

Optimering af fremtidens dataanalyser

Dato & Information

Til:	Afdelingen for Husdyr/Heidi Lund Hyttel
Udarbejdet af:	Lasse Magnussen
Dato:	18. december 2025
Emne:	Optimering af data lifecycle management procedurer

RESUMÉ OG ANBEFALING

Formål:

Notatets formål er at kortlægge og anbefale systemer, der effektivt, sikkert og omkostningseffektivt kan indhente og lagre store mængder video- og billeddata fra forsøgsinstallationer.

Hovedanbefaling:

Det anbefales, at der arbejdes videre med fundene fra scenarie 3 (Hybridløsning).

Økonomisk Implikation:

Etableringen af en hybridløsning kan være omkostningstung. Afhængigt af det endelige databehov kan prisen for hardware og lagringsomkostninger svinge fra få tusinde til millioner af kroner. Et forsigtigt estimat for den anbefalede løsning indebærer indkøb og installation af en professionel NAS-enhed, som kan facilitere lokal lagring og bearbejdning. Priserne starter fra 45.000 kr. - 90.000 kr., hvor der stadig skal medregnes en månedlig omkostning for langsigtet cloudlagring.

1. Formål og Baggrund

Dette notat er udarbejdet som en delleverance til AP1 i projektet "Optimering af fremtidens dataanalyser". Formålet er at kortlægge, informere om og anbefale systemløsninger, der kan sikre en effektiv, sikker og omkostningseffektiv indhentning, adgang til og langsigtet lagring af de stigende mængder video- og billedmateriale fra forsøgsinstallationer.

For at garantere en nuanceret analyse af og løsning på problematikken er der opstillet tre scenarier, som alle i et vist omfang kan løse eller optimere den nuværende dataindsamlingsprocedure. Disse scenarier er:

1. **Optimering af nuværende praksis:** Fastholdelse af systemet, men fokus på procesoptimeringer.
2. **Systemskifte:** Fuld overgang til enten Cloud- eller lokal NAS-lagring (Network attached Storage). En NAS-enhed er en specialiseret server, som kun håndterer datalagring og deling.
3. **Hybridløsning:** En kombination af lokal styrke og cloud-skalerbarhed.

Baggrunden for notatet er et behov for optimering af databehandlingsprocedurerne, da forsøgsdesignene løbende bliver mere komplekse med mindre marginaler, hvilket resulterer i eksplosivt voksende datamængder. Dette repræsenterer en stor udfordring, da lagring af store datamængder udgør en betydelig omkostning.

Udfordring

Der er i øjeblikket flere udfordringer i ambitionen om at sikre effektiv dataindhentning, omkostningseffektivitet og sikker lagring af data. Her skinner særligt tre problematikker igennem:

- Der er dårlig netforbindelse i staldene, som komplicerer eventuel live-adgang til og analyse af data.
- Det er dyrt at lagre de store datamængder i skyen, hvorfor data nuværende transporteres manuelt på harddiske.
- Den dårlige forbindelse og omkostningstunge lagring resulterer i flaskehalse, hvor data ikke kan tilgås eller bearbejdes rettidigt.

Der er flere systemer i spil, hvorfor problemet er blevet inddelt i to specifikke spørgsmål:

1. Hvilket system kan facilitere en nem indhentning af og adgang til data?
2. Hvilket system kan sikre en effektiv, billig og sikker lagring af store datamængder?

2. Scenarie 1: Behold nuværende system og optimér processer

Dette scenarie tager udgangspunkt i "the least disruptive practice". Den grundlæggende proces opretholdes, såsom brug af Azure Blob Storage, lejet kameraudstyr (og dertilhørende kontroludstyr) og delvis manuel håndtering, men der iværksættes en række målrettede optimeringer for at løse problemerne med netværksdækning, omkostninger og sikkerhed.

2.1 Effektivitetsoptimeringer

Den primære udfordring for effektiviteten er det ustabile netværk og de manuelle arbejdsgange. Hvis løbende-lagring og fjernadgang skal blive en realitet, kræver det et teknisk løft.

