SAP Analytics in a Cloud Native World

Don Loden, Protiviti

SAPinsider
Las Vegas

2023

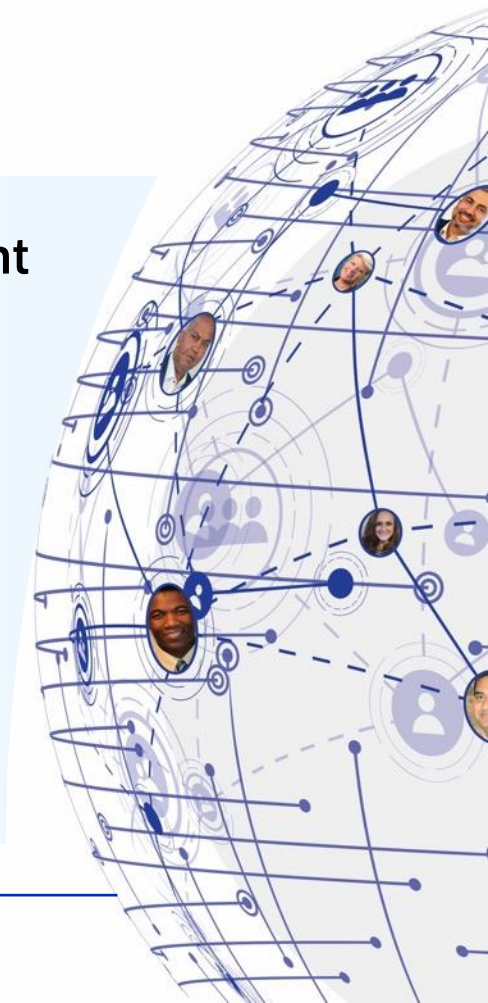
SAPinsider

In This Session

Deploying SAP data solutions in a modern cloud environment provides a fast, scalable and repeatable process for implementing full SAP stacks with unparalleled capabilities.

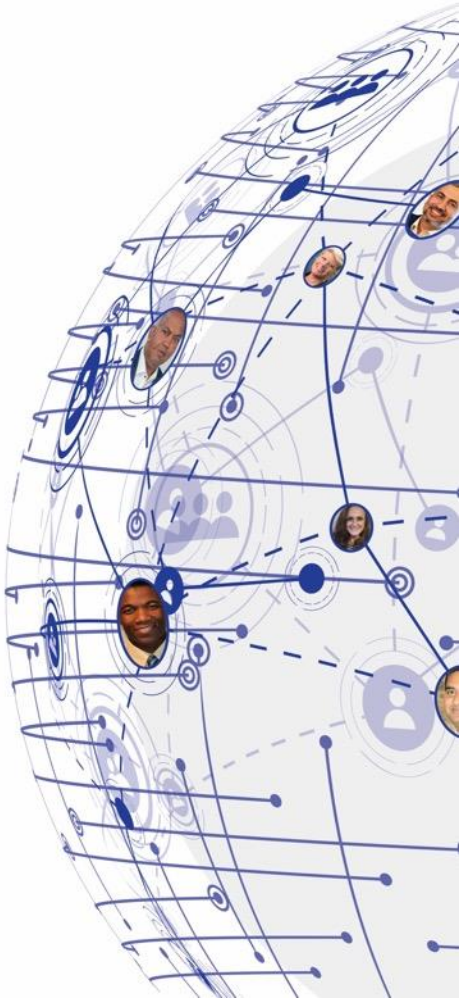
However, this topic can be confusing with many new terms that are foreign to a traditional SAP teams.

We will discuss some common terms to level set and take closer look from a modern SAP customer lens…



What We'll Cover

- Defining Common Terms
- What do Modern Cloud Architecture Look Like for SAP Customers?
- Leave Data at Rest? Why?
- Wrap-Up



Going to the cloud can be confusing...

Let's define some common terms

You may hear terms like…

Hyperscaler. What is a hyperscaler, and what should you look for in a hyperscaler for SAP needs?

A cloud provider that involves the ability to seamlessly provide and add compute, memory, networking and storage resources. ***Look for INTEGRATION with your existing technologies!***

- Common hyperscalers are Amazon, Google, and Microsoft in the SAP ecosystem
  - Many SAP customers have one or many of these in their landscapes already. Multi-cloud is already common and will become more common in the future
  - It is not just about price… each have their respective niche for example:
    - Microsoft: Many ready to use low/no code services but may not be as configurable
    - Amazon: Great for developing applications as they offer so many options and configurations
    - Google: Recently offered Cortex Data Foundation for SAP

You may hear terms like… (Cont.)

