
Ådne Helmersen

En Sammenligning av Microsoft Azure og
Amazon Web Services

Forstudierapport

1. Introduksjon – hensikten med dokumentet

Denne rapporten utarbeides som en del av forstudiet til bacheloroppgaven jeg skriver våren 2021. Oppgaven er selvvalgt og går ut på å sammenligne to plattformer for skyteknologi, «Amazon Web Services» og «Microsoft Azure». Denne forstudierapporten skal gi deg som leser av bacheloroppgaven oversikt over hva jeg skal gjøre gjennom dette prosjektet, den skal gi svar på hvorfor jeg har valgt akkurat dette temaet og hva jeg ønsker å oppnå med oppgaven. Jeg vil presentere en risikoanalyse for å gjøre deg klar over hvilke utfordringer man kan støte på og redegjøre for timeantall som er satt av til oppgaven, altså hva oppgaven kommer til å koste. Til slutt vil jeg presentere hvordan framdriften skal dokumenteres underveis og hvordan prosjektet organiseres.

2. Bakgrunn for prosjektet

Skyteknologi lar oss ta i bruk IT-infrastruktur på en mest mulig effektiv måte ved å sentralisere datakraft for deretter å distribuere denne rundt om i verden. Datakraften kommer fra datasenter og distribueres ut til sluttbrukerne av leverandørenes tjenester ved hjelp av internett. Dette gjør at vi nå kan se på IT-ressurser på samme måte som vann og elektrisitet. Bedrifter og organisasjoner kan unytte seg av og oppnå store fordeler ved å basere sin infrastruktur på skyteknologi, og jeg skal i denne oppgaven sammenligne plattformene til de to største leverandørene på markedet i dag; Amazon og Microsoft. Plattformene AWS og Azure kan defineres som et sett av tjenester som til sammen vil kunne brukes til å gjenskape den opprinnelige infrastrukturen til en bedrift, samtidig som det holder informasjon sikret og gir mulighet for mer effektiv dataprosessering. Disse skal sammenlignes for å kartlegge hvor brukervennlige de er, hva de koster, hvordan infrastrukturen kan skaleres og hvor tilgjengelige disse er. Alle som tenker å effektivisere sin tradisjonelle IT-infrastruktur vil få nytte av å lese denne oppgaven.

2.1 Beskrivelse av problemer og behov

Utgangspunktet for denne oppgaven vil være å gjøre det enklere for bedrifter å ta et valg i det infrastrukturen deres skal oppdateres. Overgangen til skybasert infrastruktur vil i store bedrifter være en kostbar og tidkrevende prosess, hvor man gjerne leier inn eksterne for å hjelpe til med omveltningen. Argumentene for å gå over til skybasert infrastruktur er mange – Man betaler for de ressursene man bruker, kapasiteten vil alltid kunne skaleres etter det behovet som finnes til enhver tid, man får tilgang til tjenester som gjør det enklere å få innsyn i data som ellers ville krevd store ressurser å få prosessert osv. Ved første øyekast er det altså ikke mye som tilsier at dette er noe man ikke har noen nytte av. Dette er grunnen til at fokuset i oppgaven vil ligge på å velge riktig plattform, framfor å bruke mye tid på å forklare fordelene med slik teknologi. Det er likevel noen ting man må tenke på når denne overgangen skal gjennomføres, og disse er knyttet til kostnader og lagring av sensitive opplysninger utenom bedriftens egne vegger. Dette skal vi se nærmere på i risikoanalysen.

2.2 Kort om dagens systemer og rutiner

Den tradisjonelle IT-infrastrukturen er typisk basert på et «on-site»-serverrom som tar seg av all dataprosessering for bedriften. I tillegg har man gjerne en andre, og kanskje tredje backup av disse dataene et annet sted som forsikring mot tap. Store bedrifter har som regel en intern IT-enhet som tar seg av drift av serverrom i tillegg til alt annen maskin og programvare som bedriften benytter seg av. Her eier ofte bedriftene maskinvaren selv og IT-enheten har ansvar for at det fungerer som det skal.