Stabilisering af Netværk (Optimering af upload, analyse og adgang):

For at kunne sikre den bedste adgang til data fra en Network Video Recorder (NVR), med ca. 12 kameraer (nuværende installation) i HD (720p, 15fps), kræves en stabil upload-hastighed på ca. 35Mbps. NVR-enhedens funktion er, at den modtager, optager og gemmer videoerne fra kameraerne.

- *Løsning 1: Optimering af 4G/5G-infrastruktur:* Det anbefales at vælge TDC NET som netværksudbyder, da de i øjeblikket dækker landdistrikterne bedst grundet brug af lavfrekvente bånd (700/800MHz), der trænger bedre igennem forhindringer.
- *Løsning 2: Hardware-opgradering:* Der kan investeres i professionelle routere med failover-funktionalitet (dobbelt SIM/eSIM), der automatisk skifter fra TDC NET's netværk til TT-netværket ved udfald. For at maksimere signalet kan routeren kobles til en ekstern MIMO-antenne (Multiple Input, Multiple Output), placeret udendørs, højt og rettet mod nærmeste mobilmast, så den kan forstærke signalet.
- *Løsning 3: Satellit (LEO):* Starlink kan levere 10-30 Mbps og der forventes højere hastigheder i 2026. Det er en dyr løsning med en startpris på ca. 8.500-14.000 kr. i hardware og et løbende abonnement på minimum 2.900 kr./md. ved høje datakrav. Denne løsning kræver frit udsyn til himlen, hvilket gør den sårbar overfor vejrforhold.

Optimering af Procedurer og Kommunikation:

De manuelle processer er sårbare over for kommunikationsfejl.

- *Koordinering:* Der skal etableres faste procedurer for, hvornår data hentes, hvor det lagres og hvad det kaldes, for at sikre overensstemmelse mellem kontor og udekørende teknikere.
- *Fælles Inventarliste:* Oprettelse af en fælles liste over tilgængelige data og harddiske kan begrænse flaskehalse.

- *Filhåndtering i Azure:* Da omdøbning af filer i Azure ikke er understøttet direkte og kræver flere "operationer" i Blob Storage (kopier, navngiv og slet), skal der aftales en fast procedure ved fejl for at undgå spildtid.

Standardisering af Hardware:

Den nuværende lejeaftale af hardware til installationerne hos Jysk IT medfører varierende kvalitet i lys, lyd og billede, da udstyret ikke er fast. Dette komplicerer arbejdet med maskinlæringsmodeller og giver en svingende databyrde.

- *Minimumskrav:* Der skal defineres en standard for udstyret i samarbejde med leverandøren for at definere datakrav, så den endelige databyrde kan reduceres.
- *Eget Setup:* Alternativt kan der indkøbes et standardiseret setup (NVR, 12 HD-kameraer, PoE switch, harddiske), hvor en billig løsning starter fra ca. 18.000-28.000 kr., hvilket sikrer ensartethed, men kræver oplæring i nye systemer. Dette kan hurtigt blive en dyr post, hvis data skal bearbejdes løbende eller live, da begge scenarier kræver væsentligt dyrere og mere hårdføre installationer (hårdt miljø). Hvis en NVR skal køre analyser, baseret på SEGES algoritmer og modeller, løbende eller live, så kræver det en dedikeret server med god video management software (VMS), som samtidig understøtter en åben platform.

2.2 Omkostningsoptimeringer

De nuværende omkostninger ved Azure Blob Storage kan blive unødigt høje. Særligt interaktion, lagring og udtræk af data er dyrt, hvilket reducerer brugen og dermed funktionaliteten. Det er derfor vigtigt at fokusere på datastyring.

Data Lifecycle Management (Lagringsniveauer):

Der er store besparelser ved at udnytte Azure's lagringsniveauer effektivt.

