SAS Viya with Singlestore

Status, plans and demo

Jonas Lie-Nielsen

FANS Trondheim, 1. November 2023

Agenda

1. Arkitekturoversikt - hvordan funker dette
2. Compute – S2
   a. Demo rapport, ytelsesforbedringer
   b. Kode eksempel med bruk av base views og pushdown
   c. Plan for hva som skjer i November - work tables ++
   d. Eksempel - work tables
3. VA real time demo
4. VA - større tabell demo
5. Planer fremover VA

SAS Viya with SingleStore

Viya Foundations

Compute
SAS/ACCESS Interface to SingleStore

- Connectivity from Compute
- SAS Procedures Pushdown
- SQL Pass-Through facility
- SQL Functions
- Bulk Load

CAS
Standard Data Connector

- Serial data transfer (including multinode support)

SingleStore database is integrated with SAS

Viya with SingleStore

CAS
Viya with SingleStore Data connector

- Connectivity from CAS
- Order by/Group By Push down
- Calculated columns Push down
- MLA Action Push down
- SAS Meta data stored in db

In-Database
Technologies for SingleStore

- Parallel data transfer
- Model Scoring
- Data streaming

SAS Viya with SingleStore Use Cases

Reduce CAS Footprint

Support HUGE amounts of Data

Reduce CAS Node pool

Support limited memory scenarios

Replace Hadoop

Companies are exploring alternatives to Hadoop because of:

- Complexity
- Hardware Utilization
- Scaling Costs
- Old technology

Reduced Storage Costs

Leverage bottomless architecture

30% to 40% storage cost reduction

Automatic storage tiering

10:1 cost ratio

Move from SPDS

SPDS placed on No New Sales list Jan 2021

~1,000 licenses

Strong migration case to Viya with SingleStore

SAS Viya Singlestore Vision and Status

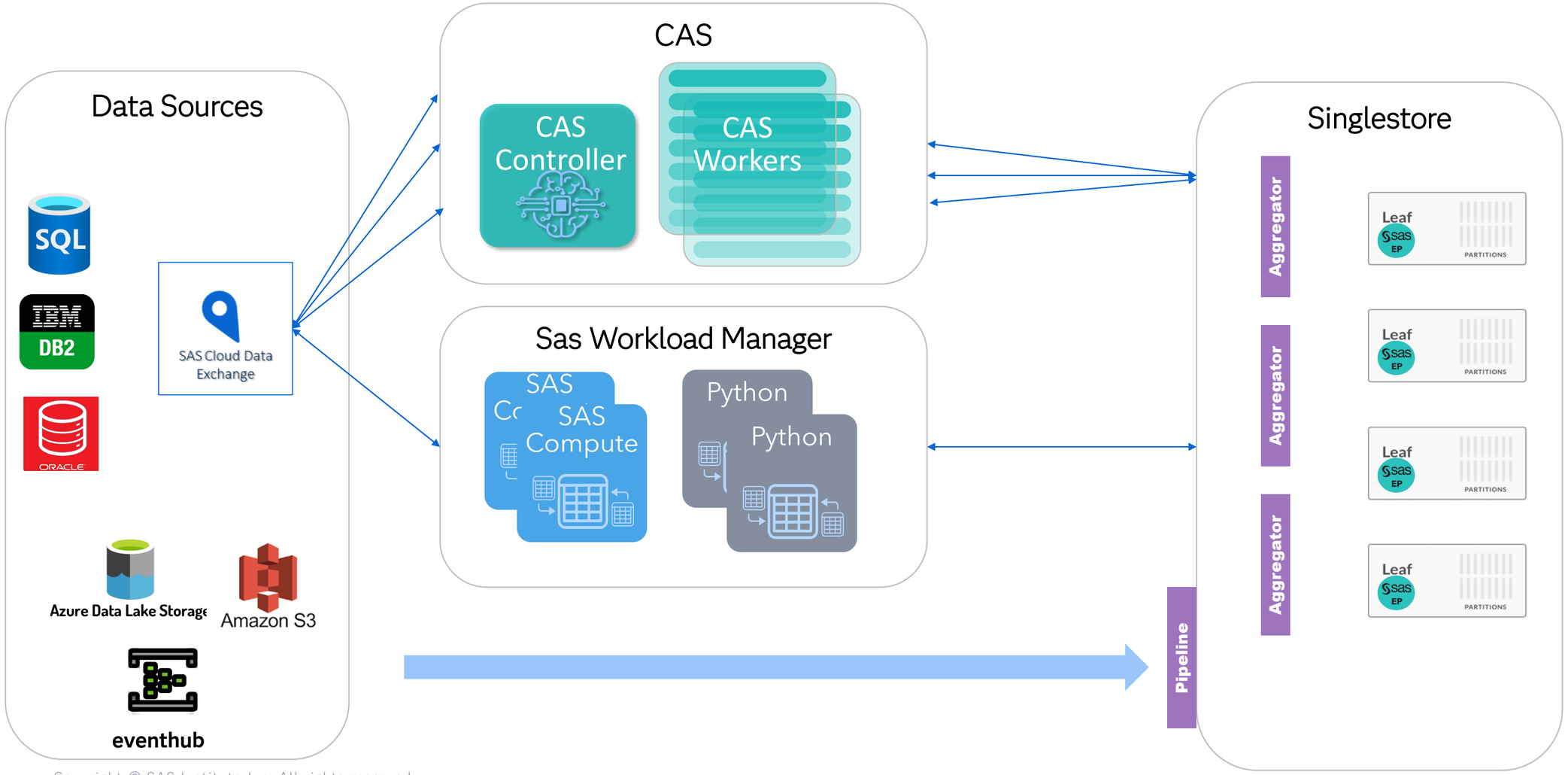
Status per the 2023.10 release

2023.10

- Enable Singlestore as an integrated Data platform for SAS Viya
- Make all data in S2 available in CAS without loading the data, but pushing down cas actions or stream the data
- Enable read/write between Viya and S2 at the same speed as internal read/write in SAS9
- Building real-time dashboards on S2 tables
- Pushdown of sql from compute to S2
- Pushdown of datastep from Viya to S2
- Support of S2 views in CAS

Data flow architecture

SAS Viya with Singlestore



Deliverables – Key Transformations (Posten example)

DI Transform	Studio step	Steerco plan for 80-90% readiness	Next Update
Extract	Query	23Q4	2023.09 (+4.8%)
Join	Query	24Q1	2023.09 (+3.2%)
SPDS Table Loader	Load Table	23Q4	2023.11
Append	Union Rows	23Q4	2023.09
Table Loader	Load Table	23Q4*	2023.10 (+76%)
SCD Type 2	Implement SCD	24Q1	Planning
Sort	Sort	24Q2	Planning
Splitter	Branch Rows	24Q2	Planning
Lookup	Lookup	24Q3	Planning