Public or private cloud? Both options have pros and cons. Which is right for me?

*In simple terms: **Private** is controlled by a single organization; **Public** is a subscription service that is offered to any and all customers who desire the same services*

- Some analogies:
  - Private cloud is like a house where Public cloud is like an apartment building
  - Private cloud is like watching movies you own on media (DVD) Public cloud is like a streaming video service
- If security is of utmost concern, then Private cloud is a good option, but it is typically more expensive

You may hear terms like... (Cont.)

You have BI or a data warehouse(s). Why should you be thinking about data fabric?

- You have spent lots of time and resources on data warehouses and even data lakes... *Data Fabric is a collection of tools and practices to 'knit' the data topologies together*
  - Some means are achieved by virtualization others by accelerators or API layers that ease combination of disparate data
  - Data warehouses (EDW) are connected databases that pull, transform, and store harmonized data. Data Fabric receives data from the EDW.

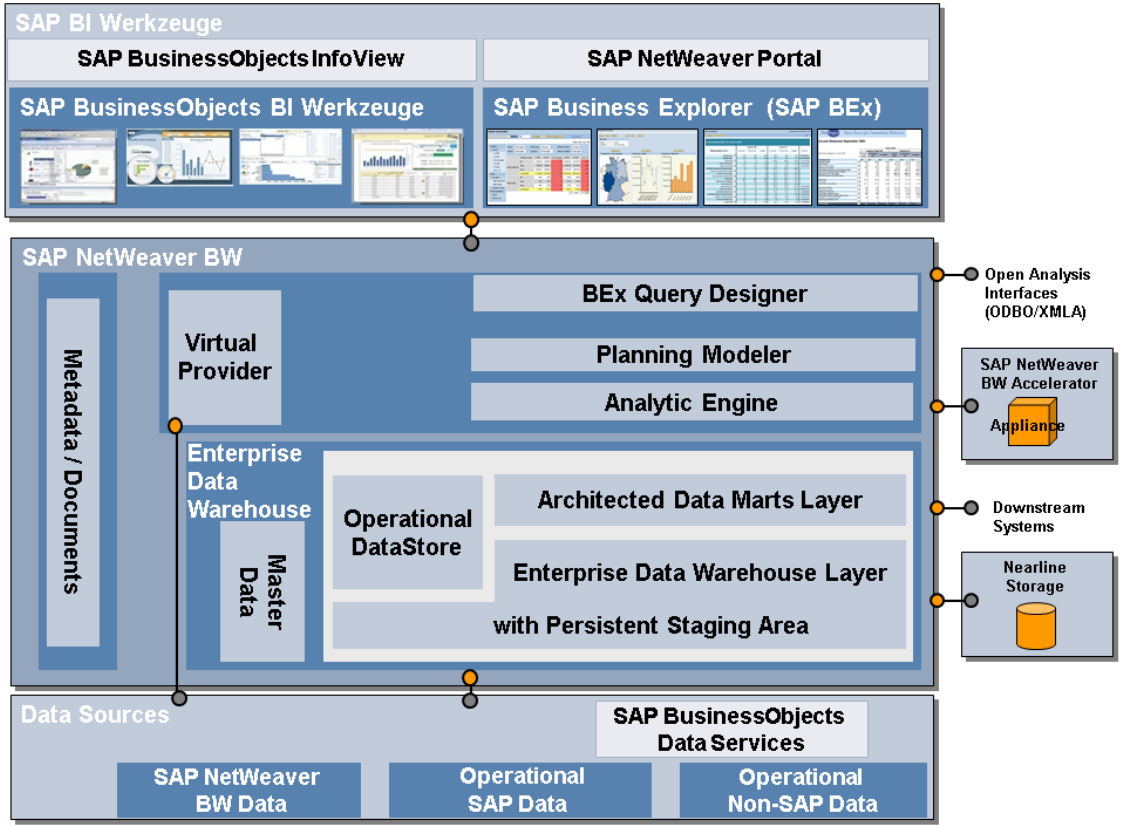
SAP HANA Brought DBA Topics to Basis… Cloud Brings Multi-Platform

