

Anskaffelsesstrategi for tungregnemaskin og lagringssystem for vitenskapelige data

Innhold

1. BAKGRUNN OG FORMÅL MED ANSKAFFELSENE	4
2. OM BEHOVET	4
2.2 Spesifikt for HPC	5
2.3 Spesifikt for datalagring	6
2.4 Forholdet mellom anskaffelsene	9
3. ØKONOMI OG ORGANISASJON	9
3.1 Økonomisk effekt	9
3.2 Organisasjonsmessig effekt	9
3.3 Kostnader ved anskaffelsene	9
3.3.1 Prosjektkostnader	9
3.3.2 Investering HPC	10
3.3.3 Investering datalagringsutstyr	10
3.4 Finansiering	11
3.5 Avhengighet til andre anskaffelser og prosjekter	11
3.6 Mottaksprosjekt	12
4. PLANLAGT GJENNOMFØRING AV ANSKAFFELSENE	12
4.1 Organisering av anskaffelsene (HPC og Datalagring)	12
4.2 Tidsplan HPC	13
4.3 Tidsplan Datalagring	13
4.4 Fremtidige behov	13
5. MARKED OG KONKURRANSE	14
5.1 Oversikt over mulige leverandører	14
5.1.1 HPC	14
5.1.2 Lagringssystem	14
5.2 Regelverket for offentlige anskaffelser	15
5.3 Hovedregel om anskaffelsesprosedyre	15
5.3.1 Anskaffelse av HPC	15
5.3.2 Anskaffelse av datalagring	16
5.4 Kontraktstype og betingelser	16
5.4.1 Generelt	16
5.4.2 HPC	16
5.4.3 Lagringssystem	17
5.5 Samfunnsansvar	17
5.5.1 Miljø	17
5.5.2 Etisk handel (varekjøp)	17

5.5.3 Sosialt ansvar (tjenestekjøp)	17
6. RISIKOANALYSE OG KRITISKE SUKSESSFAKTORER	18
6.1 Kritiske suksessfaktorer	18

1. BAKGRUNN OG FORMÅL MED ANSKAFFELSENE

UNINETT Sigma2 (Sigma2) har ansvaret for innkjøp og drift av den nasjonale e-infrastrukturen for beregningsvitenskap i Norge og tilbyr tjenester innen tungregning (HPC-maskiner, Notur) og datalagring (NorStore). Sigma2 tilbyr tjenestene til individer og grupper involvert i forskning og utdanning ved norske universiteter og høyskoler, samt andre organisasjoner og prosjekter med offentlig finansiering.

HPC og datalagring har, på bakgrunn av teknologisk utvikling og kostnad for vedlikehold, en levetid på 4-5 år. Formålet med denne anskaffelsen er å starte utskifting av e-infrastrukturen for fortsatt å sikre brukerne et tilbud som dekker deres behov når det gjelder funksjonalitet og kapasitet basert på en formålstjenlig og moderne teknologi til en lavest mulig pris.

Sigma2s styre har besluttet at dagens fire tungregnemaskiner (HPC) skal reduseres til to. De to maskinene med de største vedlikeholdskostnadene skal erstattes først.

For å oppnå en integrert, oversiktlig og brukervennlig infrastruktur hvor man reduserer behovet for å flytte data, samt legger grunnlaget for utvikling av nye brukervennlige tjenester, har styret besluttet at den nasjonal infrastrukturen skal baseres på en datasentrisk infrastruktur hvor det er tett samordning mellom data og utstyr for tungregning og øvrige tjenester.

Denne anskaffelsesstrategien omfatter anskaffelse av:

- Den første av de to tungregnemaskinene (HPC), heretter kalt A1.
- En liten maskin, kalt "L", for spesielle formål, for eksempel for å teste og utvikle parallelle applikasjoner.
- Et lagringssystem som skal erstatte dagens lagringssystem (NorStore).

2. OM BEHOVET

2.1 Behovsundersøkelse

Som del av forberedelse til denne anskaffelsen er det gjennomført behovsundersøkelse (som en nettundersøkelse), samt gjennomført brukermøter med de viktigste brukermiljøene.

Tilbakemeldingene fra behovsundersøkelsen og besøksrunden er at forskerne som benytter dagens systemer i all hovedsak kan få sitt behov for regnekraft og datalagring dekket gjennom en *utvidelse og fornyelse av de teknologiske plattformer som tilbys i dag*.

Det ble imidlertid gitt tilbakemelding på at brukermiljøene er for dårlig for dem som har behov for å teste og utvikle parallelle applikasjoner, og at det må tas grep for å forbedre dette. Videre er det og uttrykt ønske om en bedre integrasjon mellom HPC-systemer og datalagring for å oppnå bedre brukervennlighet. En slik integrasjon vil og gjøre det enklere å tilby nye tjenester som f.eks. dataanalyse.