- *Strategi:* Data skal klassificeres, og der skal være faste procedurer for sletning og arkivering. Lagringsniveauerne skal bruges aktivt, så "Hot" kun bruges til aktive projekter, "Cool" til data, der skal tilgås månedligt, og "Archive" til data, der sjældent eller aldrig bruges.
- *Besparelse:* Lagring i "Archive" kan være op til 16 gange billigere end "Hot" (ca. 5.685 kr./md. modsat 91.543 kr./md. ved 1PB). Automatisering af denne flytning er en lavpraktisk, men yderst effektiv optimering.

Datakomprimering:

Filer skal komprimeres før upload. Ved en rekodning af video til H.265 (HEVC) format kan filstørrelsen reduceres med 50-65% uden betydeligt kvalitetstab. For 20TB data betyder det en reduktion til ca. 8 TB, hvilket direkte minimerer lageringsomkostningerne. Dette kan automatisk faciliteres på moderne NVR-installationer. Databyrden kan reduceres yderligere ved at indstille kameraerne til kun at optage ved bevægelse.

Reducering af Hardware-slid:

Transport af harddiske slider på udstyret. Ved at benytte live-upload eller "Cloud Bridge"-funktionalitet i moderne NVR-systemer (f.eks. Hikvision/Axis) minimeres behovet for fysisk håndtering. Hvis fysisk lagring fortsat er nødvendigt, bør der skiftes til Surveillance-grade HDD (billigere, tåler overskrivning, sårbare for stød) eller Enterprise SSD (dyrere, men stødsikre).

2.3 Sikkerhedsoptimeringer

- **Minimering af fysisk risiko:** Live-upload eliminerer risikoen for, at data går tabt, hvis en harddisk tabes eller beskadiges under transport.
- **Backup:** Nuværende NVR-RAID niveau, sikrer kun mod lokale problemer, ikke mod tyveri eller lokal ulykke, hvor alle harddiske beskadiges. Direkte upload til skyen eller en central server sikrer "off-site backup", hvilket er en god sikring mod datatab. RAID (Redundant Array of Independent Disks) er en teknologi, der kombinerer flere harddiske til én enhed for at forbedre datalagring, enten ved at øge hastigheden, give datasikkerhed (redundans, hvor data gemmes på flere harddiske) eller begge dele.

Opsummering af Scenarie 1 (Fordele & Ulemper)

- *Fordele:* Kræver mindst omstilling af arbejdsgange ("Least disruptive practice"), udnytter eksisterende kompetencer i Azure-plattformen, optimeringer kan implementeres trinvist, og der er lave startomkostninger.
- *Ulemper:* Fortsat stor afhængighed af internet i ustabile yderområder, løser ikke problemet med dyre "egress fees" (udtræk af data) ved lokal bearbejdning, og kræver disciplinerede manuelle procedurer.

3. Scenarie 2: Erstat systemet (Cloud eller NAS)

I dette scenarie erstattes den nuværende lagringspraksis med enten en ren cloud-løsning eller en ren lokal serverløsning (NAS/SAN). Begge løsninger faciliterer en praktisk og sikker adgang til data. Hvor Cloud repræsenterer den mest effektive (praktiske og sikre) løsning og NAS-løsningen den mest langsigtet besparelse, ved begrænset databyrde.

3.1 Ren Cloud-løsning (AWS, Azure, Google)

Her uploades alt direkte til skyen, og lokal hardware minimeres.

- Fordele: Uendelig skalerbarhed, praktisk adgang til data (reduktion af flaskehalse), minimal drift, ingen vedligeholdelse af servere og høj sikkerhed (automatisk redundans).
- Ulemper: Omkostninger ved brug og særligt "egress fees", internet-afhængighed og "Vendor Lock-in" (svært at skifte udbyder).

3.2 Ren NAS-løsning (Lokal Netværkslagring)

Her etableres centrale lagerversere (NAS) hos SEGES, som løbende indhenter data fra installationerne i staldene, så det kan lagres og tilgås nemmere af alle.

