



ONE SIZE DOES NOT FIT ALL

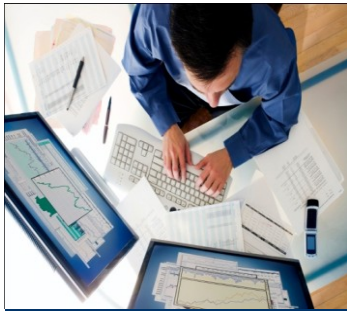
PER AXEL AAMOT

PRESALE, SYBASE NORWAY – SAP DATABASE & TECHNOLOGY

16. FEBRUAR 2012

STRATEGI

FEM MARKEDER



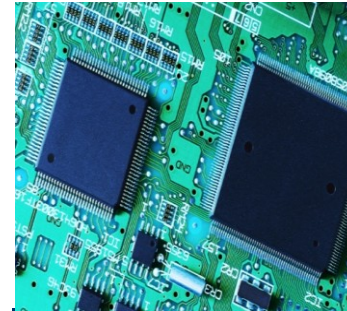
APPLICATIONS



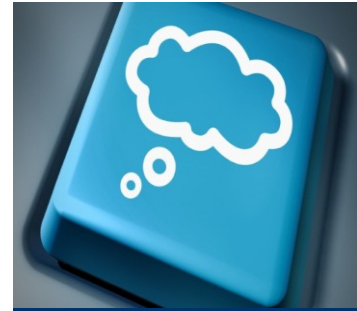
ANALYTICS



MOBILE



TECHNOLOGY
& DATABASE



CLOUD

HANA

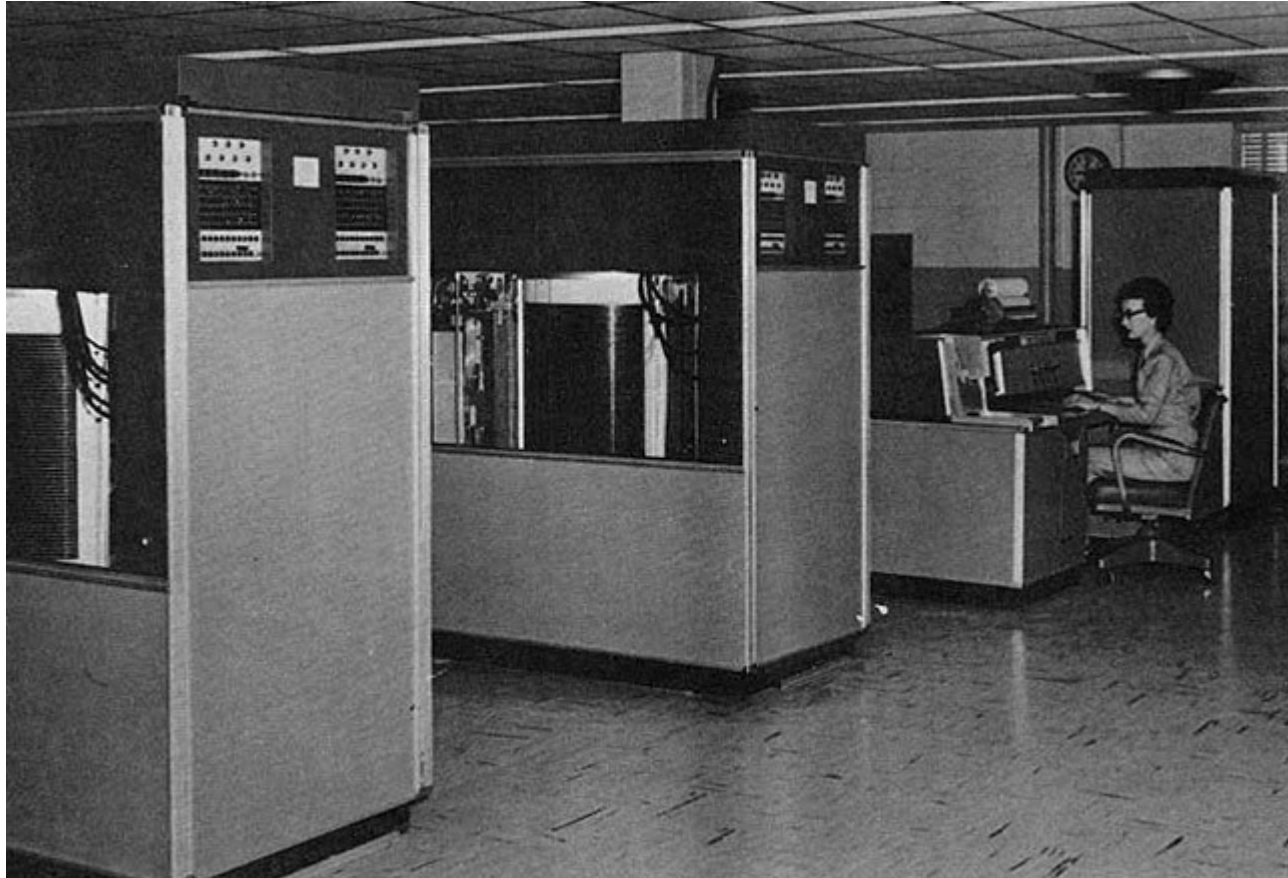
SERVICES

AGENDA

- Historikk, trender og produktutvikling
- Arkitektur for lagring av store datamengder
 - Kolonne- eller radbasert database
 - Minne, lokal disk , SAN, Cloud
 - Sammenkobling, konsolidering eller distribuering av data
- Løsninger og teknologier – hvor passer de inn i Big Data?
- Eksempler og referanser

HISTORIKK OG TRENDER

TILBAKEBLIKK



- IBM RAMAC 305 system med 2 IBM 350 diskette (1956→)
- \$10.000/MB

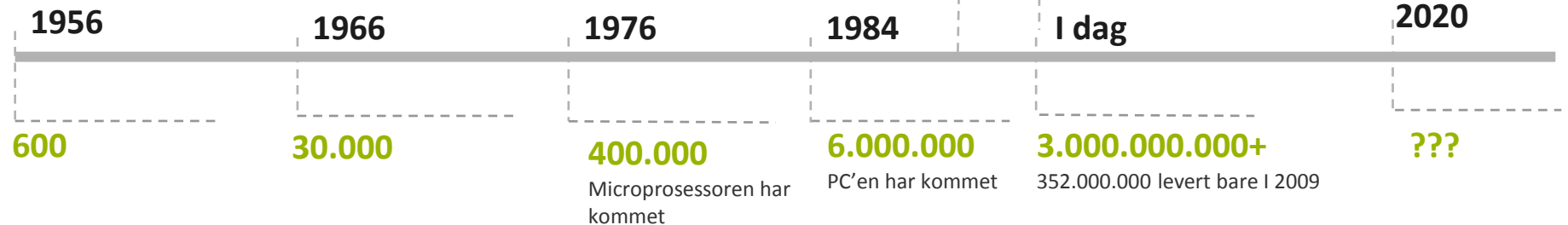
“DATAEKSPLOSION”

2000: 300 millioner internet brukere

2010: 1.7 milliarder brukere

5 milliarder brukere

Antall datamaskiner i verden



PDP-8 (1965)



Xerox Alto (1973)



IBM PC (1981)



Datavolum (generert, replikert, lagret)

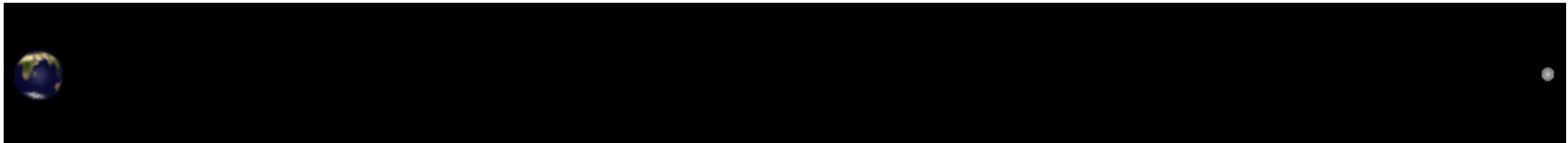
- 2005: 130 Exabytes
- 2009: 800 Exabytes
- Volum doubles hvert år – **1.8 Zettabytes** (10^{21}) i 2011

