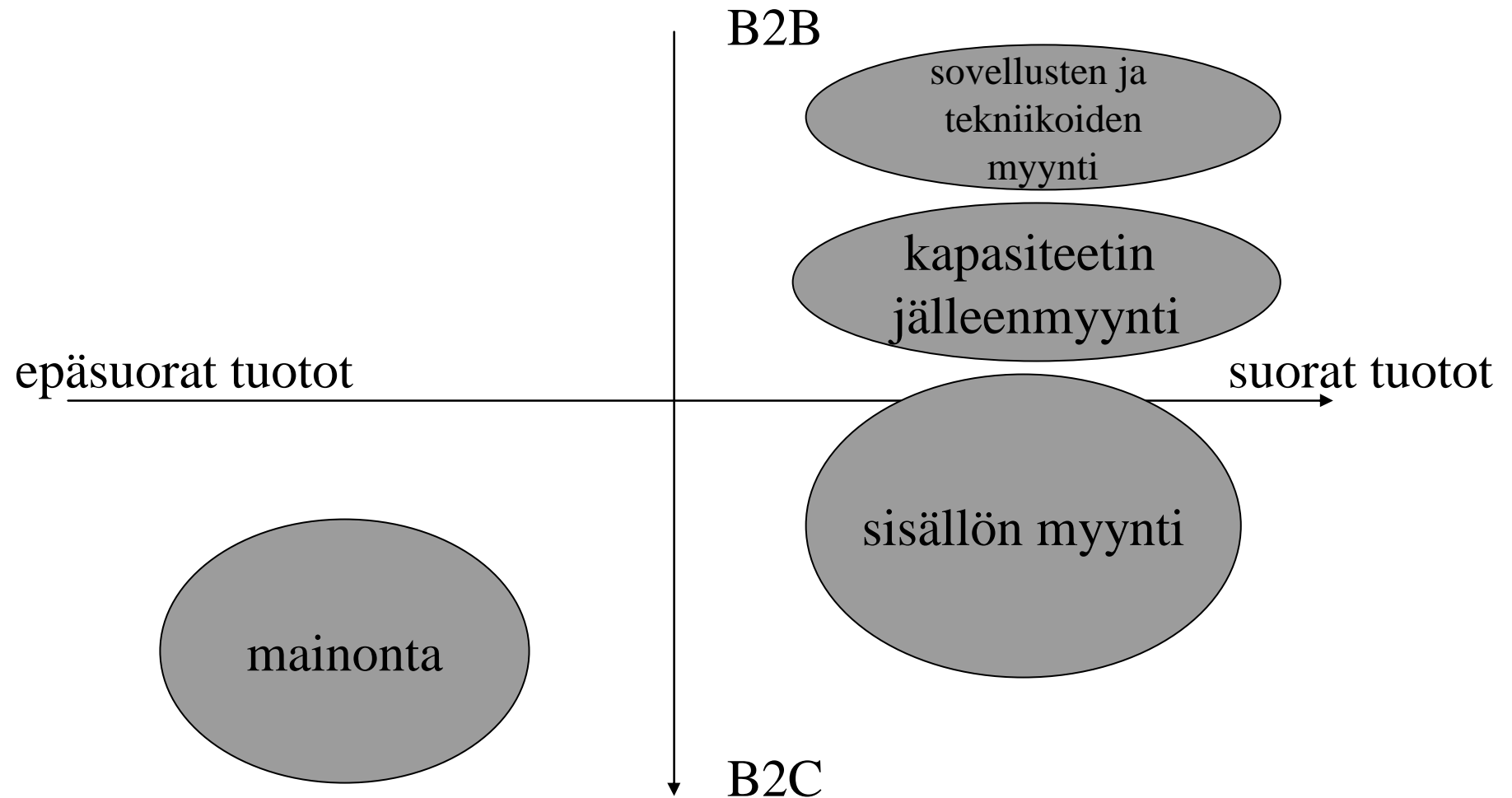
.NET Framework (ohjelmakirjasto)

Visual Studio.NET (sovelluskehitystyökalut)

Taloudellisia perusteita

- Luodaan ”Network externalities”
- Olemassa olevien resurssien tehokas käyttö
- Koko tietojenkäsittelyn tuottavuuden selkeä kasvu

Liiketoiminnan mahdollisuuksia



Kirjallisuutta

Oram, Andy (2001) Peer-to-Peer: Harnessing the Power of Disruptive Technologies. O'Reilly.

Barkai, David (2001) Peer-to-Peer Computing. Intel Press.