



ID2221 Data-intensiv data-behandling 7,5 hp

Data-Intensive Computing

Fastställande

Betygsskala

A, B, C, D, E, FX, F

Utbildningsnivå

Avancerad nivå

Huvudområden

Datalogi och datateknik

Särskild behörighet

Undervisningsspråk

Undervisningsspråk anges i kurstillfällesinformationen i kurs- och programkatalogen.

Lärandemål

Kursen kompletterar kurserna i distribuerade system med ett fokus på bearbetning, lagring och analys av massiva datamängder. Den förbereder studenterna på examensarbetsprojekt för masterexamen, och för forskarutbildning inom området data-intensive computing sys-

tems.

Huvudsyftet med denna kurs är att erbjuda studenten en solid grund för att förstå storskaliga distribuerade system som används för att lagra och bearbeta massiva datamängder.

Mer specifikt kommer studenter efter avslutad kurs att kunna

- förklara arkitekturen och de egenskaper hos datorsystem som behövs för att lagra, utsöka och indexera stora datavolymer
- beskriva olika beräkningsmodeller för att bearbeta stora datamängder för statiska data (batchbearbetning) och för data i rörelse (stream processing)
- använda olika beräkningssystem för design och implementation av icke-triviala analysmetoder på massiva data
- förklara olika modeller för att schemalägga och allokera resurser för beräkningsuppgifter på stora datorkluster
- förklara avvägningarna som görs vid design av effektiva algoritmer för att bearbeta stora datamängder i en distribuerad beräkning.

Kursinnehåll

Ämnen:

- Distribuerade filsystem
- “No-SQL”-databaser
- Skalbara meddelandesystem
- Exekveringsmaskiner för Big Data: Map-Reduce, Spark
- Högnivå-frågor och interaktiv bearbetning: Hive och Spark SQL
- Stream processing
- Graph processing
- Skalbar maskininlärning
- Resurshantering.

Examination

- TEN1 - Tentamen, 4,5 hp, betygsskala: A, B, C, D, E, FX, F
- LAB1 - Programmeringsuppgifter, 3,0 hp, betygsskala: P, F

Examinator beslutar, baserat på rekommendation från KTH:s handläggare av stöd till studenter med funktionsnedsättning, om eventuell anpassad examination för studenter med dokumenterad, varaktig funktionsnedsättning.

Examinator får medge annan examinationsform vid omexamination av enstaka studenter.

När kurs inte längre ges har student möjlighet att examineras under ytterligare två läsår.

Skriftlig tentamen. Laborationsuppgifter.

Etiskt förhållningssätt

- Vid grupparbete har alla i gruppen ansvar för gruppens arbete.
- Vid examination ska varje student ärligt redovisa hjälp som erhållits och källor som använts.
- Vid muntlig examination ska varje student kunna redogöra för hela uppgiften och hela lösningen.