Vil nå 35 Zettabytes innen 2020

1.8 ZETTABYTES (ZB)



32 GB iPad
0.9cm tykk



”DATAEKSPLOSJON”

- Hver dag skapes det **2.5 Exabytes** data
- **75%** av informasjonen i det digitale univers skapes av mennesker
- Selskaper er ansvarlig for **80%** av informasjonen i løpet av sin levetid
- Antall filer som inneholder informasjonen i det digitale univers vil øke med **75x** i løpet av neste tiår, mens antall IT-administratorer som forvalter disse vil øke med **1.5x**
- Mengden data et typisk selskap vil administrere vil vokse **35x** frem til 2020
- Bare **½** av dataene som burde være sikret er det

TRENDER

- **SSD** priser begynner å nærme seg den magiske grensen \$1/GB
- Grense for lagringstetthet fremdeles ikke nådd
- Data Analytics som **SaaS**
- **Hadoop**
- Mobilitet og mobile devicer har gått fra å være novelty til del av IT-strategi
- **Crowdsourcing** (ex. Wikipedia)
- **Virtualisering** – 2010 første år fler virtuelle servere levert
- Cloud – **2%** av IT-budsjetter i dag; **2015: 20%** av informasjon vil komme i kontakt med Cloud tjenester

ARKITEKTURER

INNSAMLING OG LAGRING AV STORE DATAMENGDER

LAGRING AV DATA - DMBS

- **Rad-orienterte** databaser
 - Utviklet for OLTP-verden
 - Styrke i å håndtere transaksjoner fra mange brukere og ivareta constraints
 - Svakheter rundt håndtering av spørringer, spesielt når disse blir ad hoc og for ustrukturerte data, og med tung OLTP-trafikk samtidig
- **Kolonne-orienterte** databaser
 - Utviklet for OLAP
 - Styrke i å håndtere (ad hoc) spørringer og laste store datamengder raskt
 - Svakheter hvis brukt til OLTP

CLUSTERLØSNINGER



Ivareta SLA

- Beskytte mot nedetid ved node-feil ved at andre noder i cluster kan fortsette uten å være affektert
- Ivareta SLA under tung last gjennom lastbalansering av arbeid på tvers av flere noder



Maksimere ressursutnyttelse

- Konsolidere flere applikasjoner på et cluster og maksimere utnyttelse av ressurser og redusere under-utnyttet maskinvare
- Utnytte redundant maskinvare ved å distribuere last på tvers av nodene i et cluster



Redusere infrastruktur-kostnader

- Rulle ut cluster på standard, hyllevare maskiner, og redusere innkjøps-kostnader og support
- Gi en plattform som lett kan bli utvidet og vedlikeholdt ved å sette noder online og offline etter behov

LAGRING AV DATA

- Hadoop

- HDFS er en av byggestenene i Hadoop (den andre sentrale er MapReduce) – et distribuert filsystem for lagring av data
- NameNode for håndtering av filsystem metadata
- DataNodes for lagring av faktiske data
- Kjører på standard maskinvare/lagringsmedium

- Cloud

- Data lagret online i virtualiserte datalager drevet av 3.parts leverandører
- Selskaper kjøper eller leier kapasitet
- Ytelse? Pålitelighet? Sikkerhet? Ansvar? (CONOPS)



IN-MEMORY COMPUTING

- Fysisk I/O er ofte fremdeles den største flaskehalsen

- "Vanlig harddisk"

- Disk latency ligger på noen (5-10) ms (access time)
 - Overføringshastighet (når dataene er funnet) rundt 100 MB/s



- SSD

- Latency rundt 0.1 ms
 - Overføringshastighet 100-500 MB/s

- Minne (SDRAM)

- Latency rundt 10 ns
 - Overføringshastighet i størrelsesorden 15.000 MB/s



FLYTTING AV DATA

- Kopiering, backup og restore
- Meldings-utveksling, filoverføring, ETL
- Data-replikering – synkron og asynkron. Zero loss og korrupsjon
 - Database-replikering
 - Disk/blokk-replikering
- Konsolidering vs federation



ONE SIZE DOES NOT FIT ALL

Krav om lengre lagring av data

+

Ustrukturerte data: bilder, tekst, lyd og video

+

Vedlikehold av data for BI

+

Flere transaksjoner

DATAVEKST



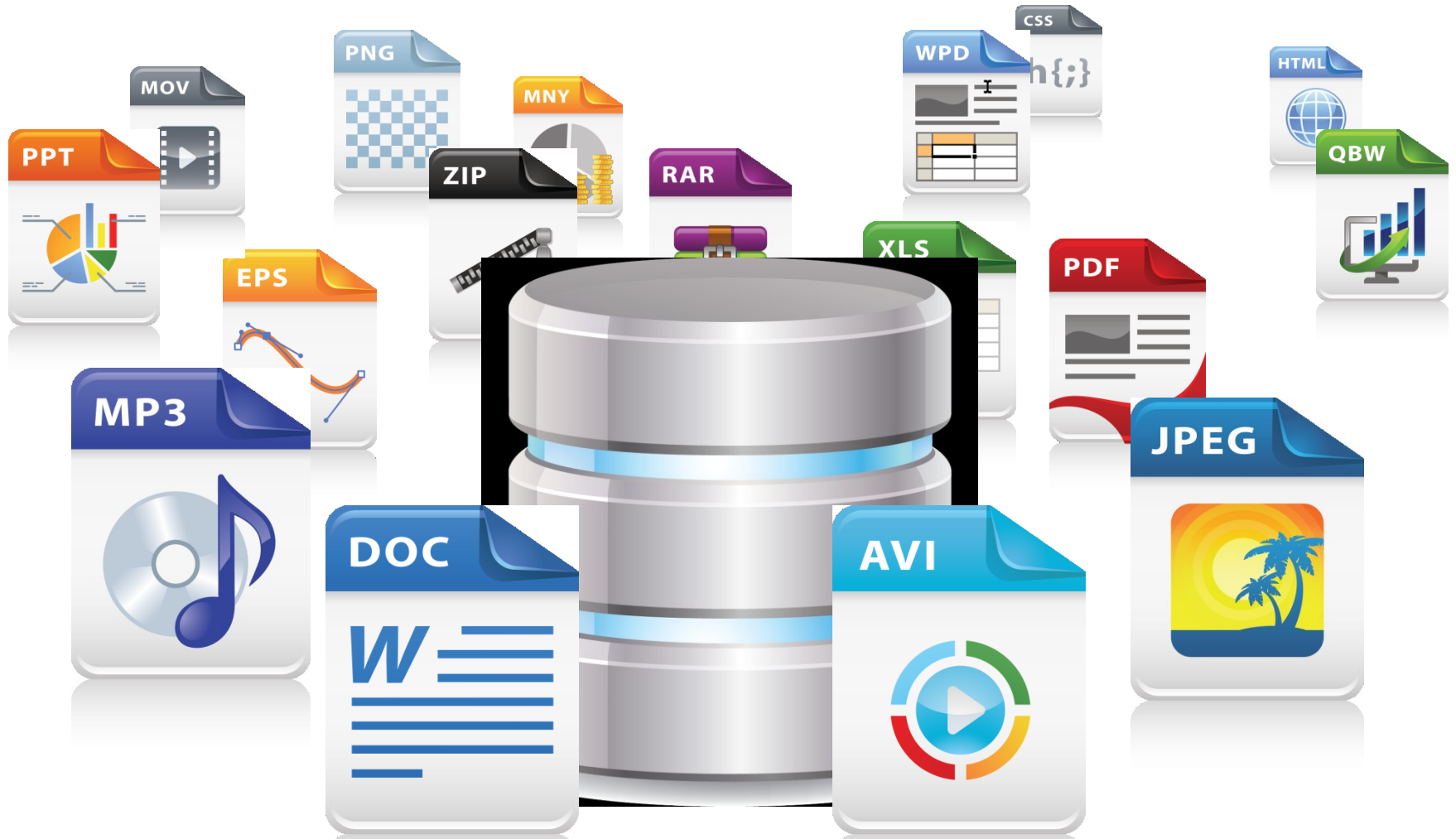
MULTIPLISER MED:

- x Replikerte Servere
- x Backup
- x Utviklings- og testmiljøer

**= EKSPONENTIELL
DATAVEKST**



HÅNDTERING AV USTRUKTURERTE DATA



SYBASE ADAPTIVE SERVER ENTERPRISE

- **Rad-orientert** DBMS, i versjon 15.7
- "Vokst opp" på Wall Street
 - 24 av topp 25 Globale banker og 46 av topp 50 bank/finans institusjoner
 - >50% av Wall Street-transaksjoner går gjennom ASE
- **Strategisk** transaksjonsmotor for SAP
- Nøkkel-egenskaper:
 - Sikkerhet og kryptering
 - Partisjonering
 - Virtualisering og clustering
 - In-memory teknologi

ASE 15.7 OG BIG DATA

- God håndtering av tekst og store objekter
- Nyskapning innen applikasjonsstøtte

Fleksibel håndtering av ustrukturerte data

Lagre store datavolum kostnadseffektivt

- Optimalisert lagring av ustrukturerte data
- Komprimering
- Redusert behov for temp space

Ytelse og skalerbarhet på parallell H/W

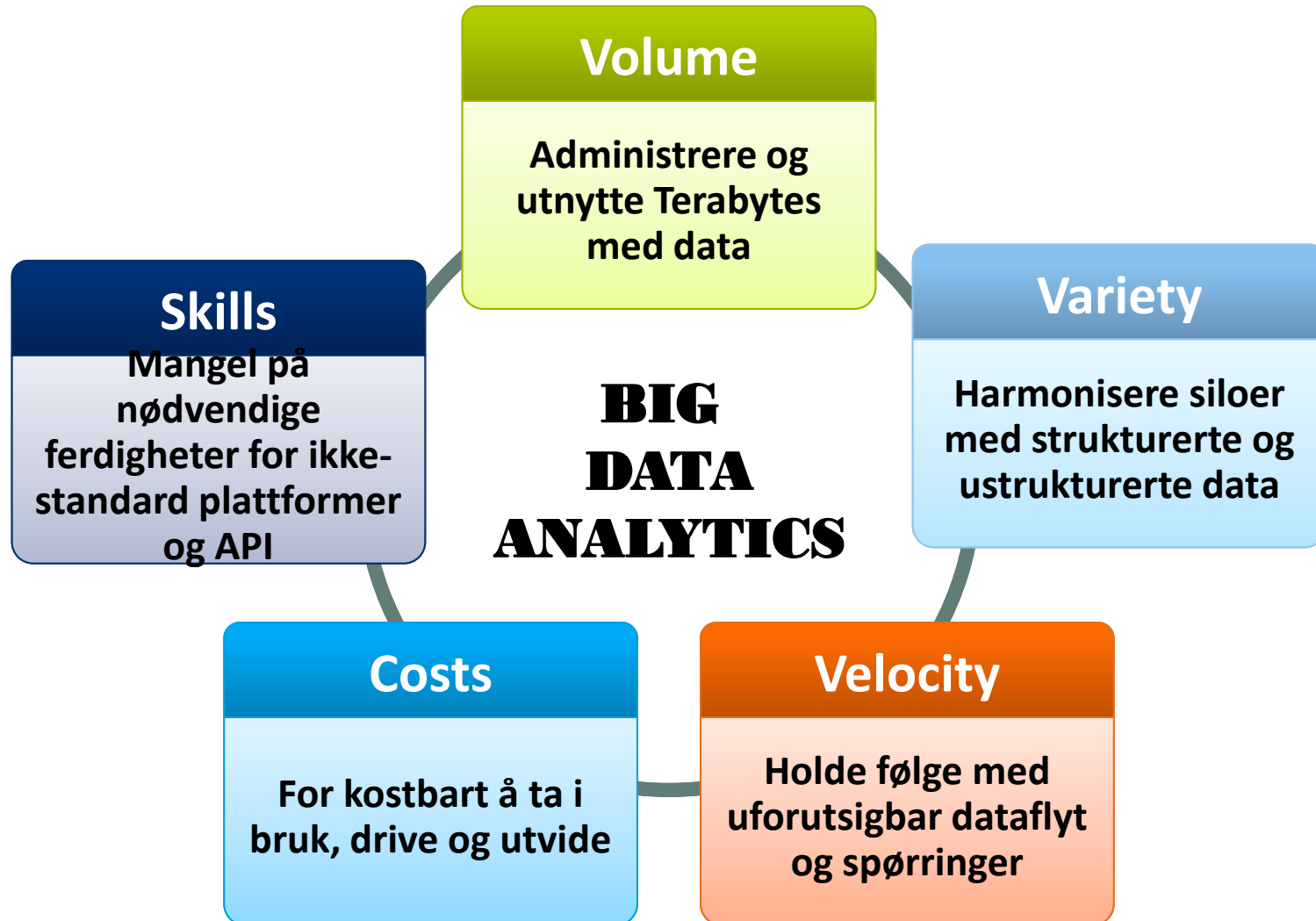
- Tråd-basert kjerne for parallell maskinvare
- Forbedringer rundt spørringer og system-funksjoner
- Støtte for blandet last (OLTP + OLAP)

BIG DATA ANALYTICS

- “The core applications for **Big Data** are in extremely scalable **analytics**, where the extremes involve processing a heavy volume, high velocity, staggering variety, and unpredictable variability of data types” – Forrester Research, Nov. 2011
- 3 viktige dimensjoner i forhold til Data Latency:
 1. Dataenes tidspunkt
 2. Hendelsens varighet
 3. Beslutningstid

