

# SUNET distribuerad lagring

## Förstudierapport

### Inledning/Bakgrund

Under 2016 genomförde Sunet tillsammans med ett antal av de större nationella och internationella infrastrukturerna (SND, SNIC, SciLifeLab, Onsala Rymdobservatorium, EISCAT\_3D och MaxIV) en dialog kring behovet av samordning av långsiktig lagring av forskningsdata. Behoven av samordning involverar både metadata, identifierare, protokoll, APIer samt även tekniska lösningar för lagring av digitala objekt på permanent lagringsmedium (lågnivålagring).

Under 2017 genomförde Sunet en förstudie tillsammans med bla Göteborgs Universitet, Uppsala Universitet, SND och SciLife Lab (bla BILS) i syfte att utforska möjligheterna till en distribuerad lagringslösning där befintliga lokaler och resurser på lärosäten används för att sänka kostnaden.

Som resultat av denna förstudie ämnar Sunet etablera ett utvecklingsprojekt i syfte att etablera en *distribuerad* tjänst för objektlagring. Denna kan fungera som lagringsinfrastruktur vid behov av att kunna lagra digitala objekt på ett så billigt och enkelt sätt som möjligt. Detta dokument är underlag till detta utvecklingsprojekt i samverkan med berörda parter.

Förslaget bygger på en **distribuerad** tjänst i meningen att lagringsplattformen (servrar, lagringsmedium, mm) framförallt kommer att placeras hos de lärosäten som har störst behov av lagring. En sådan tjänst kommer i någon mening vara **federerad** genom att vissa lärosäten kommer att kunna bidra med resurser för inplaceringen (datorhall, el, kyla osv) till tjänsten, i utbyte mot lägre kostnad för lagring av större volymer. Denna tjänst, "Sunet distribuerad lagring", kommer att erbjudas enligt en fast prislista till alla Sunet-anslutna organisationer på samma sätt som alla andra tilläggstjänster, t.ex. Box, Zoom etc.

Tjänsten "Sunet distribuerad lagring" kommer att etableras som en vidareutveckling av Sunets befintliga tjänster i inom gruppen "Sunets molntjänster". Den viktiga och stora skillnaden är att "Sunet distribuerad lagring" kommer att ägas helt och hållet av Sunet till skillnad från Sunets befintliga molntjänster där ägandestrukturen ser annorlunda ut. Dock kommer "Sunet distribuerad lagring" att integreras tekniskt så att det blir möjligt att enkelt flytta tjänster och data mellan Sunets befintliga molntjänster.

Utgångspunkten för "Sunet distribuerad lagring" är ett långsiktigt åtagande för Sunet så länge det finns intresse och behov hos Sunets kunder.

### Sammanfattning

Förslaget är att etablera en federerad/distribuerad lagringstjänst med följande egenskaper:

- Plattformen består av ett antal *lagringsnoder* som presenteras via en gemensam tjänsteportal.

- Användare får tillgång till tjänsteportalen via SWAMID-inloggning, som också kan styra fakturering och uppföljning.
- All hårdvara och mjukvara som används i tjänsten ägs av Sunet.
- Lagringsnoder placeras på de lärosäten som beställer minst 2 PB (PetaByte) lagring.
- För att uppnå lägsta möjliga kostnad vid stora volymer förutsätts att lärosätet har möjlighet att erbjuda el, kyla och plats i kvalitetssäkrad datorhall.
- Kapacitet kan beställas i ett antal "nivåer" kopplade till olika pris.
- Lärosäten som inte vill ha en egen lagringsnod kan välja att köpa kapacitet från en av de befintliga lagringsnoderna.
- Lagringsnoderna ansluts direkt till Sunets nät, SunetC, med minst 100 Gbps och belastar inte lärosätets nätverk eller uppkopplingen till Sunet.
- Lagringsnoder kan vid behov anslutas direkt till lokala beräkningsresurser.
- Tekniken bygger på CEPH och OpenStack. Lagrings-APIet blir i första hand Amazon S3.
- Applikationer/tjänster som SND utvecklar kommer att anslutas direkt till lagringsnoderna för maximal prestanda.
- Utgångspunkten för "Sunet distribuerad lagring" är ett långsiktigt åtagande för Sunet så länge det finns intresse och behov hos Sunets kunder.