- Fordele: Lynhurtig dataadgang over lokalt netværk (LAN), mulighed for live-adgang, ingen løbende licensomkostninger for plads (kun engangsinvestering) og fuld kontrol over data.
- Ulemper: Høje startomkostninger til hardware, internt driftsansvar (backup/opdateringer) og dyrere skalering sammenlignet med skyen.

4. Scenarie 3: Hybridløsning (Anbefalet strategi)

Dette scenarie kombinerer styrkerne fra NAS og Cloud for at eliminere svaghederne ved det enkelte scenarie. Det anbefales at implementere lokale servere til aktiv databehandling og at bruge skyen til langvarig arkivering, da lagring af store datamængder over 500 TB lokalt bliver meget økonomisk udfordrende.

Strategi:

1. **Lokal Behandling (NAS/Server):** Der opstilles en lagerserver (NAS), som enten indhenter (kræver optimering af nuværende forsøgsinstallationer), eller fungerer som manuelt lagringspunkt for den indsamlede data. Dette sikrer nem løbende adgang til data og reducerer omkostningerne ved cloudlagring.
2. **Cloud til Arkiv:** Når data er færdigbehandlet, eller rådata skal sikkerhedskopieres, flyttes det automatisk til billig cloud-storage (Archive Tier).
3. **Netværksoptimering:** Understøttes af den, i Scenarie 1, beskrevne netværksoptimering (4G/5G Failover/MIMO), som kan forbedre synkroniseringsmuligheder med installationer i staldene.

Opsummering af Scenarie 3 (Fordele & Ulemper)

- Fordele: Optimal lokal hastighed, ingen flaskehalse, billig adgang til aktive data, automatisk upload til skyen for langsigtet lagring, minimerer dyre "Hot"-lagringsomkostninger og undgår "egress-fees".
- Ulemper: Øget kompleksitet (kræver administration af to systemer), risiko for synkroniseringsfejl og begrænset lokal lagerplads (afhængig af hardware). Dette scenarie løser kun problematikkerne omkring sikring af praktisk og billig lagring. Hvis live eller periodisk automatisk analyse, lagring og adgang til data på skal sikres, så kræver det en optimering af installationerne i stalden, som pointeret i scenarie 1.

5. Sammenligning af Scenarier

Scenarier	Initialinvestering	Fordele	Ulemper
1. Optimering af Status Quo	Lav til ingen: Mulighed for besparelser med nuværende licenser.	- Minimal omstilling. - Begrænset oplæring. - Trinvis optimering. - Lave omkostninger. - Nem implementering.	- Løser ikke effektivitetsudfordringer ved indhentning og lagring af data. - Fortsat ustabil internet. - Risiko for datatab ved manuel håndtering. - Flaskehalse forbliver en udfordring.
2a. Systemskifte: Cloud	Lav (indledende): Men drift kan koste 6.000-100.000 kr./md. for 1PB lagring afhængigt af niveau.	- Uendelig skalerbarhed. - Praktisk adgang til data. - Minimal drift. - Høj sikkerhed.	- Høje løbende omkostninger ved brug. - Stor internet-afhængighed. - Vendor Lock-in. - Højt krav om lifecycle management.
2b. Systemskifte: NAS	Høj: Hardware kan koste 250.000 kr. og op efter for store løsninger.	- Lynhurtig dataadgang. - Ingen løbende licenser. - Billig lagring på sigt. - Fuld kontrol over data.	- Høj startinvestering. - Internt driftsansvar. - Dyrt at udvide kapaciteten.
3. Hybrid-løsning (Anbefalet)	Middel - Høj: En billig/enkel NAS-enhed starter fra 11.000-28.000 kr., men kan stige til 250000kr. og op ved enterprise skala. Hvis adgang til installationer i staldene skal optimeres ift. live adgang, analyse og lagring af data, kræver det en væsentlig investering i hardware	- Optimal lokal hastighed. - Ingen flaskehalse. - Automatisering af lagring. - Minimerer dyre "Hot"-omkostninger. - Undgår egress-fees.	- Kompleksitet (to systemer). - Synkroniseringsrisiko. - Løser ikke selvstændigt praktiske udfordringer i staldene uden optimeringer og stabilt netværk.