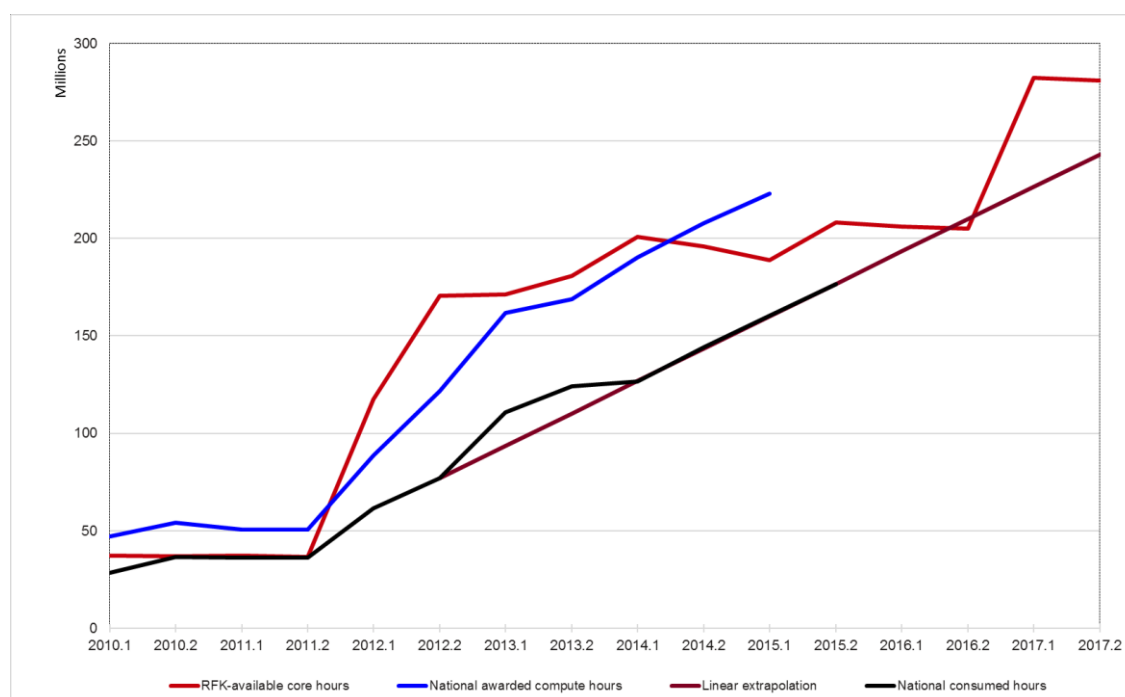
2.2 Spesifikt for HPC

Tabell 1 viser en oversikt over dagens HPC maskiner.

HPC maskin	Type	Kapasitet (cpu-timer <u>pr. år</u>)	Ytelse (Tflop)	Produksjonsstart	Årlig vedlikeh. (NOK)
Abel	Kapasitet	76 413 480	181,6	Oktober 2012	447 188
Hexagon	Kapabilitet	102 807 360	107,9	April 2012	2 654 175
Vilje	Kapabilitet	113 012 760	268,3	Oktober 2012	4 372 000
Stallo	Kapasitet	78 804 960	195,7	Oktober 2012	390 000

Tabell 1. Dagens HPC ressurser, kapasitet, ytelse, oppstartsdato og årlige vedlikeholdskostnader

Som angitt i Tabell 1, er alderen på dagens HPC-maskiner passert 3 år. Fra og med år 4 må det betales betydelige vedlikeholdskostnadene, slik at det fra et pris/ytelses-perspektiv ikke er lønnsomt å beholde disse maskinene. Planen er derfor – som del av denne anskaffelsen – å erstatte to av maskinen med en ny (A1) i begynnelsen av 2017 og de to andre gjennom enda en ny (B1) i begynnelsen av 2018. Bakgrunnen for dette er den gjennomførte behovskartleggingen, kartlegging av bruksmønster og projiserte kapasitetsbehov (se kurven nedenfor).



Figur 1. Kapasitets- og forbruksutvikling pr. periode

En planlagt økt satsing på avansert brukerstøtte forventes å gi økt behov for både beregnings- og lagringsressurser.

A1 skal erstatte de to HPC-maskinene (Hexagon og Vilje) som har høyest vedlikeholdskostnader og skal ha en ytelse og kapasitet beregnet ut fra det samlede behovet disse to maskinene dekker, inkludert en stipulert økning fram til mars 2018.

I tillegg har brukerundersøkelsen dokumentert et behov for å anskaffe en mindre maskin, heretter kalt "L" (Lavterskelmaskin), som skal benyttes til undervisning, samt til programvareutvikling og testing av

(parallele) applikasjoner. "L" skal være mest mulig lik produksjonsmaskinen A1, slik at det er enkelt å flytte mellom disse to.

Maskinene som skal erstattes er såkalte kapabilitetsmaskiner, (se Tabell 1). Dette er maskiner som har strenge krav til forsinkelse og båndbredde i kommunikasjonen mellom prosessorkjerner. Derfor skalerer den aggregerte båndbredden lineært i forhold til antall kjerner på slike kapabilitetsmaskiner. I dag kjører HPC-maskinene "jobber" som benytter opp til 4000 kjerner samtidig. Vår kravspesifikasjon og evaluering av tilbudene i forhold til denne, vil avgjøre om A1 og skal være en kapabilitetsmaskin eller en kapasitetsmaskin.

Det er beregnet at A1 må ha en kapasitet på ca. 200 millioner CPU-timer (pr. periode), noe som vil tilsvare ca. 50 000 kjerner med en utnyttelse grad på 0,9. [Ytelsen er foreløpig beregnet til ca. 1 PFLOPS.](#) L-maskinen vil være i størrelsesorden 2-3% av [denne A1](#), dvs. ca. 1000 kjerner.

Som del av denne anskaffelsen skal det anskaffes lagring for /work, dvs. lagringen som benyttes når beregning pågår. Når det gjelder lagring forøvrig (/home og /project) skal det anskaffes av lagringsanskaffelsen.

Det er behov for å fortsatt benytte *Intel*-arkitektur som basis for de nye systemene.

Bakgrunnen for dette er at produktmerket *Intel* fremdeles er den helt dominerende markedsaktøren innenfor området. Både Sigma2s eksisterende omfattende programvareportefølje, samt våre planlagte nye anskaffelser av programvare, vil fungere best og billigst på denne arkitekturen. Et eventuelt skifte bort fra *Intel*-arkitektur vil bety at kompetansen på denne arkitekturen opparbeidet gjennom mange år både i brukermiljøene og hos IT-personell mister mye av sin verdi og kompetanse på ny arkitektur må etableres. Fortsatt bruk av Intel-arkitektur vil derfor, etter vår mening, *sikre mest mulig effektiv ressursbruk* slik lov om offentlige anskaffelser § 1 anviser.