- Before SAP HANA a DBA just maintained the database that supported SAP but was very ‘Abstracted’ from the application layer…
  - Abstraction was by design for client-server deployments (R3/ECC)
  - DBA topics where more backup and recovery
  - Maintenance and patching
- After SAP HANA the database is an integral part of the ERP in S4
  - The ERP function can exploit means and performance that only HANA can offer
  - This comes at a price: DBAs must understand the HANA DB componentry at much deeper level past the application layer.
- Cloud is much the same: Now you must think it native application data needs across other platforms when data is shared with SAP – *Many new things to learn!*

What does a Modern SAP Data Landscape Look Like ... taking data fabric concepts into account?

Where Did We Come From… Legacy SAP Landscapes

- Typical BW processing layer
  - *Complex but directional most data from SAP and some from third party systems*



Courtesy of SAP

Why are we Here?
For Innovation and the Art of the Possible

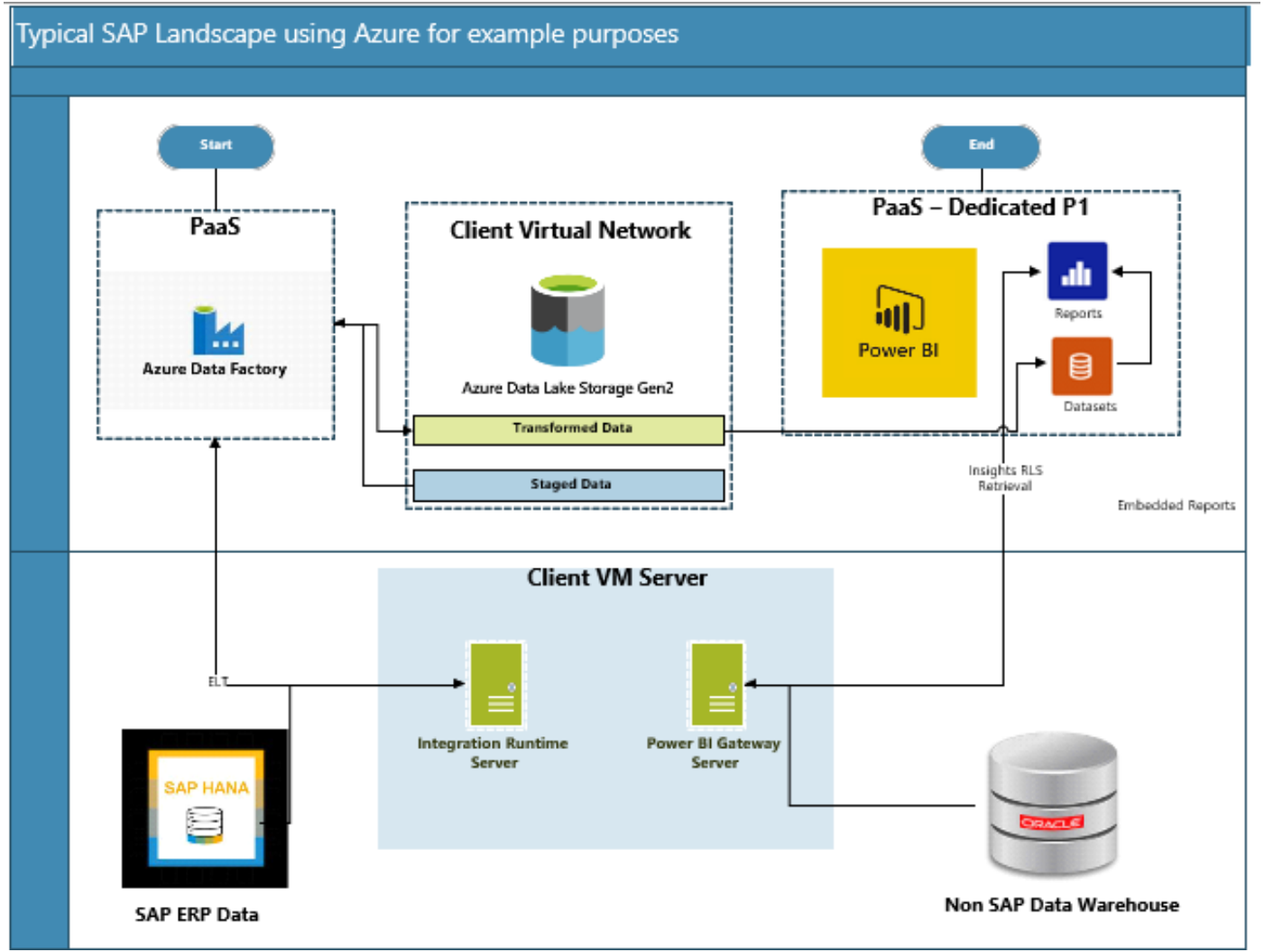
Cloud computing fosters the ability to access information in almost limitless size and complexity, empowering organizations to fail fast so they can innovate quickly…

