

Referat

Ekstraordinært møde i den tekniske arbejdsgruppe den 29.06.2026

Dato: 29. juni 2026 Mødeleder: Bettina Lundgren
Sted: Virtuelt Microsoft Teams Sekretær: Rikke Korshøj Andersen

Punkt	Ca. tid	Aktivitet
-------	---------	-----------

- | | | |
|----|-------------|--|
| 1. | 14.30-15.15 | 1. Ny storageplatform, herunder: <ul style="list-style-type: none">1.1 Kvalificere kapacitetsbehov<ul style="list-style-type: none">• Nuværende dimensionering• Risikobilledet for den eksisterende platform og behovet for ny storage• Fremtidig dimensionering1.2 Kvalificere udbudsmateriale |
| | | 2. Evt. |

(B)– beslutning; (D) – drøftelse; (O) – orientering.

Medlemmer af den tekniske arbejdsgruppe - deltagere

Mads Sønderkær, senior Bioinformatiker, Aalborg Universitetshospital, Region Nordjylland (afbud)
Jacob Lind-Galsgaard, IT-arkitekt, Digitalisering og IT, Region Nordjylland (afbud)
Søren Vang, ledende bioinformatiker, Aarhus Universitetshospital, Region Midtjylland
Henrik Hammer Jordt, områdeleder, Digitalisering og It, Region Midtjylland
Mads Aagaard Jørgensen, bioinformatiker, Ph.D., Sygehus Lillebælt, Region Syddanmark
Martin Larsen, bioinformatik/klinisk laboratoriegenetiker, Ph.D., OUH, Region Syddanmark
Thomas Lau Lindestrand-Hansen, Cand. Scient., Ph.d., SUH-Køge, Region Sjælland
Michael Jørgensen, IT-arkitektur, Koncern Digitalisering, Ringsted, Region Sjælland
Frederik Otzen Bagger, ledende bioinformatiker, Rigshospitalet, Region Hovedstaden
Søren Zachariassen, sektionschef, IT-arkitektur, CIMT, Region Hovedstaden
Bettina Lundgren, vicedirektør, SDS
Nicolas Rapin, chefkonsulent, NGC, SDS
Mikael Kronborg Christophersen, chefkonsulent, NGC, SDS
Astrid Munk Pedersen, specialkonsulent, NGC, SDS (afbud)
Jacob Drasbeck, sektionsleder, HPC & Cloudplatform, SDS
Mette Bruun Kolbye, sektionsleder, NGC, SDS
Emma Mølgaard Engelbreth, NGC, SDS
Rikke Korshøj Andersen, specialkonsulent, NGC, SDS (sekretær for arbejdsgruppen)
Gæster: Ole Halfdan Larsen, Region Midtjylland, repræsentant fra Fagligt råd for genomisk medicin (forperson)
Anders Thomsen Lunde, specialkonsulent, Region Syddanmark
Anne Kirstine Winding, chefkonsulent, Region Hovedstaden (afbud)
Cecilie Bisgaard-Nøhr, chefkonsulent, Danske Regioner (afbud)

Pkt. 1 Ny storageplatform v/Bettina Lundgren

Ny storageplatform, herunder:

1.1 Kvalificering af kapacitetsbehov

- Nuværende dimensionering
- Risikobilledet for den eksisterende platform og behovet for ny storage
- Fremtidig dimensionering

Referat:

Bettina bød velkommen til mødet og orienterede om, at vi har inviteret gæster med hhv. forperson fra fagligt råd for genomisk medicin Ole Halfdan samt medlemmer af fællessekretariatet for den nationale bestyrelse for drift af genomdatabasen, hvorfra Anders Thomsen Lunde deltog. Tak for deltagelse med kort varsel.

Bettina gav herefter ordet til Nicolas Rapin, som præsenterede den nuværende dimensionering med udspecificering af analyse og dataplatform.

Nuværende dimensionering

Præsentationen illustrerede bl.a., at data er levende, så der kan være forskelle i data, alt efter hvilken dag de opgøres. Endvidere, at 1 MB primærdata i praksis fylder 2–3 MB grundet backup, snapshots og kopier.