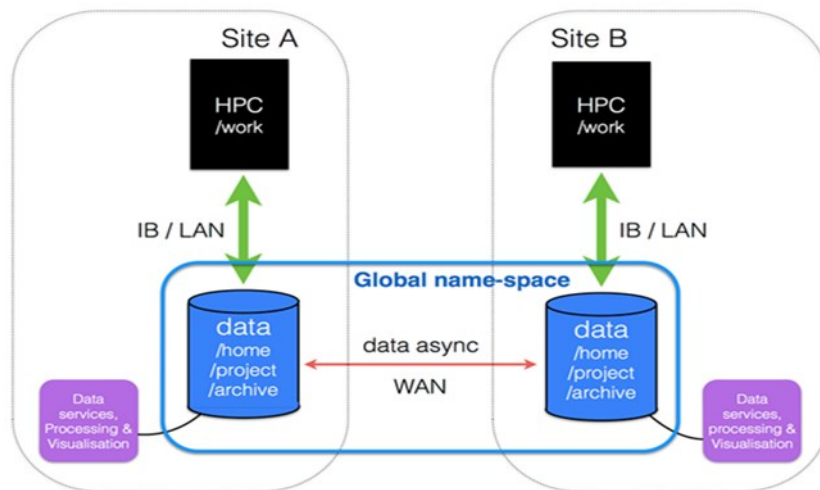
Det forventes et teknologiskifte fra *Intel* i perioden mellom anskaffelsen av A1/L og ny HPC-maskin, kalt B1. Teknologiskiftet er forventet å medføre en vesentlig reduksjon i pris/ytelse forholdet og nye muligheter når det gjelder forenklet og effektiv bruk av akseleratorteknologi (GPU), men Intel har ikke gitt informasjon om hvor stor reduksjon i pris/ytelse som kan forventes. I og med at B1 er planlagt anskaffet bare et år etter A1, er det naturlig å ta hensyn til dette når A1 spesifiseres, spesielt når det gjelder antall noder¹ med akseleratorteknologi.

2.3 Spesifikt for datalagring

Nåværende lagringssystemer for vitenskapelige data og nødvendig lagring tilknyttet HPC-maskinene er anskaffet uavhengig av hverandre og er i dag helt adskilte infrastrukturer. Data og datatilgjengelighet er avgjørende for enhver tjeneste i den nasjonale infrastrukturen. Med den betydelige veksten av data, og da spesielt ustrukturerte data, vil det ikke lenger være forsvarlig eller håndterlig å måtte overføre store mengder data over forskningsnettet.

Den foreslåtte arkitekturen prioriterer å levere dataene med lavest mulig forsinkelse til de tilknyttede ressursene som implementerer tjenestene. Det medfører at den nye tungregnemaskinen og lagringssystemet må plasseres på samme fysiske lokasjon. Se Figur 2.

¹ En node består av flere cpu'er med mange kjerner

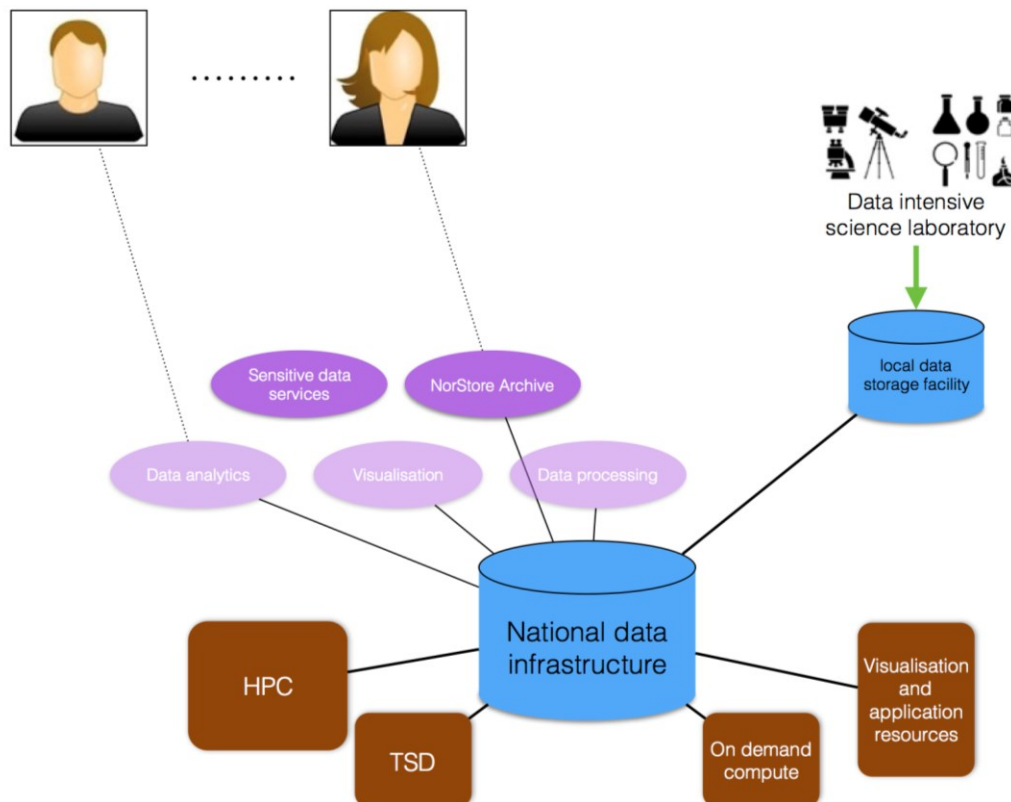


Figur 2: Samme fysiske lokasjon for HPC og datalagring, men forskjellig lokasjon for A1 og B1, med respektive datalagringseenhet.