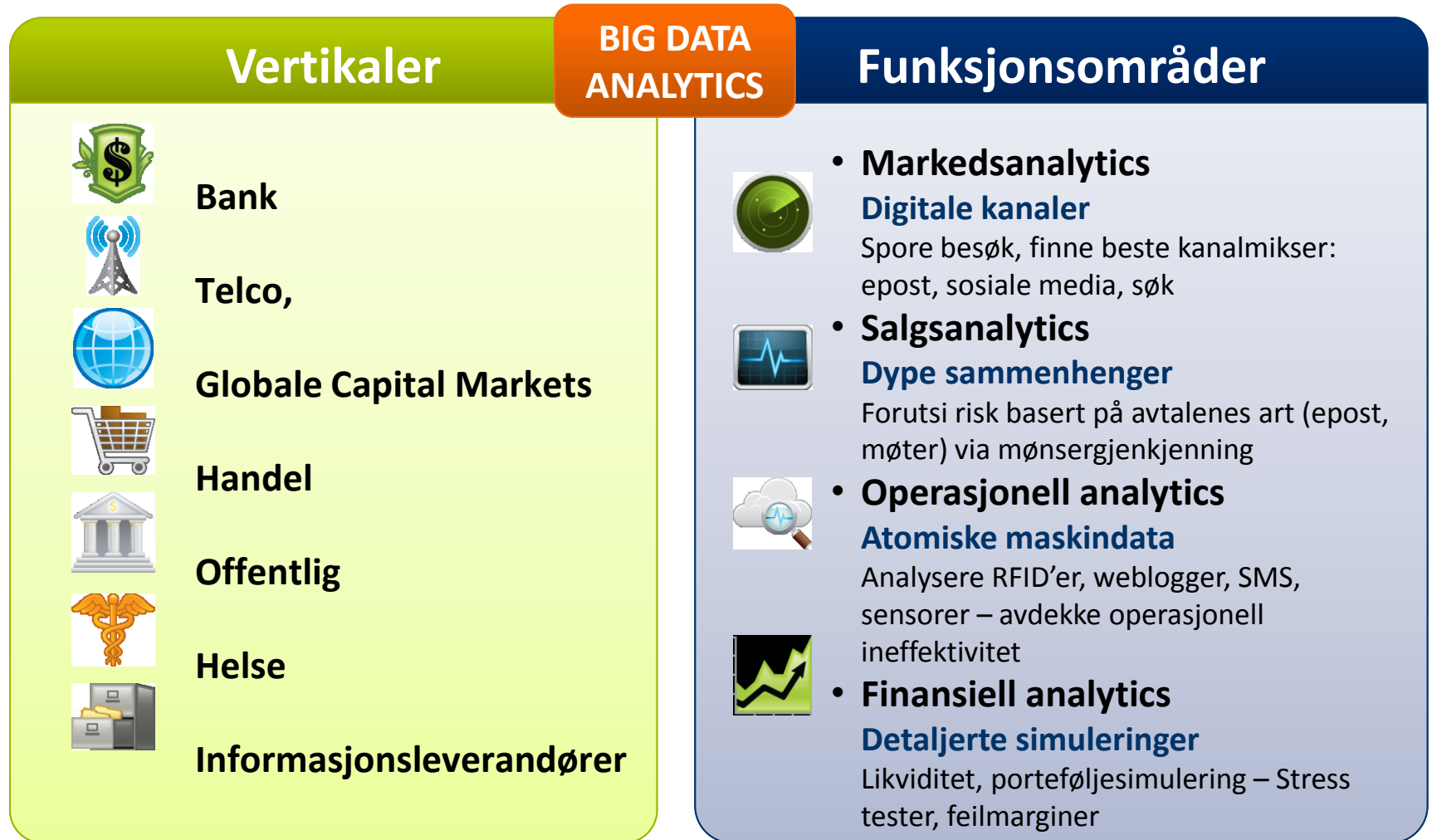
BIG DATA ANALYTICS UTFORDRINGER

Håndtere volum, variasjon, hastighet (velocity), kostnader og kunnskaper



BIG DATA ANALYTICS BRUKSOMRÅDER

Utbredt innen data-intensive vertikaler og funksjonelle områder



SYBASE IQ

Moden, enterprise-nivå analytic DBMS

MARKEDSLEDER

- Industriledende ytelse og skaleringsmuligheter
- Anerkjent EDW markedsleder av Gartner, Forrester
- Teknologipioner med 10+ patenter

BRUK

- 4500+ installasjoner hos 2150+ kunder
- ~200 nye kunder pr år (siste 4 år)
- Konsistent 96%+ kundetilfredshet

MOMENT

- 2 x DW vekstrate i markedet (siste 4 år)
- Regelmessige produktoppdateringer
- v15, v15.1 (2009), v15.2 (2010), v15.3, v15.4 (2011)



SUNGARD®



BNP PARIBAS

Telefonica

hmv
hmv.com



COMSCORE.

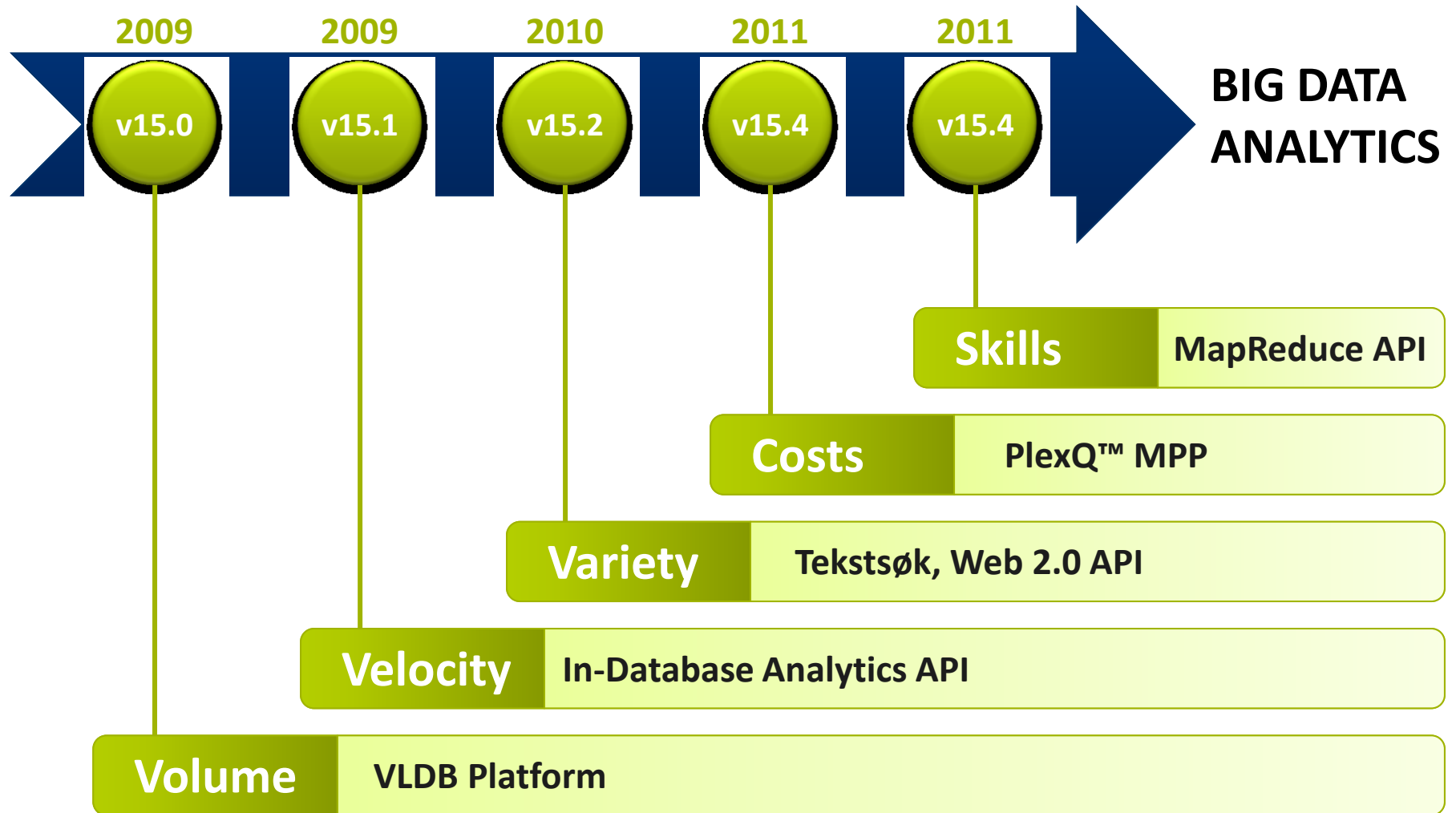


中国农业银行
AGRICULTURAL BANK OF CHINA

SYBASE® | An SAP Company

SYBASE IQ 15

En kraftig, Big Data Analytics plattform i utvikling



SYBASE IQ – KOLONNEBASERT DBMS

DRAMATISK REDUKSJON AV ANTALL DISK “HITS” – RASKERE SPØRRINGER

Spørring: Hvor mange MENN kjøper FORSIKRING i NORGE?