### Övergripande Mål och PKIer

Följande mål, med tillhörande KPIer, har använts för att designa och utvärdera lösningar för tjänsten:

- **Kostnad.** Det övergripande målet med lagringstjänsten är att på ett så kostnadseffektivt sätt som möjligt uppfylla lagringskraven från ett så stort antal tjänster som möjligt som har behov av att kunna lagra och hämta digitala objekt. Genom att bygga en generell lagringsinfrastruktur kan kostnaden amorteras över ett stort antal tjänster vilket i sin tur minskar kostnaden per lagrad GByte. Den huvudsakliga KPI:n för plattformen är kostnad per GByte per månad.
- **Närhet.** I volymer över 0.1 PB data blir nätverkskapaciteten en dimensionerande faktor. Vid en hastighet om 10 Gbit/s tar det ca 20 timmar att överföra 100 TB. Om infrastrukturen bara ska användas för arkivering räcker detta antagligen, men eftersom arkivering av data endast utgör en liten del av lagringsbehovet för forskardata är det iså fall svårt att uppfylla målet (1) om kostnadseffektivitet. För att infrastrukturen ska vara användbar för mer än arkivering måste data alltså vara tillgängligt med högsta möjliga överföringshastighet mätt i Gbit/s. Den praktiska gränsen ligger idag kring 50-70 Gbps.
- **Enkelhet.** En lagringstjänst som riktar sig till forskare måste vara enkel att använda och måste gå att integrera med existerande processer och applikationer. En möjlig KPI för detta är % forskare som använder tjänsten som del av de som har möjlighet att använda tjänsten.

## Förslag på Tjänstebeskrivning

Lagringstjänsten består av följande delar:

- En portal där behöriga användare kan skapa och administrera "sina" lagringsresurser. Användare loggar in i portalen med SWAMID-inloggning.
- Ett antal lagringsnoder placerade på de lärosäten som beställer (antingen direkt eller aggregerat) 2 PB lagring eller mer.
- Ett API (AWS S3) som ger tillgång till enskilda logiska lagringsytor i form av digitala objekt och samlingar av digitala objekt. Detta API används av tjänster och applikationer för att lagra och hämta data som lagrats i tjänsten.
- Ett gränssnitt för direkt anslutning av servrar/virtuella maskiner till en lagringsnod – som driftas antingen i Sunet eller i lärosätets regi.
- En fakturerings/redovisningsprocess som medger projektbaserad uppföljning och fakturering av de resurser som används i systemet.

## Införandeprojekt

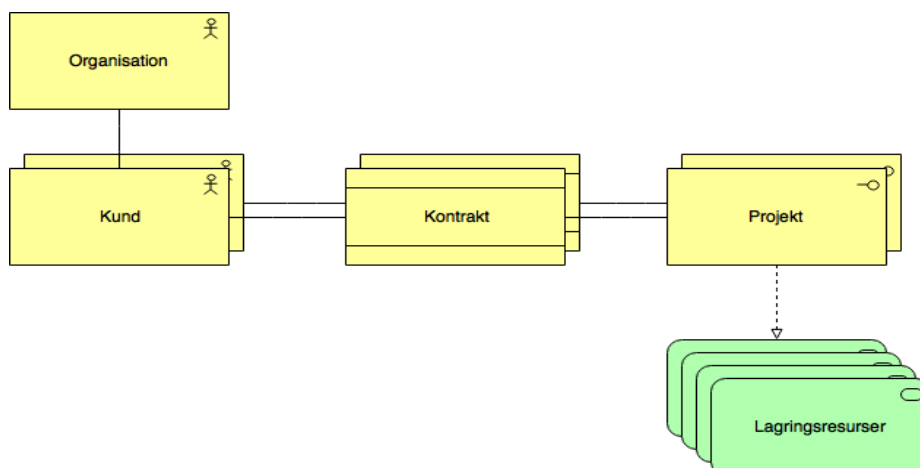
SUNET kommer att etablera ett utvecklings- och införandeprojekt för tjänsten "Sunet distribuerad lagring". För detta projekt kommer SUNET att etablera en styrgrupp som i huvudsak består av de lärosäten som under förstudiefasen indikerat att man är intresserad av tillräckligt stora lagringsvolymerna för att motivera etablering av en lagringsnod. SUNET kommer att tillsätta en projektledare och även etablera en referensgrupp i huvudsak bestående av personer som representerar framtida användare av tjänsten eller som besitter värdefull teknisk kunskap.