Overordnede krav til lagringssystemet er at det skal være *skalerbart*, *pålitelig* og *fleksibelt*.

Skalerbart betyr at det skal være enkelt og rimelig å utvide lagringskapasiteten, *pålitelig* betyr at data ikke tapes/mistes og *fleksibelt* betyr at det skal være enkelt å utvikle nye tjenester.

Figur 3 viser et konseptuelt bilde av den datasentriske arkitekturen med lagringssystemet, øvrige ressurser og de ulike tjenestene knyttet sammen. Det vil også åpnes for å tilknytte eksterne datakilder (eksempelvis laboratorier). Tjenestene vil stille ulike krav til lagringssystemet og de tilhørende ressursene, hvor beregningstjenesten er én av disse tjenestene.



Figur 3. Datasentrisk arkitektur

Anskaffelsen må omfatte og spesifisere kravene fra alle nåværende tjenester og det man vet om fremtidige tjenester. Brukerundersøkelsen, intervjuer og erfaring fra de siste fem års drift av NorStore-tjenesten utgjør grunnlaget for disse kravene.

Det er behov for to typer datalagring; i) ytelseslagring og ii) replikert kapasitetslagring.

i) Ytelseslagring er optimert for å levere høy lese- og skriveytelse, og vil i hovedsak være en ressurs som benyttes av lagringssystemet når data skrives eller modifiseres fra eksempelvis HPC-ressursen (typisk prosjektdata). Mindre filer, som gjerne aksesseres hyppig (både leses og skrives), vil trolig også ligge på denne lagringsressursen (eksempelvis programkode og konfigurasjonsdata som ligger under /home).

ii) Kapasitetslagring er optimert for høy volumtetthet men med tilsvarende leseytelse som ytelseslagring. Dette lageret vil blant annet inneholde større datafiler og arkiverte data. Majoriteten av forskningsdata (volumdata) vil lagres på denne ressursen.

Lagringsarkitekturen som er valgt er basert på en distribuert infrastruktur. Tradisjonelle NAS og SAN systemer har vært basert på RAID teknologi hvor mye av logikken er innebygd i utstyret.

Tabell 2 gir en oversikt over lagringsbehov i 2015 for både Notur og NorStore, bortsett fra /work (som brukes til mellomlagring under beregningskjøringer ved HPC anleggene). Lagringsressurser for /work for ressursene A1 og L kjøpes som del av HPC anskaffelsen fordi det ofte kreves en ytelse som overgår de øvrige dataene. Denne lagringen skal være knyttet opp direkte mot HPC anlegget slik at regneressursene ikke forsinkes av manglende lagringsytelse.

Lagringsområder i eksisterende infrastruktur	PiB (netto)	Type datalagring
/project	1	ytelse
/project	2	kapasitet
/archive	0.1	kapasitet
/project (notur)	1	ytelse
/home	0.5	ytelse
/project (tape)	1.5	kapasitet

Tabell 2. Estimert fordeling av dagens ressursbehov (i 2015)

Installasjon av lagringssystemet, som må skje samtidig med leveranse av A1 og L. Det vil skje en gradvis utbygging der man initielt vil dekke A1 og L sitt behov, fram til en total netto kapasitet på 10 PiB. Fordelingen av kapasitetsbehov mellom de to typene datalagring er basert på estimater og en forventning om at den nye infrastrukturen skal være mer kostnadseffektiv. Utgangspunktet for estimatet er dagens (2015) forbruk, minus en liten besparelse ved å samlokalisere data i én infrastruktur. (Summen av PiB i Tabell 2 er derfor ikke lik summen av 2015 netto kapasitetsbehov i Tabell 3). Erfaringsmessig fordobler volumet av forskningsdata seg på ca. 2.5 år.

Med utgangspunkt i dette, samt omsøkte, tildelte og forbrukte ressurser er det foretatt en beregning av hvor stort data volumet vil være ved forventet produksjonsstart i første kvartal av 2017 (se Tabell 3). Deler av lagringsbehovet som i dag er dekket av lokale prosjektområder ved HPC-anleggene er også inkludert.

Type datalagring	2015 Netto kapasitetsbehov	2017 Netto (Brutto) kapasitetsbehov	2019 Netto (Brutto) kapasitetsbehov
Ytelseslagring	2 PiB	3 (15) PiB	4 (20) PiB
Replikert kapasitetslagring	3.6 PiB	7 (20) PiB	14 (39) PiB

Tabell 3. Oversikt datalagringsressurser

2.4 Forholdet mellom anskaffelsene

Anskaffelse av HPC og datalagring må gjennomføres som to separate anskaffelser. Den viktigste årsaken er, at markedet for datalagring preges av svært mange leverandører med til dels ulike løsninger. Derfor antar vi at det vil kreve *dialog* med de mest aktuelle datalagringsleverandørene for å kunne velge den beste løsningen. Dialog med leverandørene krever en annen anskaffelsesprosedyre, *konkurranse med forhandling*, i motsetning til HPC-anskaffelsen som kan gjennomføres som *åpen anbudskonkurranse*.

Men anskaffelsene har innbyrdes avhengighet. Datalagringsløsningen omfatter deler som HPC maskinen bruker (/home og /project). Derfor må anskaffelsen av HPC- og datalagring skje synkronisert. For å sikre synkronisering gjennomføres anskaffelsene i samme prosjektet.