Utfordringene med en slik infrastruktur er flere. De fleste har nok opplevd å måtte vente på å få tilgang til et nettsted på grunn av stor pågang, for eksempel når man skal inn og sjekke skattekortet sitt, eller kjøpe billetter til et stort arrangement. Årsaken til denne ventetiden er at antall forespørsler til serverne er større enn det de klarer å håndtere innen akseptabel tid og nettleser må da vente på svar. Løsningen på dette har tidligere vært å skalere opp ved å installere flere webservere slik at man er rustet for slik pågang. Dette fungerer, men nå sitter bedriften med en maskinpark som er skalert for større pågang enn det de ellers har gjennom året, eller døgnet. Skyteknologi løser dette ved hjelp av automatisk skalering. Når etterspørselen blir større tas det i bruk flere servere i datasenteret slik at kapasiteten blir stor nok til å håndtere pågangen. Når behovet går ned igjen skaleres dette ned automatisk. Dette henger igjen sammen med kostnader; nå slipper bedriften å investere i kraftigere maskinvare for å håndtere den periodevis økte pågangen, men betaler kun for den kapasiteten de behøver til enhver tid. Azure og AWS har hver sin prismodell som vil være et sentralt element i denne oppgaven.

Sikkerhet og oppetid er også noe vi må se på i denne sammenheng. Med intern drift av egen maskinpark er man mer utsatt for nedetid hvis noe skulle skje med serverne. Man har selvfølgelig løsninger på dette fra før med bruk av backup-servere, men også her kan det bli nedetid i tillegg til at man selv er ansvarlig for å kjøpe inn nytt utstyr og sørge for oppetid. I datasentrene ligger data på flere servere, slik at opptil flere servere kan bli skadet uten at man mister informasjon, i tillegg til at det ikke koster noe for bedriften å ordne opp i dette. Det er leverandøren som må håndtere slike hendelser og dekke kostnadene når maskinvaren ryker.

3. Prosjektmål

Dette kapittelet av forstudierapporten tar for seg hva jeg ønsker å oppnå med prosjektet, med utgangspunkt i beskrivelsen av dagens systemer på slutten av forrige kapittel. Dette kapittelet blir brukt til å forklare motivasjonen og resultatet av at en bedrift går gjennom en oppdatering av sin infrastruktur. Det tas altså ikke utgangspunkt i Azure vs Amazon her, men heller årsaken til at man velger å gjøre en endring.

3.1 Effektmål

Punktene under beskriver motivasjonen en bedrift kan ha for å oppdatere sin IT-infrastruktur til å ta i bruk skyteknologi fra en ekstern leverandør:

- Øke sikkerhet mot tap av data
- Oppnå lavere kostnader knyttet til drift av IT-systemer
- Realisere tjenester for innblikk i kompliserte data som tidligere ville blitt for kostbart

3.2 Resultatmål

Effektmålene i forrige punkt beskrev utgangspunktet for en gitt bedrift med en utdatert IT-infrastruktur, som vil kunne oppnå fordeler ved å ta i bruk skyteknologi. Punktene under sier noe om hva som skal gjøres for å oppnå effektmålene eller hvorfor de blir oppfylt ved å ta i bruk teknologien.

- Lagre data på flere servere i datasenter hvor leverandør er ansvarlig for oppetid
- Benytte seg av auto-skalerting slik at kapasiteten hele tiden møter bedriftens behov.
- Ta i bruk tjenester for data mining og BI som en del av det virtuelle nettverket for enklere å få innblikk i kompliserte og store datamengder.

3.3 Prosessmål

Siden dette er et studieprosjekt er det naturlig at jeg nevner prosessmålene for prosjektet. Dette er målene for prosessen jeg skal gjennom for å produsere sluttproduktet som er bacheloroppgaven min. Siden jeg skriver oppgaven alene, vil disse målene kun gjelde for meg.