Dvs., at en fuldt bearbejdet patient: $\sim 489 \text{ GB} \times 60.000 \approx$ op til $\sim 29 \text{ PB}$ – mod kun 1,3 PB i en simpel regning.

Nicolas viste også slides med kategoriseringer af forbruget (se vedlagte slides).

Medlemmer af den tekniske arbejdsgruppe havde følgende bemærkninger til dimensionering:

- Mads Jørgensen: Af opgørelsen på de "op til 489 GB" pr genom fremgår det, at der bl.a. er 3 kopier af rådata i arkiv, og yderligere 2 kopier i hhv. NAS og backup – er det nødvendigt med 5 kopier af rådata, f.eks. FastQ?
Og, er det nødvendigt med 4 kopier af CRAM, som beregningen også inkluderer?
Nicolas svarede, at én kopi med back-up er tilstrækkeligt.
Jacob Drasbeck supplerede med en bemærkning om, at vi med automatisk tiering fremadrettet vil have én version af hver fil + en sikkerhedskopi. Med den nye platform er der behov for færre kopier.
- Mads Jørgensen spurgte endvidere, om der er udarbejdet en beregning for hvad storage-behovet er for et fuldt bearbejdet genom på den nye storage arkitektur?

Hertil svarede Jacob, at der ikke er lavet en sådan beregning, så vi ved ikke hvad storage-behovet pr. genom er på ny platform, men det forventes at blive mindre, end på nuværende storage platform. Nicolas orienterede om, at Longreads fylder mere, hvilket der vil blive behov for plads til fremover. Bettina supplerede med bemærkning om, at når den nye storage-løsning er valgt, vil NGC inddrage den tekniske arbejdsgruppe/fagligt råd i planlægning af målbillede for arkitektur og proces for hvilke data, der fremover skal hhv. gemmes og slettes.

Øvrige bemærkninger fra medlemmerne i den tekniske arbejdsgruppe til dimensionering:

- Mads Jørgensen: Positivt, at den nye platform giver mulighed for effektivisering.
- Søren Zachariassen: Positivt med softwaredefined platform, det giver bedre mulighed for at monitorere forbrug fremover.

Risikobilledet

Jacob Drasbeck præsenterede risikobilledet og orienterede om, at det er vigtigt at udskifte storageplatformen, da antallet af hardwarefejl er stigende og vi ser stigende omkostninger.

Med ny storageplatform får vi mulighed for at få en hardwareagnostisk platform med henblik på at reducere bindinger til bestemte leverandører.

Jacob orienterede om, at storage vokser med ca. 27% om året. NGC komprimerer og arbejder løbende med at frigøre plads. Trods dette vokser storage fortsat med 27% om året.

Throughput dimensioneres efter det faktiske peak-forbrug vi ser aktuelt.

Storage-forbrug i sammenlignelige lande

Bettina orienterede om, at forbruget svarer til det vi ser i sammenlignelige lande internationalt. I fx Genomics England ses tilsvarende 480 GB for ét genom.

Fremtidig dimensionering – bemærkninger fra den tekniske arbejdsgruppe:

- Frederik Bagger bemærkede, at indkøb af softwaredefined storage er fornuftigt, og var enig i forbruget i sammenlignelige lande, nævnte bl.a. også, at Japan har tilsvarende forbrug.
- Søren Zachariassen bemærkede, at det er vigtigt, at vi udskifter storage nu og ikke løber unødige risici.
- Søren Vang tilsluttede sig bemærkning fra Søren Zachariassen og kvitterede for gennemgangen af dimensioneringen, og nævnte, at han var overrasket over, hvor meget det hele egentlig fylder. Positivt at få indblik i kapacitet, det var der behov for.

1.2 Kvalificering af udbudsmateriale

Referat:

Jacob orienterede om, hvorfor vi går i udbud på 02.06. Det skyldes bl.a., at der er ringe konkurrence på 50.03. Samlet set ses mange flere fordele ved 02.06 end ved 50.03, herunder et bedre aftalegrundlag.