Rad-basert RDBMS

	Kjønn	Land	Forsikring
	M	NO	J
	M	SE	J
	K	DK	N
	M	NO	J
	M	DK	N
	-	FI	

$$\frac{10\text{M Rader} \times 800 \text{ Bytes}}{4\text{KB Page}}$$

= **2,000,000 I/O**

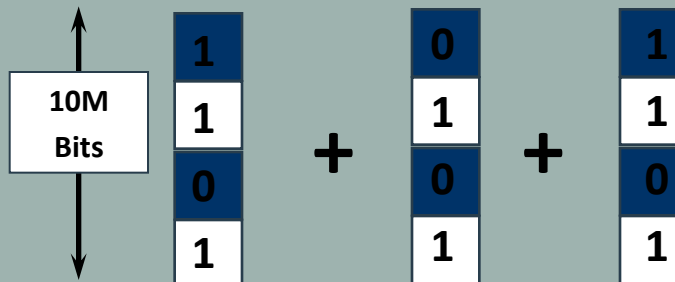
- Behandle store mengder ubrukt data
- Krever ofte full table-scan eller plasskrevende indekser/views

Kolonne-basert Sybase IQ

	Kjønn	Forsikret	Land
	M	J	DK
	M	N	NO
	K	J	SE
	M	J	NO

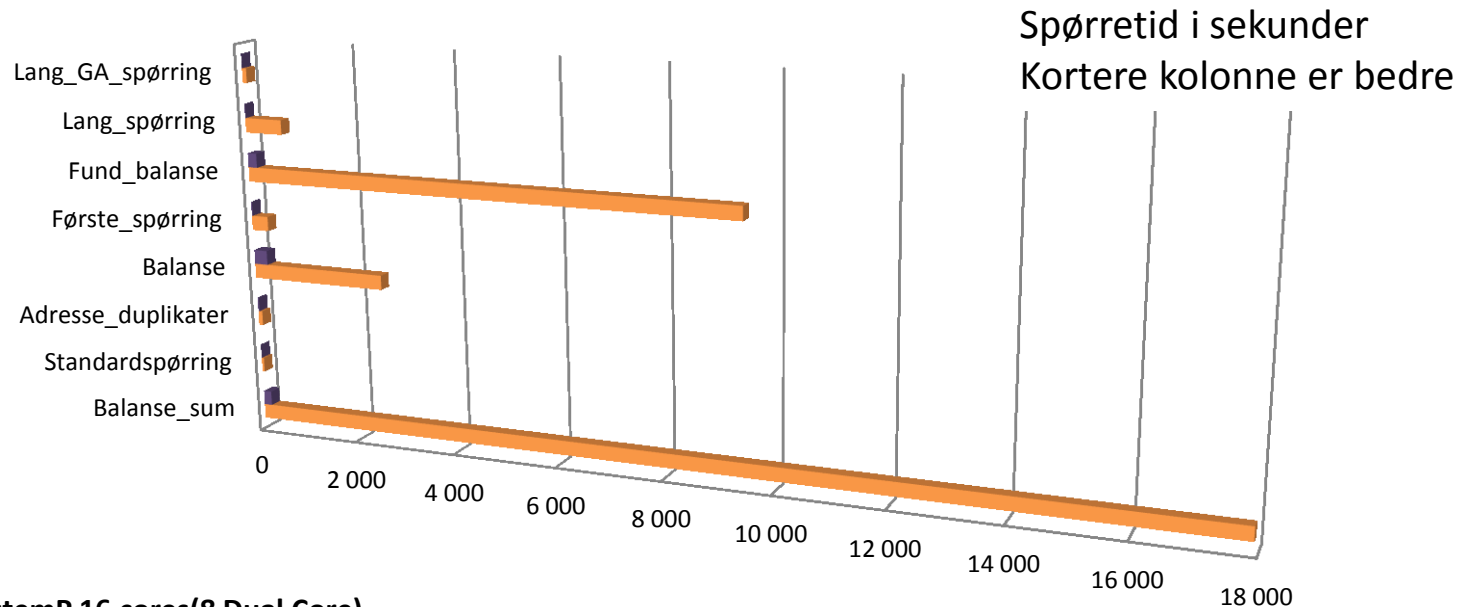
$$\frac{10\text{M Bits} \times 3 \text{ kolonner}}{16\text{KB Page}}$$

= **234 I/O**



KUNDE-EKSEMPEL

YTELSESFORBEDRING VED BRUK AV IQ



Maskinvare: IBM 570 systemP 16 cores(8 Dual Core)

	Balanse_sum	Standardspørring	Adresse_duplikater	Balanse	Første_spørring	Fund_balanse	Lang_spørring	Lang_GA_spørring
■ Sybase IQ (sekunder)	146,6	1,6	14,7	253	11,2	168,3	5,1	2
■ Oracle (sekunder)	17 927	38	71	2 589	318	9 452	723	78

Oracle Total: 519 minutter(31,196 sekunder)

Sybase IQ Total: 10 minutter (603 sekunder)

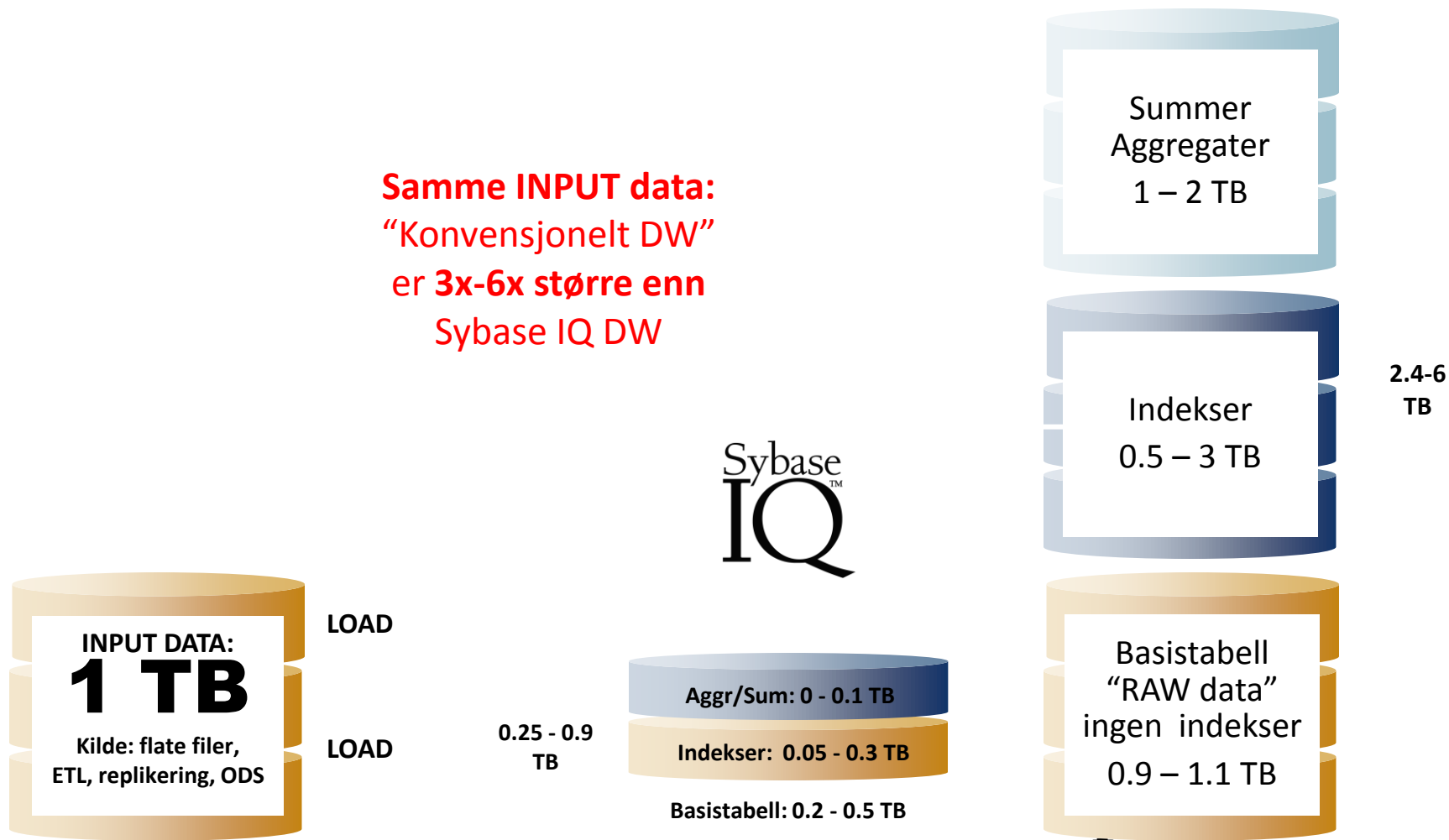
50x Ytelsesforbedring!

