

HVEM: Gunnar Brataas og Richard Sanders, SINTEF IKT

HVA: skal lage metoder og verktøy som kan bidra til å unngå datakrasj i såkalte sky-programmer

HVORFOR: Mange dataprogrammerere har for lite kunnskap om skalering, det vil si å vurdere hvor mye belastning et system tåler.

AV SIRI LINDSTAD

Når datasystemer krasjer



– Noen systemer får man ikke testet ordentlig før de faktisk må virke, så da må man bare la det stå til, sier Gunnar Brataas (t.v.) og Richard Sanders. (Foto: Ole Morten Melgård)

PROSJEKT: CloudScale. **INSTITUSJON:** SINTEF IKT i samarbeid med norske, tyske og kroatisk industribedrifter og forskningsmiljøer. **FAGRETNING:** Systemutvikling og sikkerhet. **FINANSIERING:** EUs sjuende rammeprogram, samt egenfinansiering. **ARBEIDSFORM:** Utvikling og utprøving av språk, metoder og verktøy. **VIKTIGSTE ARBEIDSREDSKAP:** Dataprogrammer og datamaskiner. **PUBLISERINGSFORM:** Foredrag, artikler, sosiale nettverk. **NETTSIDE:** www.cloudscale-project.eu. **NY KUNNSKAP:** Det kan alltid gå galt når man sjøsetter systemer. Men med gode verktøy og metoder kan man minske risikoen.



Alle skulle plutselig sjekke tallene i selvangivelsen. Det var mer enn Altinn kunne tåle i 2012.
(Foto: Scanpix/Espen S. Hoen)

Det er snart vår og dermed tid for å håndtere selvangivelsen. Hvis det funker, da. For et sikkert vårtegn de siste årene har vært at internettportalen Altinn har knelt når hundretusenvise av ivrige nordmenn har forsøkt å logge seg på.

Men hva er det egentlig som går galt, gang på gang? Ikke bare Altinn, men også en rekke andre datasystemer har det med å kollapse? Er det maskinen? Er det nettet? Er det meg?

Nei, det er de som utvikler og drifter systemene uten helt å overskue konsekvensene av mange små og store valg som blir gjort, ifølge SINTEF-forskere. De er ikke flinke nok til å skalere.

Flest ikke alltid best

Å skalere vil si å plusse på mer maskinvare på systemet, og så studere hvor mye datalast, altså hvor mange brukere, systemet klarer å håndtere.

– Tenk deg at du skal løfte noe sammen med noen. Er dere to, kan dere trolig løfte dobbelt så mye som om du var alene. Er dere tre, klarer dere muligens tre ganger så mye. Men er dere ti, blir det gjerne problemer. Dere klarer ikke å løfte ti ganger så mye som én person. Kanskje dere til og med klarer å løfte mindre enn sju personer klarer, forklarer Gunnar Brataas.

Han er forsker ved SINTEF IKT i Trondheim og har ytelse og skalering som hjertebarn og fagfelt. Nå er han og tre SINTEF-kolleger i gang med EU-prosjektet CloudScale, som handler om å lage metoder og verktøy som kan forbedre skalerbarheten til såkalte sky-programmer.

– Jeg snakket med en fra et firma som håndterte et datasystem hvor du skulle kunne lese av strømmen. Han mente at 90 prosent av dagens datautviklere ødelegger skaleringen i systemet, mens de siste 10 prosentene kan fikse det igjen. Ja, det er litt satt på spissen, men likevel noe sant, sier Brataas.

Må finne knekkpunktet

Det finnes et knekkpunkt i mange systemer der det ikke lenger nytter å pøse på med ressurser, det være seg kjappere nettoppkobling eller kraftigere maskiner. Det blir for mye som skjer på én gang, for mye koordinering på tvers.

– Denne helheten må den som programmerer, ha blick for. Den gang systemene var mindre og enklere satt sammen, var det kanskje lettere for programmererne å holde oversikt over det som skjedde under kjøring. Dermed kunne man lettere sette inn tiltak dersom det buttet imot. Nå blir gjerne avstanden mellom dem som lager systemet, og selve systemet bare større og større. Samtidig er også datasystemene mer innvevd i dagliglivet enn før. Det gjør at man kan gå på flere smell, og de får større konsekvenser.

Opp i skyen

Ja, alt var øyensynlig enklere før. Da lå all datakraftkapasiteten lokalt, på maskinen din som gjerne var koblet til en server borti korridoren. Men etter hvert som systemene krevde stadig mer ressurser, har det blitt stadig vanligere å sette ut datakraftkapasiteten i den såkalte «nettskyen». Man setter rett og slett driften ut til servere som står i eksterne datasentre – kanskje på den andre siden av kloden – og som man er tilknyttet via internett.