Projektet kommer även att få i uppdrag att hitta ett bra och beskrivande namn på tjänsten samt att slutligen etablera de tjänste- och prisnivåer som skissats nedan.

## Affärsrelationer

Lagringstjänsten kommer att drivas på samma sätt som SunetC: alla materiella och immateriella resurser ägs av SUNET och drift kan ske genom ett antal kontrakterade driftspartners under ledning av SUNET-personal.

Denna ansats garanterar långsiktighet i tjänsten eftersom SUNET inte behöver upphandla själva tjänsten, utan endast möjligtvis driften av befintliga resurser och hårdvara. Detta innebär dock att SUNET måste stå för alla investeringar vilket medför att SUNET kommer att etablera tjänsten i takt med att det finns konkreta kunder i tjänsten. Det finns också vissa miniminivåer som måste uppnås innan det är kostnadseffektivt att bygga lagring på detta sätt.



Sunet kommer att etablera tjänsten med hjälp av existerande kontrakt gentemot de flesta underleverantörer som täcker både hårdvara, mjukvaruutveckling och drift, men kommer på sikt att behöva förnya dessa upphandlingar i den takt som krävs för att hålla tjänsten vid liv. Detta arbetssätt är exakt samma som används för nätet och har visat sig fungera väl.

En viktig skillnad mot hur SunetC drivs är dock att resurser i lagringstjänsten måste kunna redovisas och faktureras på projektnivå. Ett lärosäte väljer att skriva ett eller flera kontrakt för att få tillgång till tjänsten – t.ex. kan olika större forskningsprojekt ha egna projekt medan lärosätets IT-avdelning “tar hand om” mindre projekt och enskilda forskare.

Varje lagringsresurs i tjänsten kopplas till ett *projekt* som knyts till en projektkod eller annan fakturareferens i lärosätets ekonomisystem. Resurser faktureras per kontrakt och portalen erbjuder möjlighet att ladda ner en specifikation av resurser som konsumerats, med tillhörande projektkoder för varje faktura. De lärosäten som så önskar kan koppla projektkoder via SWAMID-inloggning så att lärosätets processer för hantering av behörigheter styr både vem som får rätt att skapa lagringsytor samt hur kostnaderna för dessa bokförs.