DATAKOMPRIMERING

DRAMATISK REDUKSJON I LAGRINGSKOSTNADER OG VEDLIKEHOLD

**Samme INPUT data:
"Konvensjonelt DW"
er 3x-6x større enn
Sybase IQ DW**

Konvensjonell DBMS



CONTINUOUS INTELLIGENCE™

*ØYEBLIKKELIG INNSIKT FRA DATA SOM BEVEGER/FORANDRER SEG
HURTIG*

- **Situasjonsbevissthet**
- **Øyeblikkelige svar**
- **Bedre beslutninger: komplett innsikt i tide**



Noen ganger kan man ikke vente på rapporten

SYBASE ESP – EN RAD PLATTFORM

En raskere og billigere måte å bygge CEP applikasjoner

- Redusere/eliminere avhengighet av spesialist-kunnskaper
- Korte ned implementasjons-/utrullingstid med 75%
- Forbedre smidighet og tilpasningsevne

Plasserer sanntid innen rekkevidde

- Databasekunnskaper kan videreføres
- Krever ikke nettverksprogrammering, hendelseshåndtering osv

Non-intrusive utrulling

- Hendelses-drevet arkitektur
- Tilpasses eksisterende datamodeller
- Legges på toppen av eksisterende systemer

SYBASE ESP – 4 BRUKSOMRÅDER



Situasjonsanalyse

Analysere sanntids hendelser for mønster som identifiserer forretningskritiske situasjoner

Eksempler:

- Avdekkelse av svindel
- SLA mål utenfor grenseverdier
- Kundeoppførsel indikerer problemer

Automatiske responser

Kontinuerlig justering av forretningsprosesser for å håndtere nå-situasjon

Eksempler:

- Automatisk prising basert på markedssituasjon
- Ressursallokering i sanntid



Strømtransformasjon

Vaske og berike sanntids datastrømmer, omdanne rådata til beslutningsverdig informasjon

Eksempler:

- Vasking og beriking av markedsdata
- Ordretrender I forhold til produksjonslinje

Kontinuerlig kunnskap

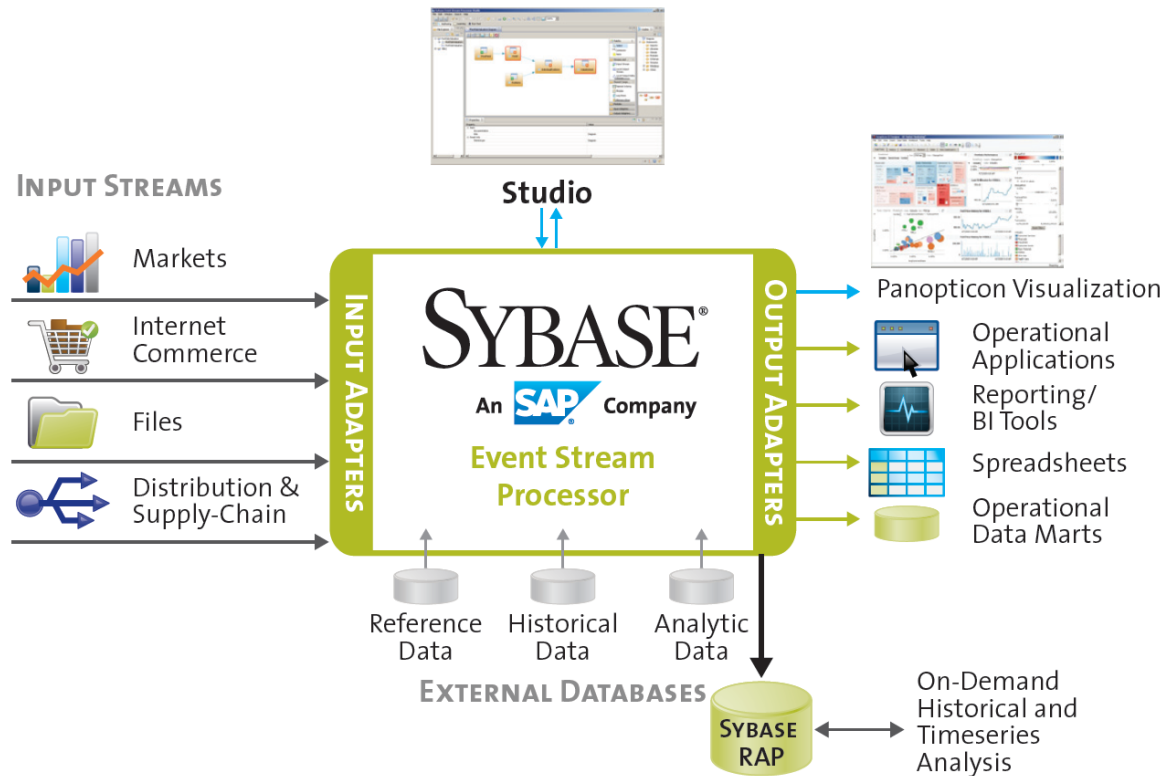
Sanntids dashboard som viser relevant informasjon for beslutningstagere

Eksempler:

- Sanntids P&L, risk & Exposure Mgmt
- Online markedsføring og tilbudsanalyse
- Systemovervåking vs ønsket ytelse