- Oppnå en karakter i faget som gjenspeiler det nivået jeg har lugget på ellers i løpet av studietiden min på NTNU.
- Oppnå økt kompetanse på feltet «Cloud Computing» da dette er noe som interesserer meg og det hører inn under det fagfeltet jeg skal jobbe med etter endt studium.
- Få god oversikt over forskjellene i produktene som Amazon og Microsoft tilbyr.
- Oppnå egenutvikling i form av økt kompetanse på å gjennomføre et omfattende prosjekt med store krav til dokumentasjon og kommunikasjon med en veileder.

3.4 Prosjektets omfang

Prosjektet skal ikke gå i dybden på fordeler og ulemper ved å ta i bruk skyteknologi. Å forklare hva skyteknologi er og hvorfor man skal bruke det vil altså ikke bli tillagt mer tid enn å gi det fokus i innledningen av oppgaven. Fokuset vil ligge på å sammenligne produktet til de 2 største aktørene på markedet, og jeg skal se nærmere på modeller for prising, hvordan infrastruktur kan skaleres og hvor tilgjengelige disse tjenestene er. Jeg skal også gjennomføre en praktisk del hvor et PHP-system blir implementert på begge plattformer. Formatet på oppgaven vil bli en informativ tekst som mot slutten går over til en praktisk gjennomføring med framgangsmåter. Oppgaven avsluttes med konklusjon som oppsummerer plattformenes styrker og svakheter.

3.5 Prosjektets milepæler og hovedaktiviteter

Navn aktivitet/milepæl	Periode/dato	Budsjettert timeantall
Forstudierapport	20.01-28.01	35
Forstudierapport og problemstilling sendes inn for vurdering	28.01	
Veiledningsmøter	20.01-20.05	6
Prosjekthåndbok (Timelister, møteinnkallinger, møtereferat og ukesrapporter med status)	20.01-20.05	70
Arbeidet med hovedoppgaven starter	29.01	
Bygge opp innholdsfortegnelse/struktur	29.01-01.02	5
Informasjonssøk	01.02-01.05	30
Oppgavedel: Teori og trender	02.02-15.03	100
Oppgavedel: Implementering	20.02-05.03	60
Oppgavedel: Resultater	15.03-10.04	120
Oppgavedel: Diskusjon	05.04-26.04	50
Oppgavedel: Konklusjon	01.05-20.05	10
Presentasjon av prosjektoppgaven	02.6	30
Innlevering Bacheloroppgave 2021	20.05	

4. Interessenter og rammebetingelser

4.1 Interessentanalyse

a. Interessenter for prosjektet:

- Oppgavestiller er forfatter (selvvalgt oppgave)
- Resultatet skal godkjennes av veileder og sensor ved NTNU
- I det daglige underveis i prosjektet er det forfatter og veileder som berøres av oppgaven
- Kommende studenter ved NTNU vil kunne ha bruk for å vite noe om dette prosjektet. Det samme gjelder for bedrifter som vil kunne benytte seg av informasjonen som blir formidlet i oppgaven
- Kun veileder har behov for å vite noe om prosjektforløpet
- Arbeidet utføres av forfatter alene
- Oppgaven leveres av forfatter. Veileder bidrar med innspill, oppfølging og retting av dokumentasjon underveis

b. Suksesskriterier for interessentene:

- Veileder må få regelmessig tilgang til dokumentasjon og framdriftsplan underveis i prosjektet. Han er avhengig av at jeg som forfatter tar eierskap i oppgaven og leder dialogen mellom oss.
- Sensor og veileder må få tilgang til delinnleveringer og ferdig oppgave innen fristen.
- Jeg må følge framdriftsplanen og sørge for at det legges ned nok timer hver uke for å komme i mål innen fristen.
- Kommende studenter og bedrifter som ønsker å lære noe av oppgaven er avhengige av at oppgaven er godt strukturert, holder fokus på problemstillingen og gir svar på de problemer som blir presentert i oppgaven. Det skal være tydelig rød tråd og det skal komme tydelig fram hvilke styrker og svakheter som finnes i produktene til leverandørene. Den praktiske biten bør forklares godt og bilder kan bli nødvendig når framgangsmåte skal presenteres.