3. ØKONOMI OG ORGANISASJON

3.1 Økonomisk effekt

En reduksjon av antall HPC maskiner fra to til fire forventes å gi reduserte kostnader til drift og brukerstøtte, budsjettmessig stipulert til ca. 20% i forhold til dagens kostnader. Det er og en målsetting å øke samlet utnyttelsesgrad ~~av systemene både ved at det er enklere å styre denne når man har færre systemer og~~ ved å legge opp til en mer trinnvis utbygging slik at forskjellen mellom anskaffet kapasitet og behov blir mindre enn ved tidligere anskaffelser. I tillegg vil en reduksjon i antall maskiner rent teoretisk gi grunnlag for en bedre utnyttelse i forhold til forbruksveksten. Samlet vil dette føre til at vi får mer forskning ut av de samlede midlene som benyttes til investering i og drift av utstyr.

Det forventes at mer avansert brukerstøtte vil øke behovet for tjenester og øke antall brukere innen e-infrastruktur fordi brukerne kan få bedre hjelp til å løse sine utfordringer.

3.2 Organisasjonsmessig effekt

I dagens modell skjer brukerstøtte og drift separat for hvert av de fire systemene. Ved overgangen til to systemer blir det nødvendig å etablere en ny modell der personell fra de fire universitetene skal samarbeide om brukerstøtte og drift. Dette skal utredes i et eget prosjekt og det forventes at det blir behov for et færre personer til disse oppgavene. Alle medarbeidere ved universitetene som allerede arbeider med drift skal imidlertid sysselsettes videre, men noen av dem må regne med nye arbeidsoppgaver. Nye arbeidsoppgaver vil spesielt være knyttet til avansert brukerstøtte, noe som vil kreve annen type kompetanse. Det betyr at samlede lønnskostnader ikke reduseres, som følge av beslutningen om å redusere antallet HPC-maskiner fra fire til to.

3.3 Kostnader ved anskaffelsene

3.3.1 Prosjektkostnader

Begge anskaffelsene vil bli gjennomført som et felles prosjekt med deltakelse fra egne ansatte, innleid personell fra de fire universitetene for å sikre tilstrekkelig teknisk kompetanse og innleie av juridisk innkjøpskompetanse fra KPMG. Tabell 4 og 5 viser budsjettene for de to anskaffelsene.

Kostnader HPC-anskaffelsen		
	2016 i MNOK	2017
UNINETT Sigma2	2.4	0.5
Innleie Konsulent	0.5	
Innleie Universitet (9 mnd.)	0.7	0.2
Sum	3.6	0.7

Tabell 4. Prosjektkostnader HPC

Kostnader Datalagringsanskaffelsen		
	2016 i MNOK	2017 i MNOK
UNINETT Sigma2	2.2	0.3
Innleie Konsulent	0.7	
Innleie Universitet	0.4	0.1
Sum	3.4	0.4

Tabell 5. Prosjektkostnader datalagring

3.3.2 Investering HPC

Som tidligere nevnt er A1 beregnet til å måtte ha 50 000 kjerner og L 1000 kjerner. I tillegg kommer kostnaden for /work. Et estimat på investeringsbehovet er basert på ca. priser fra leverandører og er vist i tabell 6.

HPC maskin A1	61,3 mnok
HPC maskin L	1,3 mnok
/work – 1,5 PB	3,0 mnok
Sum investering HPC	65,6 mnok

Tabell 6. Investeringer HPC

Tre års vedlikehold er inkludert i disse estimatene.

3.3.3 Investering datalagringsutstyr

Fra arbeidsgruppens rapport er det anbefalt at utstyret har en initiell netto lagringskapasitet på 10 PiB (se Tabell 2). Kostnadene i Tabell 7 er basert på tilgjengelige markedspriser. Prisestimatene justeres nedad med 10% årlig fra disse estimatene. Videre er det tatt med kostnader for intern nettverk/rack infrastruktur, servere for data analyse, servere for visualisering og applikasjonsnoder, samt administrative programvare for storskala datalagring. Backup vil benytte seg av eksisterende løsning de første årene.

Volumet for ytelseslagring justeres oppad årlig med 20%, mens kapasitetslagringen skales oppad årlig med 40% (i henhold til forventningen om en dobling av datavolumet hvert 2.5 år).

Lagringsbehov	2017 (fase 1)	2018 (fase 2)	2019 (fase 2)
Ytelseslagring	3 PiB	3.5 PiB	4 PiB
Kapasitetslagring	7 PiB	10 PiB	14 PiB
Kostnader	2017	2018	2019
Ytelseslagring	10,7 mnok	2,0 mnok	2,1 mnok
Kapasitetslagring	11,0 mnok	4,0 mnok	5,1 mnok
Intern infrastruktur	3,0 mnok	1,5 mnok	1,5 mnok
Data analyse	2,0 mnok		1,0 mnok
Vis-App	1,0 mnok		0,5 mnok
<u>Programvare for forvaltning Adm- Software²</u>	5,0 mnok	1,5 mnok	2,0 mnok
Backup			5,0 mnok
Total	32,7 mnok	9,0 mnok	17,2 mnok

Tabell 7. Prosjeksjon og kostnadsestimater for anskaffelse (fase 1) og utvidelse (fase 2).

3.4 Finansiering

Finansieringen av innkjøpene skjer gjennom midler fra Universitetene (Samarbeidavtalen) og Forskningsrådet. I tillegg har det vært diskusjoner med Havforskningsinstituttet (Havforsk), SINTEF og Metrologisk institutt (met.no) om interesse for å bidra med finansiering i bytte med regnetid. Per i dag er det konklusjonen følgende:

- SINTEF har nettopp gjort investering i en mindre maskin, men kan være interessert i å kjøpe regnetid senere
- Met.no ønsker en egen maskin for å kjøre sin produksjon av værmeldinger. Vår konklusjon er at det veldig liten synergi i et samarbeid om dette.
- Havforskningsinstituttet sin interesse er ikke avklart, vi venter på tilbakemelding.