Ole Halfdan spurgte, om undgåelse af vendor-lock-in er beskrevet tydeligt nok i udbudsmaterialet. Bettina svarede, at dette er beskrevet i udbudsmaterialet – NGC tjekker efter, om dette fremgår tilstrækkeligt eksplicit.

Bemærkninger fra mødedeltagerne til udbudsmateriale:

- Ole Halfdan kvitterede for et godt, brugbart materiale med en tilfredsstillende detaljegrad, og nævnte, at der fremover bør være fokus på større inddragelse af hhv. den tekniske arbejdsgruppe og af fagligt råd i processer som denne. NGC tog bemærkningen til efterretning.
- Det blev nævnt, at bestyrelsen bør forelægges en plan for håndtering af tilbud, der ligger væsentligt over budget.

Procesplan

Bettina gennemgik herefter procesplan for storageudbuddet.

Såfremt bestyrelsen tiltræder, at vi går i udbud på bestyrelsens møde den 2. juli, afholdes endnu et ekstraordinært møde i den tekniske arbejdsgruppe omkring den 21.08.2026. Rikke lægger en placeholder i kalenderen.

Afsluttende bemærkninger fra mødedeltagerne

- Søren Zachariassen bemærkede, at detaljer omkring optimering mv. ikke må blive en forhindring i anskaffelsen. Det er vigtigt, at bestyrelsen er OBS på dette.
- Frederik Bagger supplerede med en anbefaling om, at storage kommer i udbud nu. Vigtigt, at vi kommer i gang, vi er allerede forsinket.
- Der var generelt enighed om, at udbuddet ikke kan forsinkes yderligere.
- Ole bad om, at der oprettes en mail-tråd, hvor mødedeltagernes samlede bemærkninger rundsendes til gruppen her. Rikke Korshøj opretter mail-tråden umiddelbart efter mødet.

NGC vil efterfølgende udarbejde materiale til bestyrelsens møde den 2. juli, hvor de fremsendte bemærkninger kan indarbejdes.

Bettina takkede for et godt møde og for deltagelse med kort varsel i det ekstraordinære møde.

NB: Sagsfremstilling tilhørende mødet findes nedenfor.

29. juni 2026
Afdeling FDI
Sektion NGC
Initialer EMENG
Sags.nr.

Sagsfremstilling – Storage udbud, storage-strategi og kapacitetsbehov

Indstilling

Det indstilles, at den tekniske arbejdsgruppe:

- Kvalificerer kapacitetsbehov og udkast til udbudsmateriale til indkøb af ny storage-platform forud for, at materialet fremlægges til godkendelse i National bestyrelse for drift af genomdatabasen.

Resumé

National bestyrelse for drift af genomdatabasen besluttede på et ordinært møde d. 18. juni 2026, at kapacitetsbehov, storage-strategi og udkast til udbudsmateriale skal kvalificeres i den tekniske arbejdsgruppe, forud for forelæggelse ved bestyrelsen.

På mødet i teknisk arbejdsgruppe d. 17. februar bakkede arbejdsgruppen op om planerne for indkøb af ny storage-platform, jf. pkt. 6 i referat. Desuden blev arbejdsgruppen på mødet d. 20. maj orienteret om proces for storage-udbud, jf. pkt. 6 i referat. Det tidligere udbud blev annulleret, da det ikke var fuldt konditions-mæssigt.

Storage er en central og kritisk del af den national infrastruktur for personlig medicin, men per juni 2026 er ca. 63% af den installerede rå kapacitet end-of-support, 31% er end-of-life, og kun 9% er fortsat dækket af aktiv producentunderstøttelse. Den eksisterende storageinfrastruktur er fortsat i drift, men antallet af hardwarefejl er stigende, hvilket øger det samlede risikobillede.

Yderligere forsinkelse af udbuddet vil øge risikoen for driftsforstyrrelser i den nationale infrastruktur, medføre stigende support- og vedligeholdelsesomkostninger, samt vanskeliggøre

overholdelse af regulatoriske krav, herunder krav om geografisk adskilt backup og cyberrobusthed.