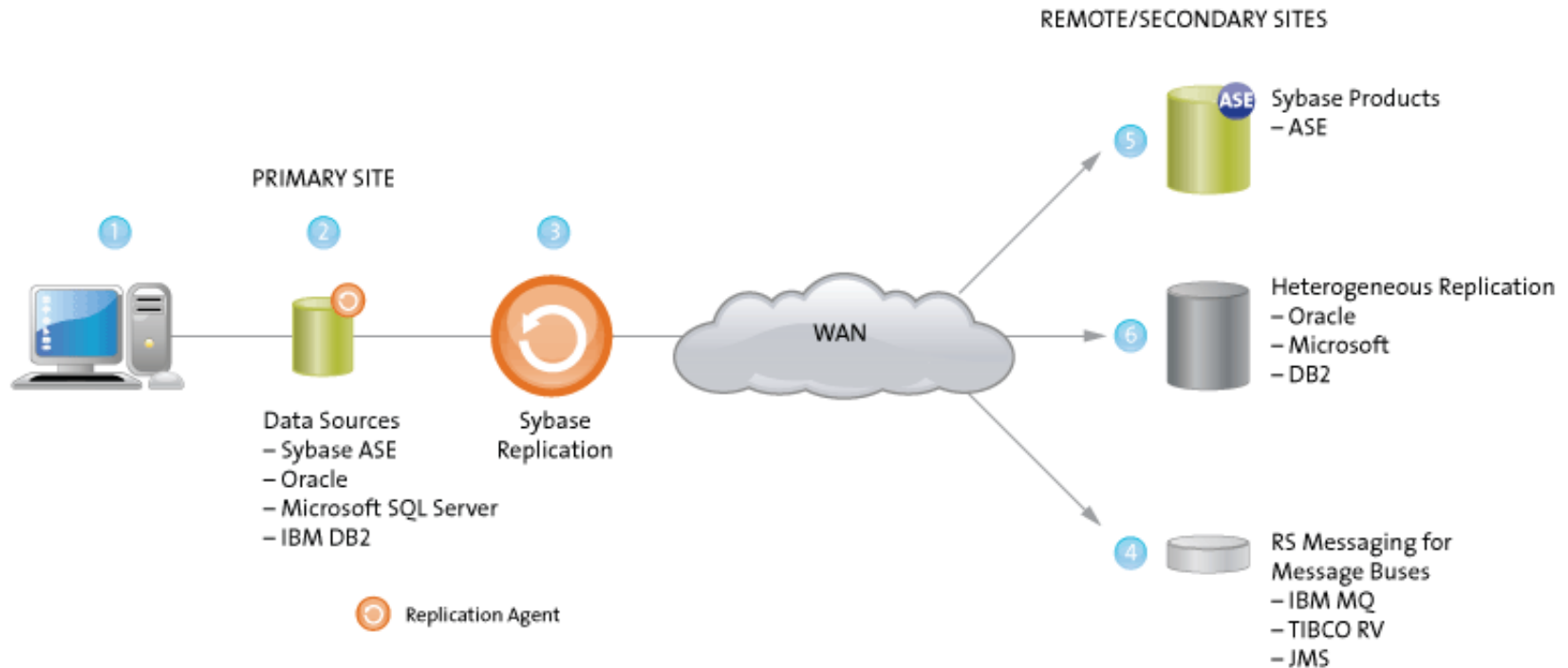


SYBASE ESP - ARKITEKTUR



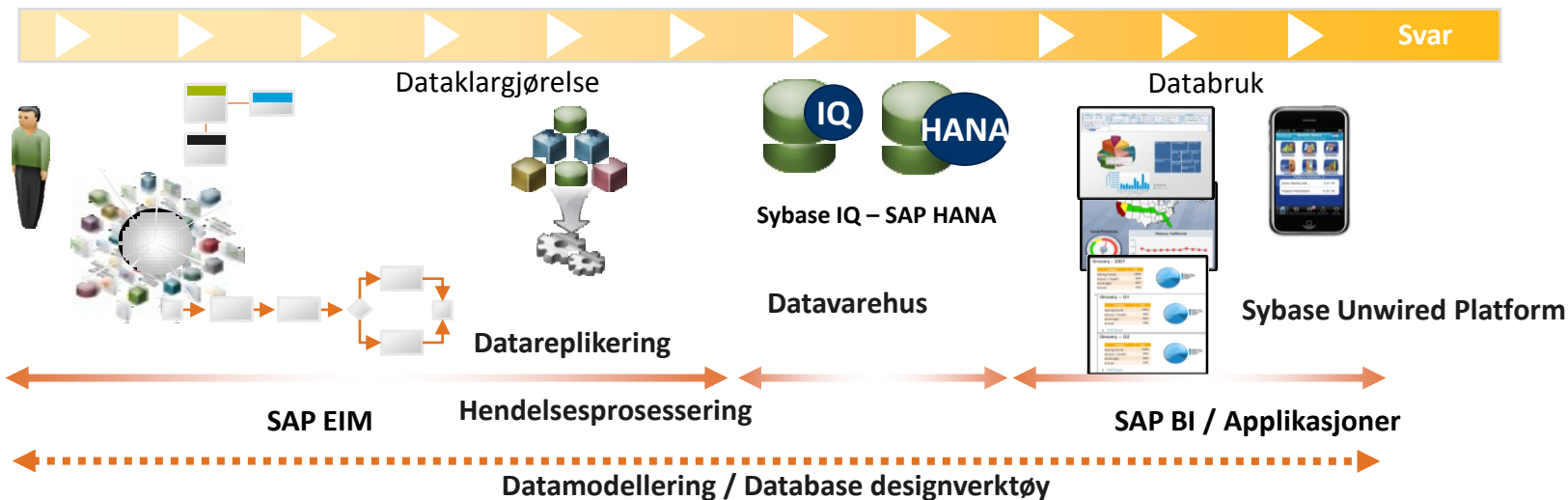
- Utvikling: Eclipse®-basert ESP Studio – CCL, SPLASH
- Integrasjon: connectivity mot standard datakilder og –konsumenter (meldingsbusser, databaser, filer, sockets, dashboard, XML, SMTP, FIX etc)
- Ytelse: hundretusenvis av hendelser pr. sekund, clustering og skalerbarhet
- Persistering av hendelser for historisk analyse via SYBASE RAP og IQ

DATAFLYT – SYBASE REPLICATION SERVER



- Ikke-intrusiv, heterogen løsning
- Gi sanntids-rapportering uten å affekttere produksjonssystem
- Effektiv datadistribusjon og –synkronisering
- Uavbrudt produksjon ved migrering

SAP OG SYBASE ANALYTICS: LEVERER EN KOMPLETT BI STACK SOM INTEGRERER BÅDE DATA-KLARGJØRELSE OG DATABRUK

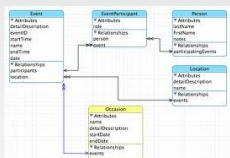


Datavarehus



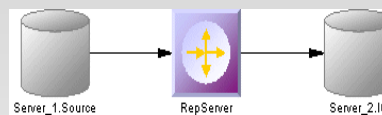
Sybase IQ
Kolonne-basert analytics DBMS

SAP HANA
Operasjonell/agile datamart

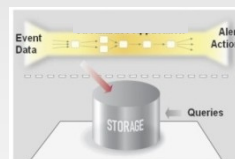


Sybase PowerDesigner
Ledende Modelleringsverktøy

+EIM



Sybase Replication Server,
Dataflyt og sanntids-synkronisering



Sybase Alert Streaming Platform (ESP), flyt av hendelsesdata for å gi operasjonell BI i sanntid

EKSEMPLER/REFERANSER

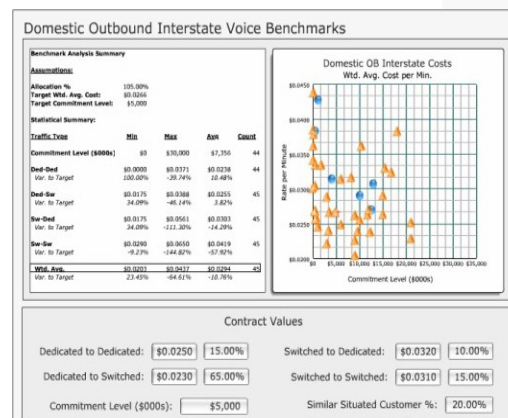
SYBASE IQ BIG DATA ANALYTICS I TELCO

Skalere med økende operasjonelle krav

Telefonica

Utfordring: Redusere kompleksiteten i eksisterende systemer for en av Europas største Telco-operatører – integrere 11 adskilte og uavhengige systemer og håndtere rask vekst i **data volume på totalt over 70 TB og 15000 spørringer pr dag fra mer enn 1000 aktive brukere.**

“The model saved between four and six times storage volume required compared to others in the market, with attendant reduction in hardware, support, administration”.