	MNOK
Prosjekt kostnader HPC	4.3
Prosjekt kostnader lagring	3.8
Investering HPC	65.6
Investering lagring	32.7
Sum Kostnader	106.4
Egne midler	31.4
Konkurranseutsatt	75
Sum investeringsmidler	106.4

Tabell 8. Investeringer HPC

3.5 Avhengighet til andre anskaffelser og prosjekter

NorStore skal plasseres på to lokasjoner, der A1 og L plasseres og der B1 plasseres. Det gjenstår å få avklart plassering for B1. B1 lokasjon bør derfor være klar til samme tidspunkt som A1 lokasjon, men en midlertidig plassering kan benyttes fordi den initielle installasjonen er forholdsvis liten.

² [Programvare for konfigurering, dataadministrasjon, sikring av data, optimalisering av ytelse, osv.](#)

3.6 Mottaksprosjekt

Som en del av disse anskaffelsene skal det i løpet av våren 2016 etableres et mottaksprosjekt som får ansvar for at organisasjonen er klar til motta det nye utstyret både når det gjelder nødvendig kompetanse, installasjon, testing, godkjenning og produksjonssetting. Personell som nomineres inn driftsteam for nytt anlegg er naturlige deltakere i et slikt mottaksprosjekt. Det må vurderes om det er nødvendig med separate prosjekt og team for HPC og lagring.

4. PLANLAGT GJENNOMFØRING AV ANSKAFFELSENE

4.1 Organisering av anskaffelsene (HPC og Datalagring)

Navn/Ressurser	Rolle / enhet	Behov/oppgaver
UNINETT Sigma2 v/Styret	Avtaleeier	Godkjenner anskaffelsesstrategien, samt verifiserer at anskaffelsene er gjennomført i henhold til anskaffelsesstrategien
Gunnar Bøe, Sigma2	Prosjekteier og ansvarlig innkjøper	Overordnede avklaringer rundt anskaffelsesprosjektet
Arild Halsetrønning, Sigma2	Prosjektleder	Ansvarlig for møter, prosesser og utførelse av anskaffelsene, i samsvar med anskaffelsesstrategien
Jørn Amundsen, Sigma2 Roy Dragseth, UiT (3) Ole Widar Saastad, UiO (2) Alexander Oltu, UiB (1) Jan Christ. Meyer, NTNU (1) Helge Stranden, UNINETT	Faglige ressurser HPC	Ansvarlig for utarbeidelse av kravspesifikasjon, herunder benchmarking, samt avklaring av tekniske krav/forutsetninger datahall
Andreas Jaunsen, Sigma2 Gurvinder Singh, UNINETT Hans Eide, UiO (2) NN, NTNU (2)	Faglige ressurser datalagring	Ansvarlig for utarbeidelse av kravspesifikasjon, samt sikre best mulig faglig samordning av krav mellom HPC og datalagring
Halvor Oseid, KPMG	Juridisk og anskaffelsesfaglig rådgivning/kvalitetssikring	Juridiske/anskaffelsesfaglige råd og avklaringer
Martin Rydland, KPMG	Prosjektstøtte	Generell bistand, møtereferater, tilrettelegging mv.

Tabell 9. Organisering av prosjektet

4.2 Tidsplan HPC

Aktivitet:	Tidsplan
Anskaffelsesstrategi ferdigstilt og godkjent	4. desember 2015
Konkurranse dokumenter godkjent	15. januar 2016
Kunngjøring Doffin/TED	18. januar 2016
Tilbudsfrist	18. mars 2016
Tilbudsevaluering	18. mars 2016 - 18. april 2016
Tildelingsbeslutning og innstilling godkjent	18. april 2016
Meddelelse tildelingsbeslutning	19. april 2016
Karensperiode	19. april 2016 - 29. april 2016
Inngåelse av kontrakt	02. mai 2016
Leveranse av L systemet	01. september 2016
Leveranse av A1 systemet	01. oktober 2016
Klar for produksjon	01. februar 2017

Tabell 10. Tidsplan HPC

Alle datoer etter tilbudsfrist er tentative.

4.3 Tidsplan Datalagring

Aktivitet:	Tidsplan
Kontraktsstrategi ferdigstilt og godkjent	4. desember 2015
Godkjent kvalifikasjonsdokumentasjon	29. januar 2016
Kunngjøring av konkurranse og oversendelse TED	29. januar 2016 (14. januar?)
Utsendelse av kvalifikasjonsgrunnlag	29. januar 2016 (14. januar?)
Frist for mottak av forespørsel om deltagelse	1. mars 2016 (15. februar?)
Evaluering og beslutning om prekvalifisering	1. mars 2016 - 11. mars 2016
Beslutning om prekvalifisering godkjent	11. mars 2016 (26. februar?)
Meddelelse om prekvalifisering til leverandør	16. mars 2016 (1. mars?)
Konkurranse dokumentasjon godkjent	18. mars 2016 (3. mars?)
Utsendelse av godkjent konkurranse dokumentasjon	18. mars 2016 (3. mars?)
Tilbudsfrist	2. mai 2016 (15. april?)
Forhandlinger og tilbudsevaluering	2. mai 2016 - 17. juni 2016
Tildelingsbeslutning og innstilling godkjent	17. juni 2016
Meddelelse tildelingsbeslutning	17. juni 2016
Karensperiode	17. juni 2016 - 27. juni 2016
Inngåelse av kontrakt	03. juli 2016
Leveranse av datalagringssystemet	02. november 2016
Klar for produksjon	01. februar 2017

Tabell 11. Tidsplan Datalagring

Alle datoer etter tilbudsfrist er tentative.