Indkøb af ny storageplatform er en bunden og tidskritisk opgave under det strategiske arbejdsprogram 2026-2028. Den tekniske arbejdsgruppe skal derfor kvalificere kapacitetsbehovet og udbudsmaterialet, så et nyt udbud kan igangsættes uden yderligere forsinkelse.

Baggrund

Regulatorisk kontekst:

Kravene til cyber-, informationssikkerhed og databeskyttelse er ved lov videreført fra Nationalt Genom Center til Sundhedsdatastyrelsen og Digital Sundhed Danmark. De regulatoriske krav skærpes løbende gennem blandt andet GDPR, NIS2, CER og de tekniske minimumskrav for statslige myndigheder. Specifikt er der krav om certificering efter ISO/IEC 27001 og ISO/IEC 27701, samt ISAE 3000 som databehandler og underdatabehandler.

Fra starten af har der politisk været fastsat en meget lav risikovillighed. Desuden kommer der fra ministerieret og Styrelsen for Samfundssikkerhed (SAMSIK) løbende skærpede krav til sikkerhed og robusthed. De skærpede krav omfatter blandt andet kravet om en geografisk adskilt offsite backup og styrket cyberrobusthed.

Forudsætninger for kapacitetsbehovet:

Det er ikke holdbart at videreføre den eksisterende platform både teknisk og økonomisk. Storage kapaciteten understøtter primært nationale myndighedsopgaver ifbm. infrastrukturen for personlig medicin, hvor en væsentlig del vedrører arkivering og langsigtet opbevaring. Dette omhandler i høj grad lovbundne krav til indberetning, opbevaring, databeskyttelse, tilgængeliggørelse og langsigtet forvaltning af genomiske data (jf. §7)¹. Disse behov eksisterer uanset om de enkelte regioner vælger at anvende den fælles infrastruktur.

Det er desuden afklaret på mødet d. 20. maj i teknisk arbejdsgruppe, at der fremadrettet indberettes FASTQ filer eller ækvivalente CRAM-filer, jf. pkt. 4. i referat.

Status – nyt storageudbud:

Det forrige udbud på SKI 50.03 modtog kun et tilbud og tilbuddet lå væsentligt højere end ordregivers budget. Erfaringerne fra udbuddet viser, at markedspriserne ligger over den forventede budgetramme. Det nye udbud gennemføres derfor på SKI 02.06, hvor et øget antal af leverandører forventes at bidrage til øget konkurrence og et mere fordelagtigt økonomisk resultat. De indholdsmæssige krav til udbuddet er fortsat de samme. Beslutning vedrørende skift af SKI aftale er truffet i SDS styregruppe vedr. indkøb af ny storageplatform efter anbefaling fra SKI og efter drøftelse med Kammeradvokaten.

¹ Lov om Digital Sundhed Danmark

Den aktuelle markedssituation for storage er præget af:

- Stigende priser, som kun forventes fortsat at stige.
- Stigende supportpriser.
- Reduceret adgang til reservedele og specialiserede kompetencer.

På de nuværende storageplatforme er der manglende producentunderstøttelse, som betyder, at reservedele bliver stadig vanskeligere og dyrere at skaffe. Platformene er i flere tilfælde låst til producentgodkendte komponenter, og omkostningerne til support og fejlretning er steget så markant, at det ofte er billigere at anskaffe nyt hardware, der er dimensioneret til de faktiske behov, end at forlænge levetiden på den eksisterende løsning. Det er essentielt at etablere en ny og fremtidssikret storage-platform, da yderligere udskydelse er med til at øge anskaffelsesomkostningerne og forstærke den eksisterende driftsrisiko.

Dimensionering af den nuværende platform:

Den nationale genomdatabase er en aktiv data- og analyseinfrastruktur, og ikke et passivt arkiv for FASTQ- eller CRAM-filer. Genomdatabasen fungerer som en national data- og analyseinfrastruktur, hvor primære genomdata lagres, bearbejdes og omsættes til klinisk og forskningsmæssig værdi. Storageforbruget omfatter derfor ikke kun primær data, men også pipeline-output, analysemiljøer, afledte dataprodukter, midlertidige arbejdsdata og databeskyttelse herunder snapshots og sikkerhedskopier.

