

dPersp – ugeseddel for uge 41

Internetalgoritmer

Ugens program

Mandag 6/10	12.15-14.00	Forelæsning: Internetalgoritmer - Google's PageRank <i>Gerth Stølting Brodal (Aud E)</i>
Tirsdag 7/10	9.15-12.00	Øvelser
	12.15-13.00	Forelæsning: MapReduce <i>Gerth Stølting Brodal (PBA Auditoriet)</i>
	13.15-16.00	Øvelser
Onsdag 8/10	14.15-16.00	Historisk perspektiv: Programmeringssprog Den afsluttende opgave over "As We May Think" stilles Kort "her-og-nu" evaluering af dPersp14 (passende skriveredskab bedes medbragt) <i>Erik Meineche Schmidt (PBA Auditoret)</i>

Om ugens forelæsninger

Denne uges forelæsninger omhandler teknikker og algoritmer der i høj grad har påvirket vores brug af internettet de seneste år – med fokus på to eksempler.

Internetsøgemaskiner

Med introduktionen af Google i 1998 blev det muligt at søge effektivt blandt de milliarder af websider der findes på internettet. Nogle af de egenskaber der var fundamentet for Google's succes var hastigheden hvormed forespørgsler udføres, at de fleste sider på internettet var indeholdt i Google's indeks, og ikke mindst at de returnerede svar indeholdt de meste relevante sider om det man søger efter. I forelæsningen beskrives de basale idéer der var grundlaget for Google's virkemåde.

MapReduce

Et af fundamentene i Google's succes er at have mange maskiner (størrelsesorden millioner) forbundet i et stort netværk, og ved at effektivt udnytte den massive parallelisme dette giver. Traditionelle software metoder til programmering af parallelle algoritmer medfører dog i praksis et meget stort overhead til håndtering af tilbagevendende problemer af parallel natur. Google introducerede derfor i 2004 en elegant softwareomgivelse kaldet MapReduce, der måske ikke kan løse alle problemer effektivt parallelt, men tillader en simpel måde at designe parallelle algoritmer, der automatisk håndterer alle aspekter der vedrører udnyttelsen af parallelisme. Der findes en open source variant af MapReduce, Hadoop, som bl.a. bruges intenstivt af Yahoo! og Amazon.