Enabling rapid business innovation

Try on possible ideas quickly allowing acceptable failures

Data fabric ideals support SAP/non-SAP combination keeping data mostly at rest

Why do I care? *Because you likely have a ton of data that does NOT exist in SAP!*



Microsoft Azure *was used for this real world client example, but* ***Amazon*** *and* ***Google*** *both have similar options and features if you are already using those providers.*

SAP *also has many compelling cloud subscription offerings, and these would simply snap into this architecture via integration or services. Our clients* ***typically blend SAP cloud options with common hyperscalers***

Leave Data at Rest?

Define “Data at Rest”…

- Data at rest means data that has reached a destination or is at an origination point and is not being accessed or used
  - Data is constantly being created, accessed and maintained in more places than ever in modern organizations
    - Cloud developed apps produce important data often content directly from end customers or consumers: *Example: Consumer apps that can track location of how, where, and when a company’s products are used*
    - SAP systems control and maintain the backbone of data that runs the enterprise: *Example: Making, distributing, and accounting on products for customers*
  - *There is a Great Deal of value in gaining insights from the constant combination of this data!*

Why Leave at Rest?

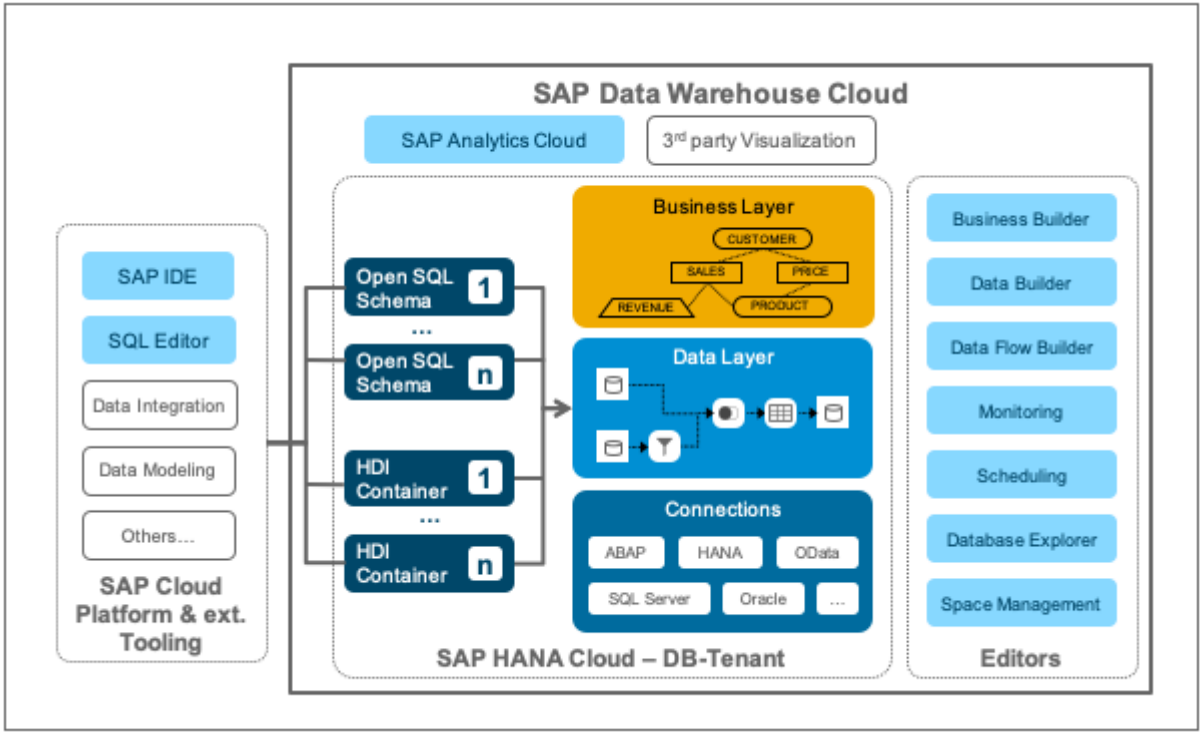
- Why? … Because Data Egress is Expensive!
  - Cloud providers will happily take as much data as you want to produce for free.
    - *The cost is when that data leaves their cloud and moves to other platforms it gets expensive*
    - This is not a big deal when there are small amounts of data moving around but *do consider this topic if you are crowd sourcing lots of information from consumers in a cloud native app and you need to combine this data with SAP data!*
- *The real answer is to Know who You are as an organization!*