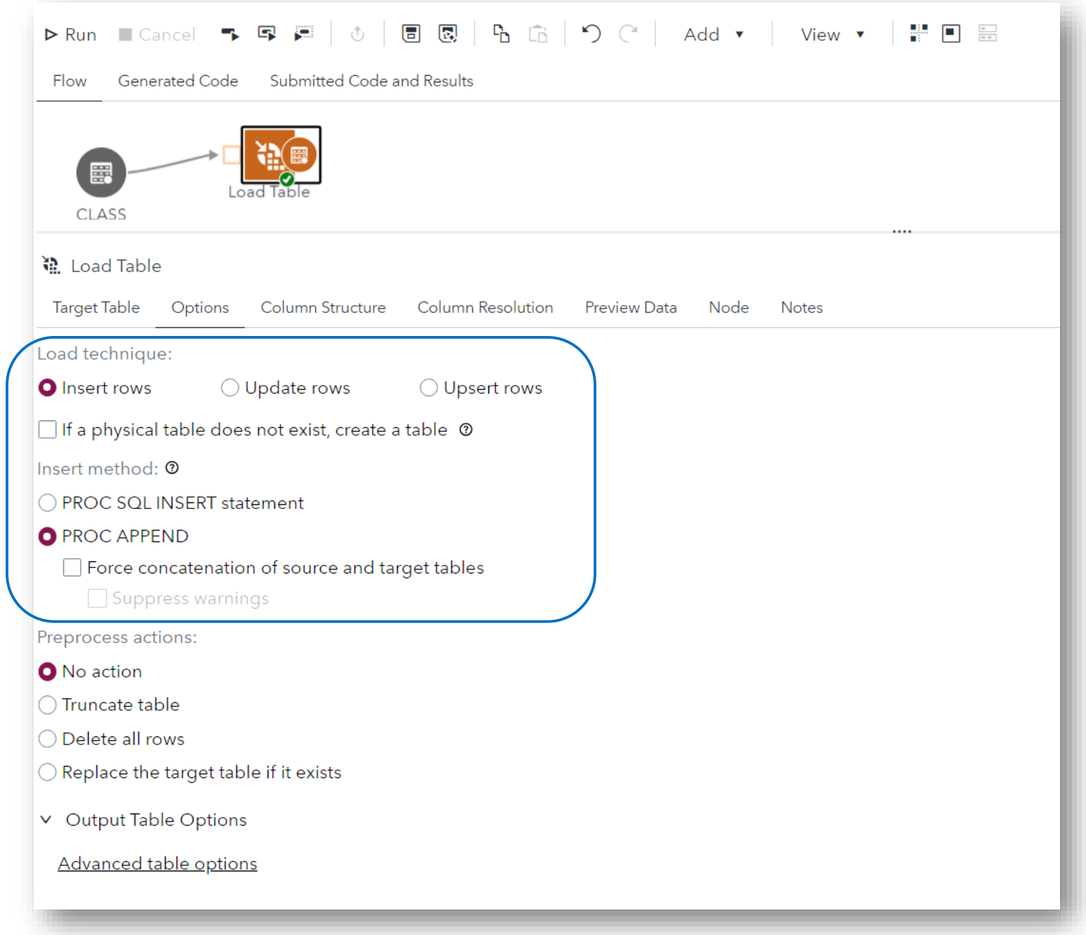
Informational
* The update planned for 2023.10 will allow to predict it with more accuracy

Legend *(percentage is included only if there is a "substantial" change)*
- Blue – update indicator (improvement in like-for-like migration)
- Green – CA run on Posten's job
- Purple – CA run on an extract of jobs that SAS has