4.4 Fremtidige behov

Det er viktig at anskaffelsene sørger for opsjoner for fremtidig utbygging og dekning av nye behov. I 2018 planlegges en ny anskaffelse som vil omfatte B1 og ytterligere utvidelse av lagringssystemet samt ny backup-løsning.

5. MARKED OG KONKURRANSE

5.1 Oversikt over mulige leverandører

5.1.1 HPC

Det fins to hovedgrupper av HPC-systemer (tungregnesystemer), det er systemer med egenutviklet (proprietært) meldingsnettverk (Interconnect) og systemer med nettverk fra en underleverandør (OEM).

De to relevante arkitekturerne innen HPC for meldingsnettverk er Infiniband (IB) og Omni-Path Architecture (OPA). Mellanox er den viktigste leverandøren av IB, mens OPA er en Intel-proprietær arkitektur.

De mest relevante prosessorarkitekturer for HPC er OpenPOWER (IBM), x86 (Intel og AMD) og ARM (Samsung, Texas Instruments m.fl.). Denne oversikten dreier seg kun om systemer med x86-arkitektur.

Systemer med proprietære meldingsnettverk benevnes også kapabilitetssystemer, mens det motsatte benevnes kapasitetssystemer. Forskjellen ligger hovedsakelig meldingsnettverkets ytelse, og dets evne til å skalere opp for å kjøre beregninger med et stort antall prosessorkjerner effektivt, med minst mulig tapt tid til meldingsformidling. For et kapabilitetssystem vil båndbredden for meldingsutveksling øke proporsjonalt med antall kjerner, mens den vil flate ut etter et visst antall kjerner i et kapasitetssystem. Enkelte leverandører leverer både kapasitets- og kapabilitetssystemer, f.eks. Cray med sine XC- og CS-systemer. Kapabilitetssystemer er betydelig dyrere enn kapasitetssystemer, både i anskaffelse og drift.

HPC-systemer kan også leveres med væske- eller luftkjøling. Et væskekjølt system har en høyere anskaffelseskostnad (faktor ca. 1.03-1.3) og høyere vekt, men lavere driftskostnader gjennom bedre mulighet for varmegjenvinning, lavere strømforbruk, mer stabil chip-temperatur og mindre fotavtrykk (tar mindre plass på gulvet) i maskinrommet.

Antall mulige leverandører av større HPC systemer er begrenset, selv om det fins mange mindre selskaper som integrerer OEM-komponenter til diverse klyngeløsninger. I oversikten nedenfor er det derfor kun tatt med leverandører man har en viss forventning om å motta tilbud fra. En langt mer omfattende liste fins i TOP500s leverandørliste på lenken top500.org/statistics/list/.

- Cray Inc. (kapabilitet, kapasitet)
- HP Inc. (kapabilitet)
- Lenovo Group Ltd. (kapabilitet)
- SGI (kapabilitet)
- Megware Computer GmbH (kapasitet)
- Huawei/Go Virtual (kapasitet)
- Dell Inc. (kapasitet)
- Groupe Bull (kapasitet)

5.1.2 Lagringssystem

Som beskrevet i seksjon 2.3 er det viktig at man velger et lagringssystem som gir de beste forutsetninger for å etablere en skalerbar, fleksibel og brukervennlig nasjonal infrastruktur. Samspillet mellom maskinvare (HW) og programvare (SW) er svært viktig for å kunne realisere dette.

For lagringsressurser er det en rekke mulige leverandører. Vi har inndelt markedet i tre grupper;

- Nøkkelferdige systemer (en leverandør integrerer systemer fra flere underleverandører,
- HW leverandører
- SW leverandører

Den første gruppen leverer komplette løsninger som inkluderer nøkkelferdige løsninger hvor HW og SW er ett integrert produkt. De øvrige gruppene er leverandører som primært tilbyr enten HW eller SW. Disse HW/SW-spesialiserte leverandørene er derimot konkurransedyktige både på kost og funksjonalitet. Derfor er det viktig at også disse kan delta i konkurransen. Majoriteten av dem har tidligere vist evne til å delta i slike konkurranser ved å samarbeide, eksempelvis Scality+HP eller Quobyte+Stack Velocity.

Nøkkelferdige systemer	Hardware leverandører	Software leverandører
DataDirect Networks	AMAX	Caringo
Dell	Hyve Solutions	Quobyte
EMC	Nextron	Scality
Hewlett Packard	QCT (Quanta)	
Hitachi	Stack Velocity	
IBM		
Oracle		

Tabell 12. Oversikt over kjente/potensielle leverandører

5.2 Regelverket for offentlige anskaffelser

UNINETT Sigma2 er underlagt forskrift om offentlige anskaffelser, FOA. Hovedregelen er at anskaffelser skal foretas ved åpen eller begrenset anbudskonkurranse. Det følger av forskrift om offentlige anskaffelser (FOA) §§ 5-1 (del II) og 14-1 (del III). I vårt tilfelle vil del III komme til anvendelse, da anslått verdi av anskaffelsen overstiger terskelverdien på 1 million kroner eks. mva., se FOA § 2-2.

5.3 Hovedregel om anskaffelsesprosedyre

5.3.1 Anskaffelse av HPC

For HPC-anskaffelsen vil prosedyren åpen anbudskonkurranse være anvendelig fordi det er mulig å kravspesifisere nøyaktig det som skal anskaffes, samt at det er et begrenset antall aktuelle tilbydere i markedet.

5.3.2 Anskaffelse av datalagring

Kompleksiteten i lagringssystemet samt kombinasjoner av leverandører gjør anskaffelsesløpet for lagring mer utfordrende. Selv om markedet er delvis kjent for anskaffelsesgruppen er det på ingen måte en fullverdig oversikt over forskjellige tilbydere og deres produkter. For å sikre en best mulig anskaffelse er det ønskelig med dialog med de mest aktuelle datalagringsleverandørene for å kunne velge den beste løsningen. Dialog krever en annen anskaffelsesprosedyre, konkurranse med forhandling, i motsetning til HPC-anskaffelsen.