Også datakraften bak et tilsynelatende enkelt spill som Wordfeud er satt bort. Bak det Scrabble-lignende spillet står Håkon Bertheussen, en tromsøværing med fem års utdannelse i datateknikk fra NTNU i Trondheim. Han lanserte Wordfeud som app 1. august 2010, og så langt har vel 20 millioner mennesker over hele verden lastet ned spillet til mobilen sin. I USA står det som Bertheussen i et intervju betegnet som 14 «ekstremt gode servere», som sørger for at trafikken flyter som den skal.

– Men hvilke funksjoner skal ligge lokalt på mobilen, og hva skal betjenes av skyen? Hva med ordlisten, som er en viktig del av spillet? Skal den lagres som en del av appen på mobilen din, eller skal det sendes en forespørsel til USA hver gang du skal sjekke et ord? Ikke vet jeg hva Bertheussen har valgt av løsninger, men slike valg har han måttet ta for at spillet skal fungere best mulig, sier Richard Sanders, også han forsker ved SINTEF IKT, og i denne sammenheng prosjektleder for CloudScale.

Det enkleste er det kjedeligste

Det er imidlertid Brataas som er den store skaleringsentusiastene av de to. Han tok i sin tid hovedfag i kunstig intelligens, men fikk siden stipend ved daværende NTH innenfor ytelse av datasystemer, og dermed var det gjort.

– Jeg jobbet en del med nettaviser tidligere, men det viste seg å være ganske kjedelig.

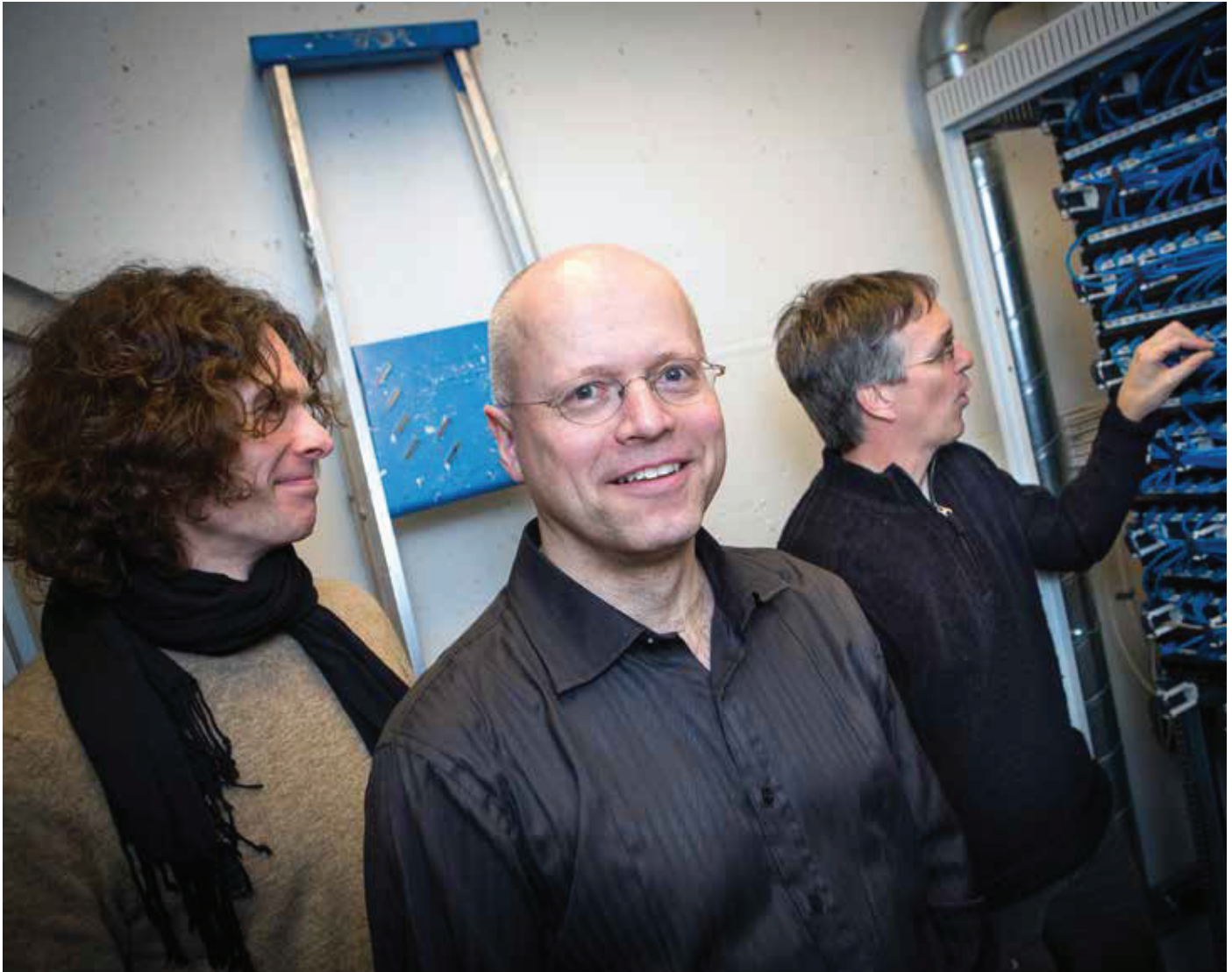
Nettaviser er nemlig enkle skaleringsmessig – antall brukere kan økes proporsjonalt med mengden dataressurser. Det er selvfølgelig bra for dem, men altså kjedelig for forskeren.