6. Anbefalinger og næste skridt

For at realisere hybridmodellen og løse de akutte problemer anbefales følgende konkrete tiltag:

6.1 Infrastruktur og Netværk

For at muliggøre fjernadgang og potentielt live-upload skal netværksstabiliteten sikres:

- *Udbydere:* Det anbefales at benytte TDC NET via professionelle routere med failover-funktionalitet (dobbelt SIM).
- *Antenner:* Routerne skal kobles til eksterne MIMO-antenner placeret højt, udendørs og rettet mod nærmeste mast.

- *Perspektiv:* På sigt (2026+) kan LEO-satellitløsninger, såsom Starlink, blive relevante, men vurderes som omkostningstunge og upraktiske til formålet.

6.2 Softwareplatform (VMS vs. NVR)

Hvis målet på sigt er at sikre løbende eller live-analyse af data, anbefales det at undgå lukkede "plug-and-play" NVR-bokse, og i stedet kigge mod en åben Video Management Software (VMS) platform installeret på dedikerede servere, som er tilkoblet installationerne i staldene, som kan varetage analyse af data og håndtere løbende uploads til en lagringsserver (cloud eller NAS) på kontoret.

- *Systemvalg:* Udbydere der tilbyder en åben arkitektur (f.eks. Milestone XProtect og Genetec).
- *Begrundelse:* En åben platform tillader integration af egne maskinlæringsmodeller (f.eks. via Python/API) og automatiseret dataudveksling, hvilket eliminerer manuelle eksportprocesser. Serveren fungerer dermed både som optager (NVR) og arbejdsstation for maskinlæring.

6.3 Datastyring og Lagringsstrategi

For at reducere omkostningerne ved datalagring skal der implementeres strammere datastyring:

- *Lagringsniveauer:* Aktivt data ligger lokalt eller i "Hot"-tier. Inaktivt data (arkiv) flyttes automatisk til arkivniveau eller tilsvarende i en cloud-løsning, hvor prisen er markant lavere.
- *Standardisering:* Fastsæt lavest mulig standard for data, for at reducere databyrde.
- *Optagelser:* Optag kun, når det er nødvendigt (f.eks. automatisk optagelse ved bevægelse).
- *Komprimering:* Videofiler skal automatisk komprimeres (f.eks. til H.265 (HEVC)).
- *Rensning:* Implementering af automatiske sletteprocedurer for data.

7. Konklusion

Implementeringen af en hybridløsning, som kombinerer en lokal NAS-enhed med langsigtet lagring i skyen, understøttet af optimeringer, professionel netværksinfrastruktur og åben VMS-software vil løse de nuværende flaskehalse, sikre bedre adgang til data, reducere langsigtede omkostninger og garantere et højt niveau af sikkerhed. Det vil frigøre teknikerressourcer fra manuel transport og muliggøre den avancerede effektive dataanalyse,

sikre nemmere adgang til og økonomisk forsvarlige lagring af data, som er projektets kerneformål.

Handlingsplan:

1. **Afdæk behov:** Hvor meget data skal kunne lagres, hvor hurtigt skal det kunne tilgås og hvor længe skal det gemmes.
2. **Standardisering:** Fastsæt en standard (så lav som muligt) for video- og billeddata, så databyrden og variation minimeres.
3. **Procedureopdatering:** Fastsæt faste procedurer for databehandling generelt og "lifecycle management". Dette er særligt vigtigt for data lagret i skyen.
4. **Undersøg præsenterede løsninger:** Med udgangspunkt i de endeligt afdækkede behov, skal det bedste system vælges.
5. **Fastlæg budget.**
6. **Find leverandører og indhent tilbud.**
7. **Pilotprojekt:** Udarbejd en testinstallation.