Konkurranse med forhandling er en unntaksprosedyre i anskaffelsesregelverket og krever en særskilt begrunnelse. Vilkåret er gitt i forskrift om offentlige anskaffelser (FOA) § 14-3. Vi antar at anskaffelsen av datalagring vil falle inn under unntaksreglene, fordi det ikke er mulig å utarbeide tilstrekkelig presise spesifikasjoner og derved ikke kunne lage en samlet prisfastsettelse (prisark) på forhånd (§ 14-3(1).b). Vi anser derfor anskaffelsesprosedyren "konkurranse med forhandlinger" som ønsket prosedyre for lagringsanskaffelsen. Imidlertid kan vi komme til en annen konklusjon ved videre utarbeidelse av kravspesifikasjonen.

5.4 Kontraktstype og betingelser

5.4.1 Generelt

Statens Standardavtaler (SSA) er utviklet for offentlig forvaltning og egner seg til ulike typer IT-anskaffelser og til kjøp av konsulenttjenester. Statens standardavtaler finnes i engelsk og norsk versjon, se www.anskaffelser.no

SSA er nylig revidert og var tilgjengelig for bruk fra 2.juli i år. Det har skjedd en del endringer i avtalene, blant annet var ambisjonen å gjøre avtalene mer balansert i plikter og rettigheter mellom kunde og leverandør. Den regnes (fortsett) som kundedvenlig, men leverandørene er vant med å prise inn risiko i forhold til kontraktsvilkårene.

Statens standardavtaler har også gode vedlegg som gjør det lettere å innarbeide de krav som oppdragsgiver har i konkurransen. Vi foreslår å benytte SSA som kontraktsgrunnlag i anskaffelsen. Spørsmålet blir deretter hvilken kontraktssmal som skal legges til grunn.

Utviklings- og tilpasningsavtalen (SSA-T) passer til kjøp av løsninger som skal tilpasses eller utvikles etter kundens behov. I motsetning til kjøpsavtalen (SSA-K) inneholder SSA-T en samarbeidsmodell mellom kunde og leverandør, faseinndeling av gjennomføringsløpet for en leveranse, samt opplegg for test og godkjenning. Avtalen åpner også for at leveransen kan deles opp i delleranser.

Statens standardavtale, kjøpsavtalen (SSA-K) brukes først og fremst til kjøp av standard utstyr og programvare, eventuelt en miks av hyllevareprodukter, som leverandøren har satt sammen for å dekke kundens behov.

5.4.2 HPC

Vi foreslår å legge SSA-K til grunn, fordi det er relativt standardisert utstyr leverandøren har satt sammen for å dekke kundens behov.

Den endelige beslutningen tas før utsendelsen av konkurransegrunnlaget.

5.4.3 Lagringssystem

Vi foreslår å legge SSA-T til grunn, fordi vi antakelig trenger en avtale med en klar og tydelig prosjektplan og et godt beskrevet samarbeid mellom kunde og leverandør, for å få til en vellykket anskaffelse.

Den endelige beslutningen tas før utsendelsen av konkurransegrunnlaget

5.5 Samfunnsansvar

5.5.1 Miljø

Innvirkning på det ytre miljø skal minimaliseres og vil i størst mulig følge anbefalingene i Energy Star for servere, www.energystar.gov.

HPC-maskin A1 og L, samt datalagringsutstyret skal installeres i datahall hos Universitetet i Tromsø hvor det ligger godt til rette for varmegjenvinning.

5.5.2 Etisk handel (varekjøp)

FNs barnekonvensjon artikkel 32 og ILO konvensjon § 138 er tatt inn i standardkontraktene.

5.5.3 Sosialt ansvar (tjenestekjøp)

Hensynet til arbeidstageres lønn og arbeidsvilkår vil bli ivaretatt i kontrakten, i henhold til anskaffelsesloven § 11a jf. forskriften om lønn og arbeidsvilkår av 8. februar 2008 nr. 112 om lønns- og arbeidsvilkår i offentlige kontrakter.

6. RISIKOANALYSE OG KRITISKE SUKSESSFAKTORER

Med risiko menes her hendelser som kan inntreffe og som har en virkning på resultatmål.

Risikofaktor	Sannsynlighet	Konsekvens	Risiko	Tiltak
HPC anskaffelse blir forsinket	1	4	4	Gjennomarbeidet prosjektplan, god prosjektledelse, bistand fra profesjonelle innkjøpere (KPMG)
Mangelfull behovsdekning bruksmessig	1	4	4	Spesifiser og gjennomføre forbedringstiltak i samarbeid med brukere
Investeringsmidler er ikke tilstrekkelige ift. behovet	2	4	8	Hensyn ta dette i ny søknadsrunde i 2016 RFK må avvise en større andel søknader
Anskaffelse av datalagring blir forsinket	2	4	8	Sikre opsjon om kjøp av lagring som del av HPC anskaffelsen

Sannsynlighet for at risikofaktor skjer (1-4)

Konsekvens av at risikofaktor skjer (1-4)

6.1 Kritiske suksessfaktorer

Følgende er våre kritiske suksessfaktorer:

- Nytt HPC system, A1, en mindre maskin, L og nytt datalagringssystem i produksjon senest 1. mars 2017
- Brukertilfredshet med Notur og Norstore målt gjennom den årlige brukerundersøkelsen er minst like bra etter, som før utskifting av utstyret
- 20 % reduserte driftskostnader
- Sikre forutsigbare og lavest mulige vedlikeholdskostnader i systemenes levetid