4.2 Rammebetingelser

Hensikten med dette kapittelet er å presentere de absolutte krav som gjelder for prosjektets gjennomføring. Disse vil ha betydning for prosessen videre og hvordan prosjektet skal organiseres.

- Absolutt siste frist for innlevering av oppgaven er 20.Mai. Da skal også all dokumentasjon være ferdigstilt: timelister med ukesrapport, fullført framdriftsplan og alle møtereferater skal være arkivert med tilhørende møteinnkallinger. All denne dokumentasjonen skal sys sammen til en prosjekthåndbok. Her følger jeg mal utlevert av NTNU.
- Krav til kostnadsramme: Prosjektet skal i utgangspunktet ta 500 timer å gjennomføre. Her har jeg en slingring på +/-10% som jeg må forholde meg til. Havner jeg utenfor dette intervallet må jeg ha en god grunn for hvorfor det skjedde, og det må dokumenteres. Ellers vil det påvirke karakteren negativt.
- Det skal lages en presentasjon av bacheloroppgaven. Denne skal presenteres for emneleder, veileder og medstudenter. Mest sannsynlig over video i slutten av Mai måned.
- Timelister skal fylles ut med ukesrapport og leveres til veileder på slutten av hver uke.

5. Kritiske suksessfaktorer

5.1 Suksessfaktorer

Faktorer som er avgjørende for om prosjektets resultat blir vellykket eller ikke:

- For at denne oppgaven skal være nyttig for en evt. bedrift så må teorien legges fram på en måte som gjør at eksisterende IT-enhet greier å knytte nåværende systemer opp mot den teknologien jeg presenterer. Det skal være mulig for en med forståelse for IT-infrastruktur å se hvilke problemer som løses ved å ta i bruk slik teknologi.
- Den røde tråen er veldig viktig i et slikt prosjekt. Alt av teknologi og løsninger må hele tiden forklares, samtidig som fokuset skal ligge på å trekke fram fordeler og ulemper med de to plattformene, som er hovedfokuset i denne oppgaven.
- Jeg må sette av godt med tid til implementering av PHP-systemet slik at jeg får dokumentert dette på en god måte. Dette av samme årsak som i første punkt; det blir enklere for IT-personell å relatere til det som skjer og dermed enklere å gjøre seg opp en mening om hvilken plattform man synes er mest hensiktsmessig.
- For å oppnå så godt resultat som mulig vil det også være viktig at jeg følger de råd som kommer fra veileder når jeg får tilbakemelding på dokumenter som produseres underveis.
- Passe på timebruken slik at denne ender opp på +/-10% av 500 timer med alt arbeidet.

6. Risikoanalyse

Tidligere I denne rapporten har jeg presentert fordeler og hva man kan oppnå ved å fornye sin IT-infrastruktur til bruk av skyteknologi, på engelsk kalt cloud computing. Som vi har sett så er det mange fordeler ved en slik overgang, og selv om den kan være svært kostbar hvis nåværende systemer er store og omfattende, så vil det gi avkastning over tid, i tillegg til nye muligheter man tidligere ikke hadde. Jeg skal nå presentere en risikoanalyse for prosjektet, hvor vi skal se litt på hva en bedrift må være klar over når en slik overgang skal gjennomføres.