— Pedro Romera,
Systems Engineering
Manager, Telefonica, Spain

BANK + FINANCE

BNP Paribas: Securities Services

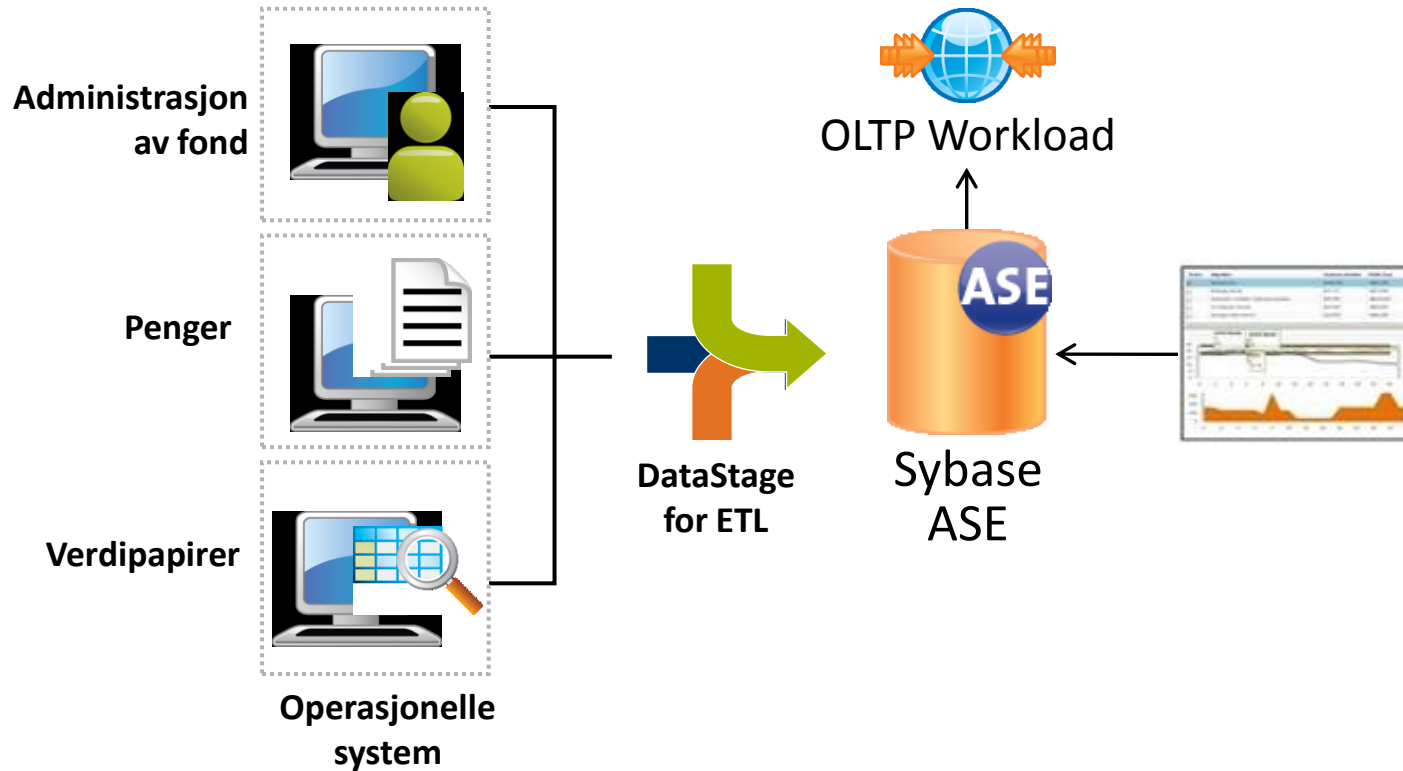


- Global tilstedeværelse: 32 land på 5 kontinenter
- Dekker over 100 markeder
- 22 proprietære lokale forvaltnings- og -clearingavdelinger
- 7,200 ansatte på verdensbasis
- 47 million transaksjoner fullført
- Administrerer USD 6,975 milliarder
- Administrerer 6,329 fond
- #5 på verdensbasis for “assets under custody”*

All figures 30 June 2011

* Kilde : Globalcustody.net — December 2010

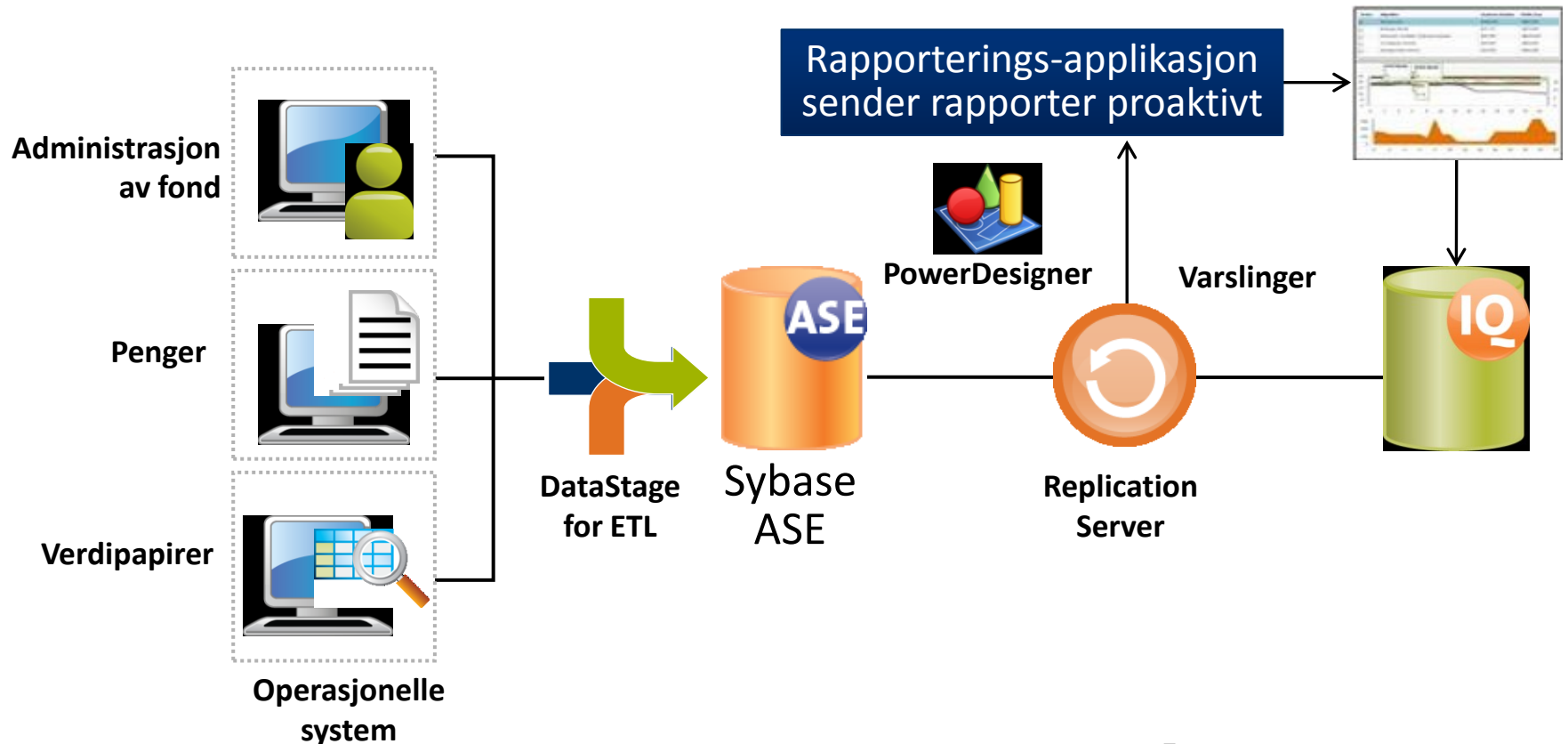
UTFORDRINGEN



LØSNINGEN



- Online kunder genererer 2X så mange ad hoc spørringer
- Rapporter kommer 3x raskere
- Frigir OLTP database ressurser til å aksellerere online transaksjoner
- Reduserer last på systemarkitekturen ved å sende ut i stedet for å polle rapporter

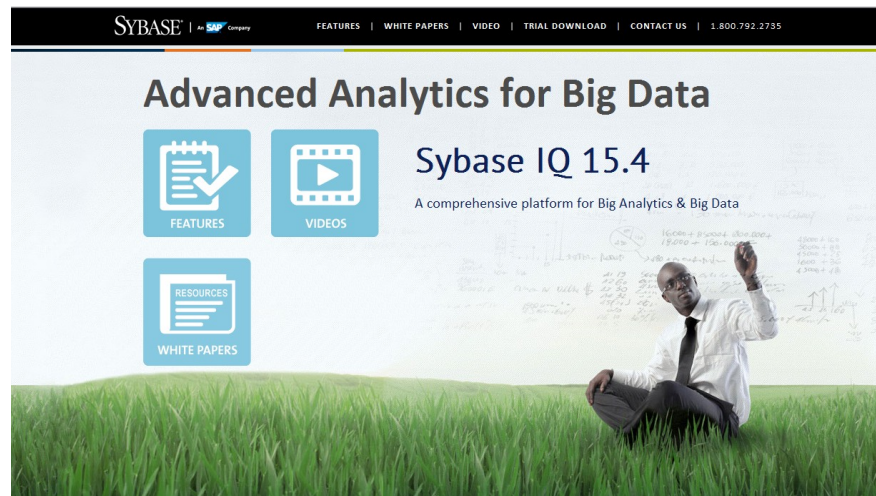


“Both Sybase ASE and Sybase IQ have proven to require very little system maintenance. They practically run on their own.”

MARC GUILLARD, DBA, BNP PARIBAS SECURITIES SERVICES

KONKLUSJON / MER INFO?

- Big Data er ingen gordisk knute men en potensiell skattekiste som kan åpnes med de rette verktøy
- <http://www.sybase.com/sybaseiqbigdata>



- Takk for oppmerksomheten

SYBASE®

An  Company