2023.10

What's New

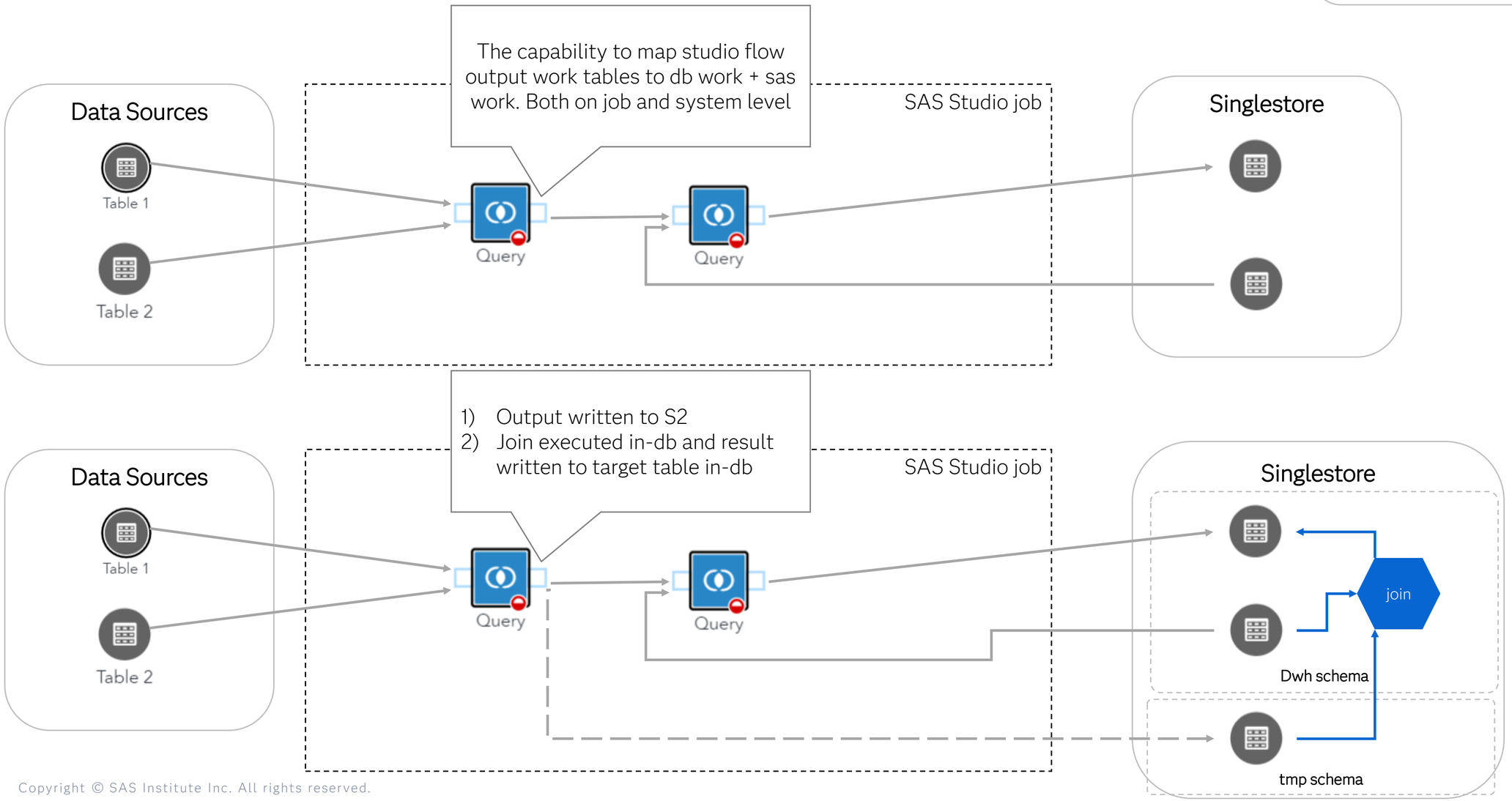
- DI Table Loader to Studio Load Table step migration – phase 2
  - PROC Append based features