- a. Bedriften som har valgt å benytte seg av en ekstern leverandør mister kontroll over eiendelen sine. Overvåking og analyse av datatrafikk som man tidligere kunne gjøre direkte mot sin egen infrastruktur vil nå måtte gjøres av interne tjenester i infrastrukturen til leverandør. Dette kan føre til problemer når det for eksempel kommer til oppbevaring og overvåking av personopplysninger.
- b. Bedriften kan ikke lenger bekrefte at data er slettet fra serverne, siden de nå ikke sitter på lagringsenhetene selv. Dette kan medføre samme problemstilling som i forrige punkt. Hvis man benytter seg av flere leverandører vil risikoen for ufullstendig sletting av data kunne øke, da oversikten synker ytterligere.
- c. Man kan være mer sårbar mot angrep hvis det skulle skje at en angriper får tilgang til skylegitimasjon. Da har angriper tilgang til implementeringen av infrastruktur til bedriften (Det er som regel kun leverandør-admin som har tilgang nettverk, systemer og applikasjoner) og konsekvensene vil kunne være store.
- d. Administrasjon av skyen krever kompetanse som nåværende IT-enhet ikke besitter. Dette kan medføre ekstra kostnader som følge av krav til opplæring eller videreutdanning. Innleie av konsulenter vil også kunne være en kostbar løsning.
- e. Man er også utsatt for tap av data som ligger lagret i skyen (I datasenteret). Fysiske katastrofer kan være en årsak til dette; for eksempel brann i datasenter. Hvis bedriften har lastet opp krypterte data til skyen, vil dette kunne medføre problemer hvis man ikke besitter krypteringsnøkkel når data skal gjenopprettes etter en slik katastrofe. Konsekvensen av en slik katastrofe vil variere etter hvor kritiske data det er snakk om. Det blir derfor nødvendig å dele opp i kategorier;

- e.1. Kundetjenester (Informasjon om kunder, resultat av spørreundersøkelser, serviceforespørsler osv.)
- e.2. Produktutvikling (Patenter, forskning osv.)
- e.3. Ledelse (finansrapporter, forhandlinger, interessentanalyser osv.)
- e.4. HR (Rekruttering, informasjon om lønn, helseinformasjon osv.)
- e.5. Salg (Tilbud, kontakter, prisinformasjon, osv.)
- e.6. Dataadministrasjon (Prosjekt-arkiv, nettverksarkitektur, passord osv.)
- e.7. Finansiell administrasjon (Informasjon om profitt, rapporter, investeringer osv.)

Nedenfor har jeg regnet ut total risiko. Man kommer fram til den ved å multiplisere sannsynligheten for at noe inntreffer med konsekvensen det får. Skalaen strekker seg fra 1 til 10 som gir 100 som høyeste total risiko.

Faktor	Sannsynlighet	Konsekvens	Total risiko
a	10	3	30
b	10	4	40
c	4	7	28
d	8	2	16
e1	5	8	40
e2	5	9	45
e3	5	7	35
e4	5	5	25
e5	5	5	25
e6	5	7	35
e7	5	9	45

7. Kost/nytte-analyse

En kost/nytte analyse skal gi leser et klart bilde av hva prosjektet eller systemet kommer til å koste. Ofte er en slik analyse laget for å gjøre det enklere å avgjøre om et prosjekt skal gjennomføres eller ikke. I denne sammenhengen må vi derfor sammenligne dagens infrastruktur med alternativet som denne oppgaven presenterer. Altså hvordan kan man spare, evt. kommer man til å tape penger på å flytte infrastrukturen sin ut i skyen. Siden dette er en generell rapport som ikke tar utgangspunkt i en gitt organisasjon vil tallene være hentet fra eksempler på nett.

Effektmål:

- Redusere bedriftens kostnader knyttet til IT-drift
- Øke sikkerheten for sensitive data
- Oppnå skalering av infrastruktur for å unngå sløsing av ressurser eller tidvis for liten kapasitet
- Gjøre det mulig å ta i bruk avansert dataanalyse
- Frigjøre ressurser som kan prioriteres i forretning framfor IT

7.1 Bortfall av direkte kostnader

De direkte kostandene ved bruk av tradisjonell IT-infrastruktur består typisk av innkjøp og implementering av maskinvare, vedlikehold av maskinparken og oppdatering av programvaren og support og drift av infrastrukturen. Det finnes ingen direkte måte å måle avkastningen på en migrering til skyen, da faktorene er såpass mange. Hvis man kun er ute etter å senke kostnadene man har med det tradisjonelle oppsettet, er det ikke sikkert dette vil skje. Det er likevel andre grunner som gjør en slik migrering attraktiv for mange bedrifter. Dette kan være ønsket om å frigjøre seg fra eget datasenter og la en ekstern aktør ta seg av drift av infrastrukturen, korte ned på tiden det tar å implementere og distribuere programvare eller ønsket om å forbedre forretningen ved å frigjøre ressurser.