– Systemer der kundene ikke interagerer med andre kunder, som hos nettaviser og nettbanker, fungerer stort sett greit. Når du logger deg inn i nettbanken, sysler du jo hovedsakelig med dine egne ting, og det er dermed lett å legge til flere tjenere etter hvert som antallet kunder vokser.

I for eksempel Altinn er det derimot en haug med prosesser som skal interagere.

– Om du bare skal sjekke selvangivelsen, er det i og for seg ikke så farlig om systemet bryter sammen og du må vente noen dager ekstra. Men for alle dem som for eksempel skal levere fersk fisk på kontinentet og behøver å få tollklarert papirer gjennom det samme systemet, blir situasjonen snart mer kritisk, sier Brataas.

Han smiler litt beskjemmet når vi spør om



– Det beste er ofte at dataprogrammererne velger å gjøre enkle ting og tenke at det er godt nok, sier Gunnar Brataas (i midten), flankert av Erlend Stav (t.v.) og Richard Sanders. (Foto: Ole Morten Melgård)

han synes det er gøy å høre om systemer som bryter sammen?

– Ja, det blir litt som når du legges inn på sykehus, og legene roper: Oi, dette er et spennende tilfelle. Kom og se!

Føre var

Å lage et datasystem og å skrive en god kode for et slikt system er en avveining mellom mange hensyn som er i konflikt.

– Det som er viktig, er å lage koder som er enkle å vedlikeholde, og som er sikre. Ofte vil en slik kode også ha god ytelse og bra skalering. Men ikke bestandig. Dessuten prøver mange utviklere ut systemet som enkeltbrukere og har dårlig oversikt over hvordan det blir når mange skal ta det i bruk, sier Brataas.

Løsningen er å lage gode modeller som i størst mulig grad gjør at du kan sjekke hva som blir konsekvensene før systemet faktisk skal virke. På den måten vet man hvilke deler av systemet som er kritiske med hensyn til ytelse og skalerbarhet.

Det er slike modeller CloudScale-prosjektet skal hjelpe til med å utvikle.

Håper på et godt rykte

CloudScale-prosjektet har nettopp startet opp og skal pågå i tre år. Det er delvis finansiert gjennom EUs sjuende rammeprogram, med partnere fra Tyskland, Slovenia og Kroatia i tillegg til Norge – både industribedrifter og forskningsmiljøer. Målet er altså å utarbeide metoder og verktøy som kan hjelpe programutviklere som ikke er eksperter på skalering, slik at de kan få oversikt over skalerbarheten til sky-programmer.

Resultatene av prosjektet skal spres gjennom artikler, på konferanser, gjennom sosiale nettverk som LinkedIn og så videre.

– Det viktige er at vi får demonstrert resultatene for potensielle brukere, slik at de får et godt rykte, og slik at bruken dermed sprer seg, sier Sanders.

Men tjene penger kommer de nok ikke til å gjøre, tror han.

– Dette er anvendt forskning, ikke produkt-

utvikling, ellers hadde EU-kommisjonen neppe brukt penger på det. Men vi håper å kunne utvikle dataverktøy med prototypekvalitet som andre kan gjøre butikk på senere. De vil i så fall neppe bruke en eneste kodelinje som vi har skrevet, men de vil lage noe tilsvarende, bare bedre og mer tilpasset egne behov. Hvis det skjer, er prosjektet en suksess. ■

SINTEF IKT

- Leverer forskningsbasert kompetanse og forskningsbaserte tjenester og produkter innenfor disse områdene: mikroteknologi, informasjonssystemer, beregningsorientert programvare, sikkerhet og sårbarhet, samt kommunikasjons- og programvareteknologi.

- Har vel 300 ansatte fra hele 28 land.