Who are you as an Organization: SAP Centric or Cloud Native first?

- SAP customers may think of themselves as SAP centric, but that may not be the case
  - Think of where your data comes from and is maintained…
    - How much of your organization is really run by and SAP system?
    - Do you get much of your insights from crowd sourcing… experience software… direct customer interaction?
    - Many SAP customers are seeing that SAP has much to offer and many compelling products, but even SAP's vast ecosystem is just a side of an everchanging story
  - It is a more simple case if all of your systems run on SAP.
    - We see some clients like that but those are the rare examples.

SAP Data Warehouse Cloud Example

- Rich Business Data Layer
  - *Takes advantage of all of the metadata and knowledge of SAP data from native means*



Strong Integration Capabilities

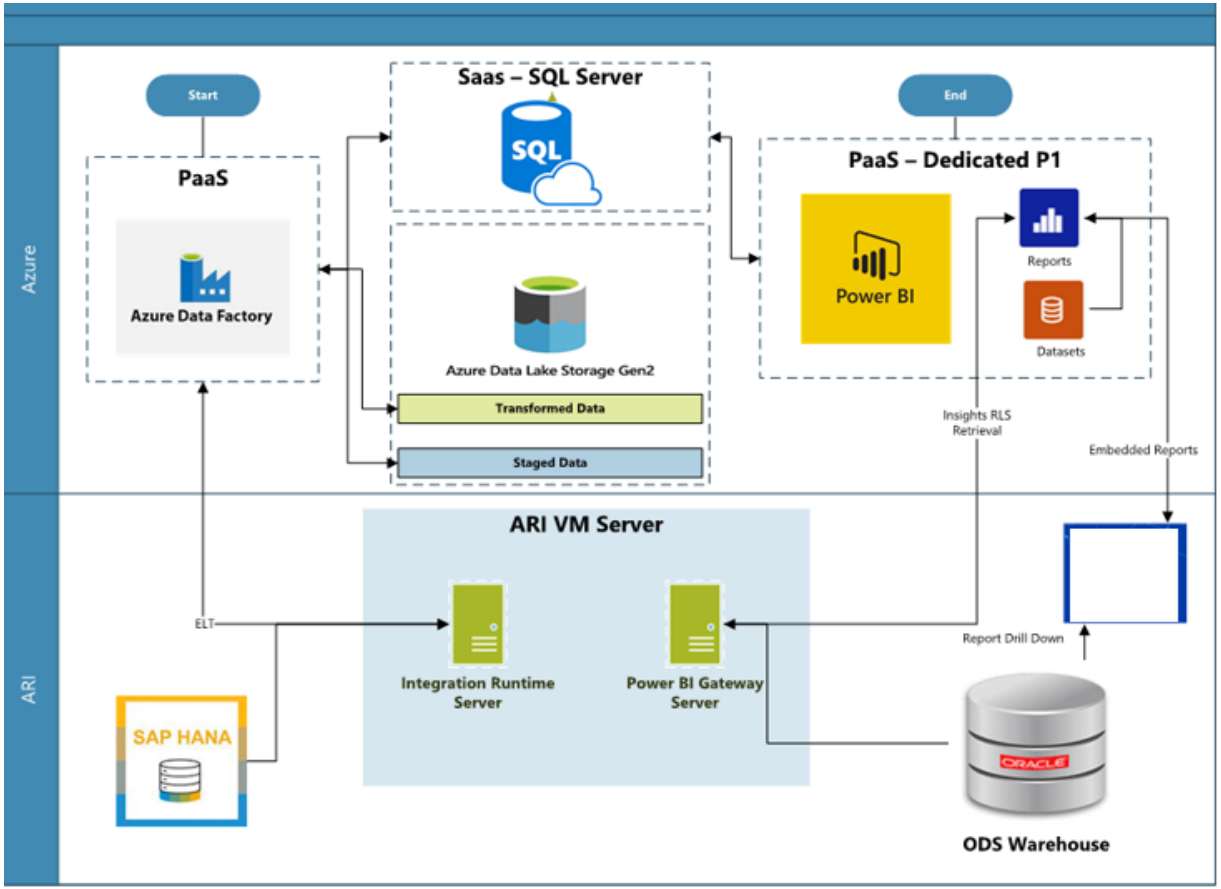
Utilizes built in APIs and connections natively