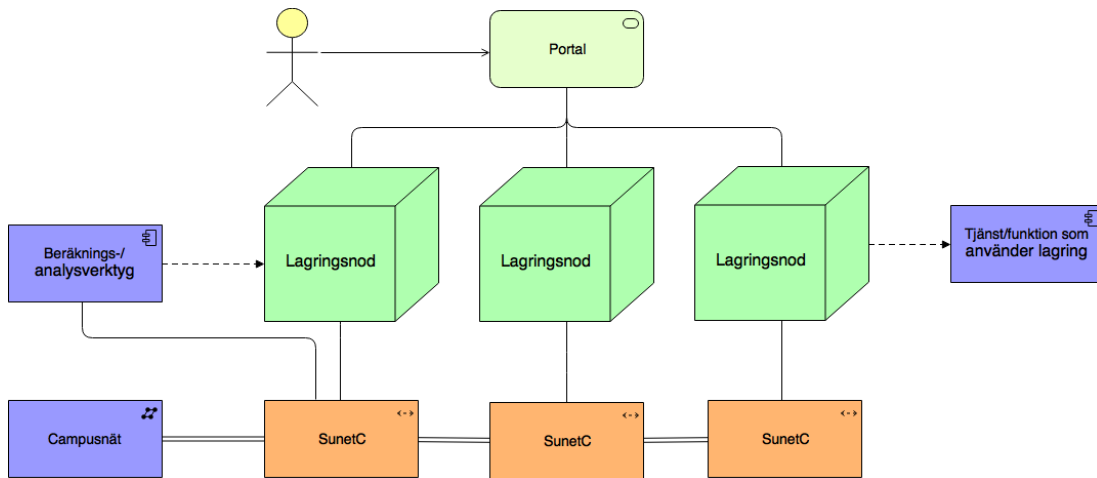
Genom att välja att teckna ett eller flera kontrakt kan lärosätet välja hur finkornig delegationen av ansvar till användargrupper ska vara - vissa lärosäten kan välja att teckna ett kontrakt och hantera den interna kostnadsfördelningen på egen hand medan andra lärosäten kan välja att teckna ett kontrakt per användargrupp.

### Tjänste-/prisnivåer

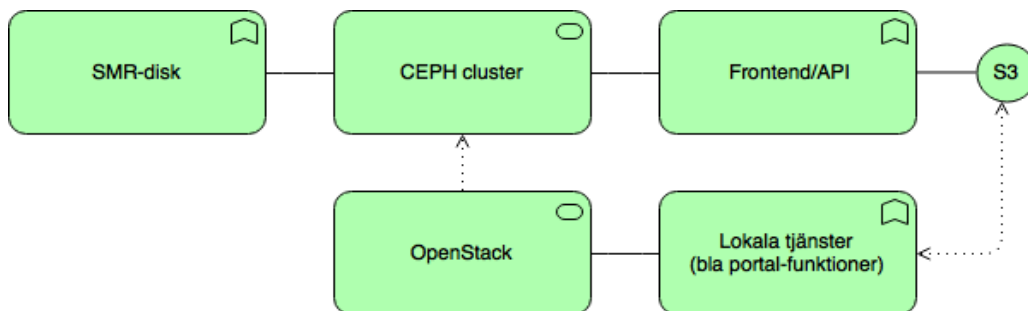
Sunet planerar att etablera 3 tjänstenivåer: liten, mellan och stor. Samtliga dessa nivåer levereras på samma sätt - via lagrings-API:et men skiljer sig åt i minsta storlek på beställningen, priser samt var lagringsresursen etableras. Dessa är dock endast kontraktsmässiga skillnader - åtkomsten av lagringsresurserna sker på samma sätt oavsett hur mycket som beställs och var dessa resurser etableras.

Nivå	Beskrivning	Pris(kr/GB /mån)
<b>liten</b>	lagring levereras i mån av utrymme från någon existerande lagringsnod optimerat för lägsta möjliga produktionskostnad	TBD
<b>mellan</b>	lagring i inkrement om 160 TB levereras från närliggande befintliga lagringsnod	TBD
<b>stor</b>	2 PB och uppåt levereras från lokal (eller närmaste om det inte finns möjlighet till lokal placering) lagringsnod	TBD