7.2 Estimerte kostnader

Som nevnt tidligere så opererer leverandørene med egne prismodeller, men alle som leverer skybasert infrastruktur til sine kunder tilbyr også prising som følger forbruket. Dette kalles ofte «pay as you go», og betyr at kunden betaler for den kapasiteten den trenger til enhver tid, på samme måte som privatpersoner betaler strømgregningen sin. Det er derimot flere kostnader

knyttet til overgangen til skyen. Vi snakker her om implementeringen, eller migreringen man må gjennom for å gjenskape dagens systemer i skyen slik at man kan fortsette driften som vanlig. Når vi skal sammenligne kost/nytte mellom tradisjonell og skybasert infrastruktur vil vi derfor komme til å se høyere kostnader de første årene knyttet til migreringen.

7.3 Sammenligning kost/nytte

Dagens systemer	År 1	År 2	År 3	År 4	År 5
Maskinvare kostnad	7400000	7400000	7400000	7400000	7400000
Maskinvare kapital	600000	600000	600000	600000	600000
Programvare kapital	900000	900000	900000	900000	900000
Total kapital	1500000	1500000	1500000	1500000	1500000
Total kostnad	8100000	8100000	8100000	8100000	8100000
Total utgift	9600000	9600000	9600000	9600000	9600000

Overgangen til sky	År 1	År 2	År 3	År 4	År 5
Arv kostnad	7600000	3800000	0	0	0
Oppsett kostnad	300000	100000	0	0	0
Migrering kostnad	2200000	3400000	0	0	0
Managed services	100000	1200000	1600000	1600000	1600000
IaaS	300000	800000	1100000	1100000	1100000
Total kapital	0	0	0	0	0
Total kostnad	10500000	9300000	2700000	2700000	2700000
Total utgift	10500000	9300000	2700000	2700000	2700000

Penger spart	År 1	År 2	År 3	År 4	År 5
Total kapital	1500000	1500000	1500000	1500000	1500000
Total kostnad	3100000	1900000	4700000	4700000	4700000
Total besparelse	1600000	400000	6200000	6200000	6200000

Analysen over er et eksempel hentet fra hjemmesiden til konsulentselskapet Capgemini

(<https://www.capgemini.com/2020/03/the-roi-of-moving-to-the-cloud/>)

8. Retningslinjer og standarder

8.1 Krav til dokumentasjon

Oversikt over de dokumenter som skal leveres i løpet av prosjektet:

- Forstudierapport (Leverer 1.utkast for tilbakemelding 28.01)
- Timeliste med ukesrapport (Veileder gis tilgang til Drive-mappe med alle timelister)
- Arbeidsavtale som skal signeres av meg og veileder før det sendes inn til instituttet
- Framdriftsplan som skal vise planlagt start og slutt for de ulike aktivitetene som inngår i prosjektet. Se kapittel 3.5
- Møteinnkallinger skal skrives før hvert møte
- Møtereferat skal skrives etter hvert møte. Referat og innkallinger skal inn i prosjekthåndboken som leveres sammen med hovedoppgaven innen 20.Mai
- Kravspesifikasjon for PHP-system som skal implementeres på begge plattformer

9. Prosjektorganisering

Siden dette er en selvvalgt oppgave jeg løser alene vil ikke dette kapittelet bli like omfattende som det ville blitt med ekstern oppdragsgiver og arbeid i gruppe. Jeg vil derfor kort oppsummere rollene som gjelder for prosjektet.

Oppdragsgiver: forfatter, Ådne Helmersen

Veileder: Helge Hafting

Oppdragsgiver med bistand fra veileder fungerer som styringskomite for prosjektet. Veileder alene vil ha ansvaret for kvalitetskontroll underveis i dette prosjektet. Sensor og veileder vil stå for kvalitetskontroll av sluttresultatet. Forfatter skal ha rollen som oppdragsgiver gjennom hele prosjektet og har ansvar for all dialog med veileder.