Den nuværende infrastruktur for personlig medicin er på 82,6 PB, hvoraf der er 70 PB tilbage efter konfiguration. Af de 70 PB er ca. 58 PB brugbar kapacitet. Heraf er kun ca. 27 PB primærdata, mens de øvrige 31 PB er til databeskyttelse i form af backup, snapshots og kopier. Det er derfor vigtigt at forstå, at 1 MB data i praksis fylder 2-3 MB. Forbruget fordeler sig på fire hovedområder:

<i>Kapacitet:</i>	<i>Område:</i>	<i>Beskrivelse:</i>
4 PB	Drift & Platform	Til management, drift og platformservices, inkl. snapshots og backup.
2 PB	Genomisk dataudveksling	Til langtidsopbevaring af primære genomdata.
12 PB	Genomisk arkiv	Til langtidsopbevaring af primære genomdata (FASTQ/CRAM/BAM m.fl.), samt tilhørende snapshots og sikkerhedskopier.
40 PB	Genomisk analyse- og dataplatform	Som produktionsmiljø med primære data, pipeline-output, Variant- og Frekvensdatabasen, analyseresultater og arbejdsdata, samt tilhørende backup.

58 PB i alt		
-------------	--	--

Genomisk analyse- og dataplatform dækker dels ca. 7,5 PB primærdata forbrugt ved Region Øst (samt 8,3 PB til backup). og dels 11,5 PB primærdata (produktions- og forskningsdata) og 12,7 PB backup.

Produktions- og forskningsdata dækker over 4 områder. 1) Ukomprimerede FASTQ-filer er genomdata fra patienter, forskerindberetning, long-read indberetning og en lille andel er til projektet Genome of Europe. 2) Pipeline-outputs omfatter bl.a. data til reanalyse, farmakogenetik og til brug ved fremtidige projekter bl.a. Pangenomgraf. 3) gVCF'er danner grundlag for de nationale vidensdatabaser, altså Frekvens- og Variantdatabasen. 4) Øvrige data vedrører projekter indenfor personlig medicin, hvor data stilles til rådighed for forskere og klinikere med henblik på for forebyggelse, udvikling af behandlinger og forbedret diagnostik.

For yderligere uddybning af data og kapacitet henvises til præsentationen (bilag 2).

Løsninger

Dimensionering af den fremtidige platform:

Den overordnede strategiske tilgang til en ny storageplatform er at etablere en enkel, robust og gentagelig platform, som understøtter organisationens nuværende og fremtidige behov uden unødigt kompleksitet.

Kapacitetsbehovet er fastlagt ud fra en effektiv dimensionering, hvor der kun investeres i den kapacitet og performance, der er dokumenteret på baggrund af faktiske driftsdata fra det seneste år. Platformen designes til lineær skalering, så kapaciteten kan udvides efter behov (just-in-time). Desuden prioriteres en stabil arkitektur fremfor 'cutting-edge' teknologier for at sikre stabilitet og kosteffektivitet. I den nye infrastruktur samles de eksisterende storage-platforme til en samlet softwaredefineret-storage-platform løsning (SDS-løsning). En SDS-løsning imødekommer krav om hardwareagnostisk arkitektur, som reducerer leverandørfhængigheden, og som dermed sikrer fleksibilitet og omkostningseffektiv storageinfrastruktur med lavere samlede omkostninger.

Den nuværende infrastruktur består af flere forskellige platforme med separate administrationsværktøjer og flere supportaftaler, hvilket øger kompleksiteten og driftsomkostningerne.

Den foreslåede nye platform dimensioneres til 70 PB brugbar kapacitet. Med ca. 58 PB anvendt brugbar kapacitet giver dette en stram ekstra buffer på 12,3 PB (inkl. indtægtsdækket virksomhed (IDV) (6,6 PB)). Denne ekstra kapacitet er ikke reserveret til fremtidig vækst i genomdatabasen, men anvendes til nødvendig driftsreserve (overhead), nationale forsknings-

og analyseaktiviteter, samt indtægtsdækket virksomhed. De 70 PB brugbar kapacitet fordeles således:

- 57,7 PB til GDB/Den nationale infrastruktur for personlig medicin.
- 6,6 PB til IDV og øvrige national forsknings- og analyse aktiviteter.
- 5,7 PB (kun 8%) til driftsbuffer (til vækst, snapshots og vedligehold herunder fejl- og hardwareudskift).