Teknisk arkitektur



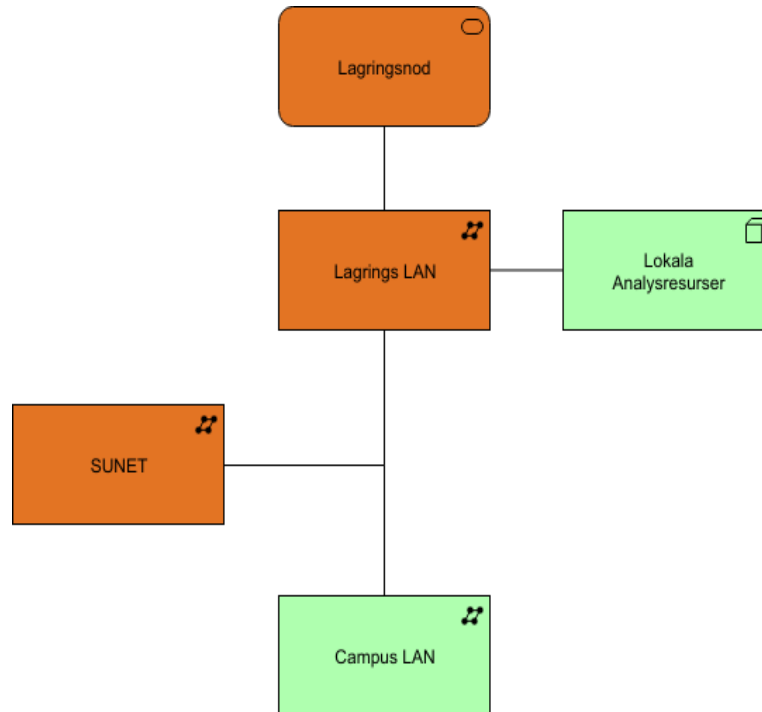
En lagringsnod består av ett CEPH-kluster frontat av ett antal API-servrar som implementerar AWS S3-gränssnittet. En gemensam tjänsteportal presenterar lagringsnoderna som *regioner* i analogi med hur AWS (Amazon Web Services) och andra publika molntjänster låter användaren välja var data lagras. Vid behov kan virtuella maskiner levereras i nära anslutning till en lagringsnod tex tjänster som behöver snabb access till lagringen.



Data lagras erasure-kodat på sk SMR-diskar i CEPH-klustret och data kan vid behov replikeras mellan olika regioner för att uppnå geografisk redundans för kritiskt data.

Lagringsnoden ansluts (via Frontend/API-servrarna) både campus-nätet och direkt till SunetC antingen med multipla 10G eller 100G interface. Anslutningshastigheten kommer att anpassas och växa med behovet

Dessutom kommer det erbjudas möjlighet att ansluta lokala beräknings- och analysresurser direkt till lagrings-nätet vilket gör det möjligt att använda lagringen för tjänster där det är viktigt att ha låg fördröjning till lagringsinfrastrukturen.



### Koppling till Svensk Nationell Datatjänst (SND)

Svensk Nationell Datatjänst (SND) koordinerar ett nätverk av 27 svenska lärosäten med syfte att förbättra tillgängligheten av svenska forskningsdata enligt FAIR-principerna. Detta inkluderar moduler för att koordinera ett nationellt repository av forskningsdata samt en nationell sökportal och metaregister.

Genom att integrera tjänsten med SND:s system för metadatabeskrivningar av dataset försäkras en stark koppling mellan data och tillhörande metadata av hög kvalitet och beskrivna med relevanta standarder. En centralisering hantering av metadata säkerställer att dessa är harmoniserade och interoperabiliteten mellan olika system ökar. Detta innebär även att dataset i förlängningen bli sökbara i SND:s nationella sökportal, samt i internationella portaler kopplade till SND, såsom t.ex. CESSDA Portal, ARIADNE och DataCite.

SND är den svenska service providern för DataCite. Genom att integrera tjänsten med SND:s tjänster för persistenta identifierare kan Data Object Identifiers (DOI) genereras för enkel citering och tillgängliggörande av dataset.

Tjänsten kan även integreras med SND:s tillgängliggörandetjänster. Detta förenklar hanteringen av utlämningsärenden från universitetet.