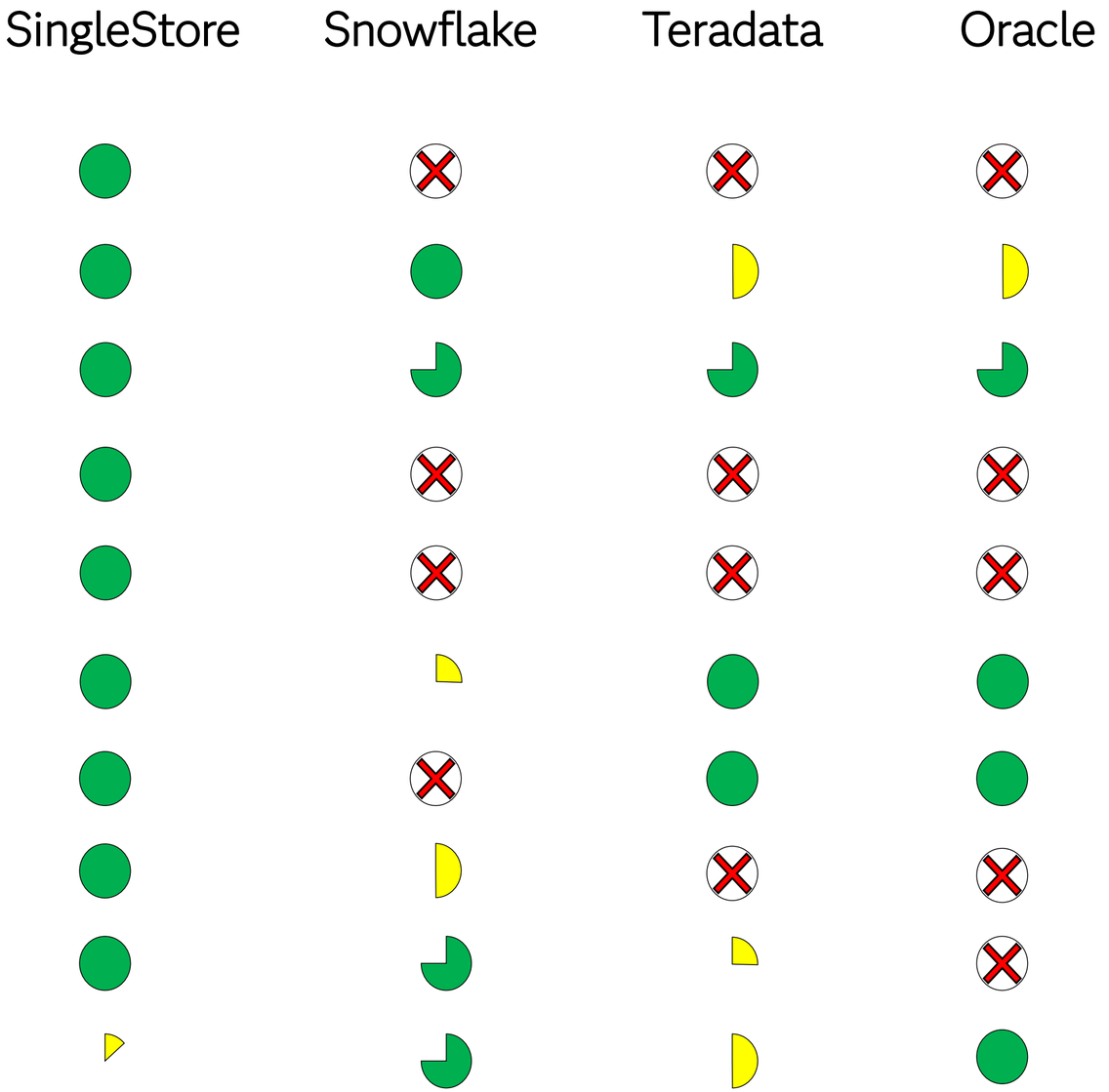
Using temporary db tables (2023.11)

Redirecting sas work or studio output library

Finding:
Pushing the join into S2 improve performance with a lot. In scenarios with big target table, we see 50-100 x



SAS Viya integration



Features and Benefits

SAS Viya with SingleStore

Performance

- Parallel, high-scale streaming data ingest
- Sub-second latency at high concurrency

Availability and Scalability

- Full HA with DR replication capabilities
- Scale to exabytes of data

Infrastructure Reduction

- SAS data sets compressed 80% +
- In memory processing with 60% reduction of RAM requirements
- De-couple and independently scale compute and storage
- Bottomless technology delivers 20-40% storage cost reduction

Open

- Eliminates SAS proprietary formats
- Bring your Open-Source analytics to SAS' commercial grade platform
- Standard SQL access

Deploy Anywhere

- Built for Kubernetes
- Deploys on-prem, in cloud, and hybrid

Security

- Data interrogation and protection tools – easily identify and protect sensitive data
- Enables encryption without loss of compression

Roadmap Priorities

Single Sign On AAD

- Build integration with SingleStore
- Enable Single Sign-On for Posten VA testing

Pushdown of aggregate functions to S2

- Enable pushdown of basic statistical functions
- N, NMISS, NDISTINCT, MIN, MAX, SUM, MEAN, MEDIAN, STD, VAR. simple.summary and simple.groupby
- Apollo pushdown of actions

Enable Views to work in VA

- Research approaches to enable computed columns, where, order by, aggregates in a view
- Determine how best to handle data combined by a view

EP support for datastep

- First support of datastep pushdown from CAS
- Secondly also from compute with the new universal EP design