Der investeres i samme kapacitetsevne som i dag, men desto længere udskiftningen af storage udskydes, desto dyrere og mere risikabel bliver udskiftningen. Bare over det seneste år er HPC-/NAS-miljøet vokset med ca. 26%, og objektstorage er vokset med ca. 27%. Det forslåede kapacitetsbehov er minimumskapacitet, for at kunne imødekomme fremtidige behov for platformen, og dimensioneringen forudsætter bedre datastyring, hvor sletning og komprimering af data er nødvendigt.

For at gøre storageforbruget bæredygtigt foreslås løbende oprydning i form af komprimering, datalivscyklus, samt en mere effektiv arkitektur med færre datakopier og smartere lagring. Der skal indføres klar dataklassificering efter formål, ejer, livscyklus og sletteregler, så der kan skelnes mellem permanent GDB-lagring, værdiskabende afledte dataprodukter og midlertidigt analyse- og arbejdsdata. Samtidig må det anerkendes, at pipelines og analysemiljøer kræver midlertidige arbejdskopier, hvorfor en vis duplikering og buffer er uundgåelig. Platformen skal understøtte genanvendelse af eksisterende hardware, hvor det er teknisk og økonomisk hensigtsmæssigt. Der igangsættes et arbejde med klar dataklassificering, sletning og komprimering samt etablering af livscyklusstyring, så midlertidige analyse- og arbejdsdata kan slettes, når de ikke længere er nødvendige. GDB og tilhørende analyseplatforme skal fortsat forstås som en datafabrik, der skaber værdi for patienter, klinikere og forskere, og ikke kun som et arkiv for rå genom filer.

Som en del af den nye infrastruktur etableres en moderne og skalerbar offsite backup- og disaster recovery-løsning, som en integreret del af den samlede storagearkitektur. Formålet er at sikre geografisk adskilt databeskyttelse af både produktions- og arkivdata, så kritiske data og services kan genetableres efter større kompromitteringer, ransomwareangreb, fysiske hændelser eller andre katastrofescenarier. Samtidig skal løsningen reducere kompleksiteten i den eksisterende backupinfrastruktur gennem fælles administration, automatisering og integration med den softwaredefinerede storageplatform. Valget understøtter de stigende krav til digital robusthed, cyberresiliens og beredskab samt behovet for en langsigtet og skalerbar beskyttelse af samfundskritiske data, se bilag 4.

Tilsvarende internationale infrastrukturer for personlig medicin, f.eks. Genomics England opererer med en infrastruktur i samme størrelsesorden som NGCs. Genomics England opgjorde

i deres årsrapport for 2022 ca. 135.000 genomer på en storage-plattform på ca. 65 PB til håndtering af både genomiske og kliniske data.²

Nærværende udbud omfatter ikke storage til en opskalering af Visionen for Bedre Brug af Sundhedsdata, da der endnu ikke foreligger en plan for, hvor meget det skal fylde. Når behovet er kendt, kan der skales op med yderligere kapacitet jf. princippet om skalerbarhed som beskrevet i strategien. Der vil være tale om indtægtsdækket virksomhed på lige fod med øvrige eksterne brugere.

Det videre forløb

På baggrund af kvalificering i den tekniske arbejdsgruppe udarbejdes en indstilling til bestyrelsen, som fremsendes med tilrettet materiale med henblik på at bestyrelsen kan godkende udbudsmaterialet, kapacitetsbehov, samt igangsættelse af udbud.

Bilag

Bilag 2: Powerpoint præsentation vedrørende storage kapacitetsbehov.

Bilag 3: Tidsplan for storage udbud.

Bilag 4: Storage- og udbudsstrategi

² [GenomicsEngland_AnnualReport_2022.pdf](#) (s. 27)