No need to build integrations

Courtesy of SAP

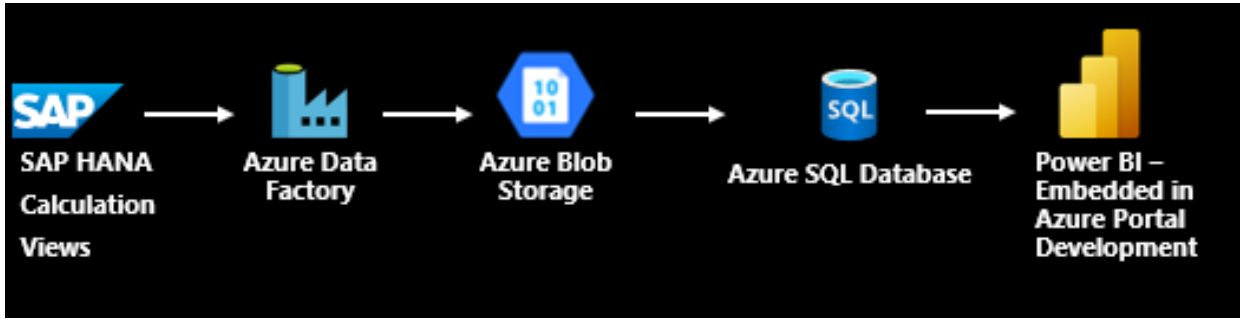
SAP Customer with Cloud Native Data Processing

- Flexibility of cloud data with native cloud tools and services
  - *Takes advantage of all of the metadata and knowledge of SAP data from native means*



SAP Customer with Cloud Native Data Processing (Cont.)

- Microsoft Azure tools deployed from another cloud native data processing example
  - Data will be copied over but then can be integrated with multiple cloud native tools
  - AWS or GCP would be similar



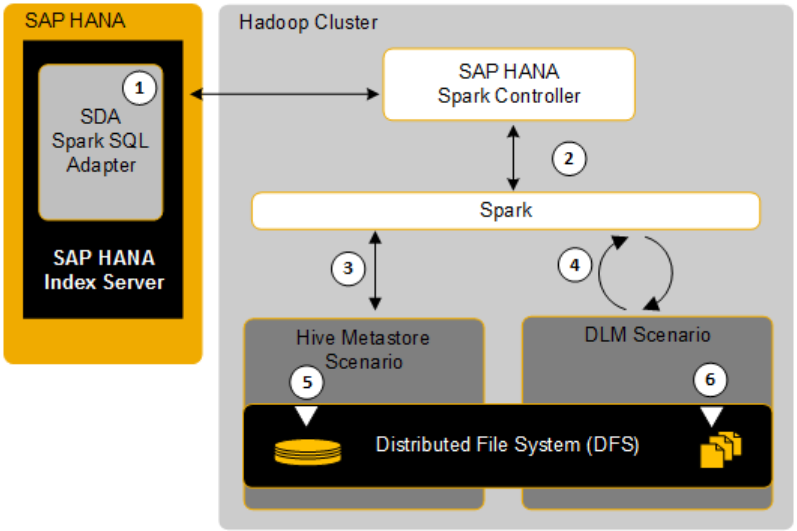
***Microsoft Azure** was used for this real world client example, but **Amazon** and **Google** both have similar options and features if you are already using those providers.*

SAP** also has many compelling cloud subscription offerings, and these would simply snap into this architecture via integration or services. Our clients **typically blend SAP cloud options with common hyperscalers

SAP Customer with SAP Centric Data Processing

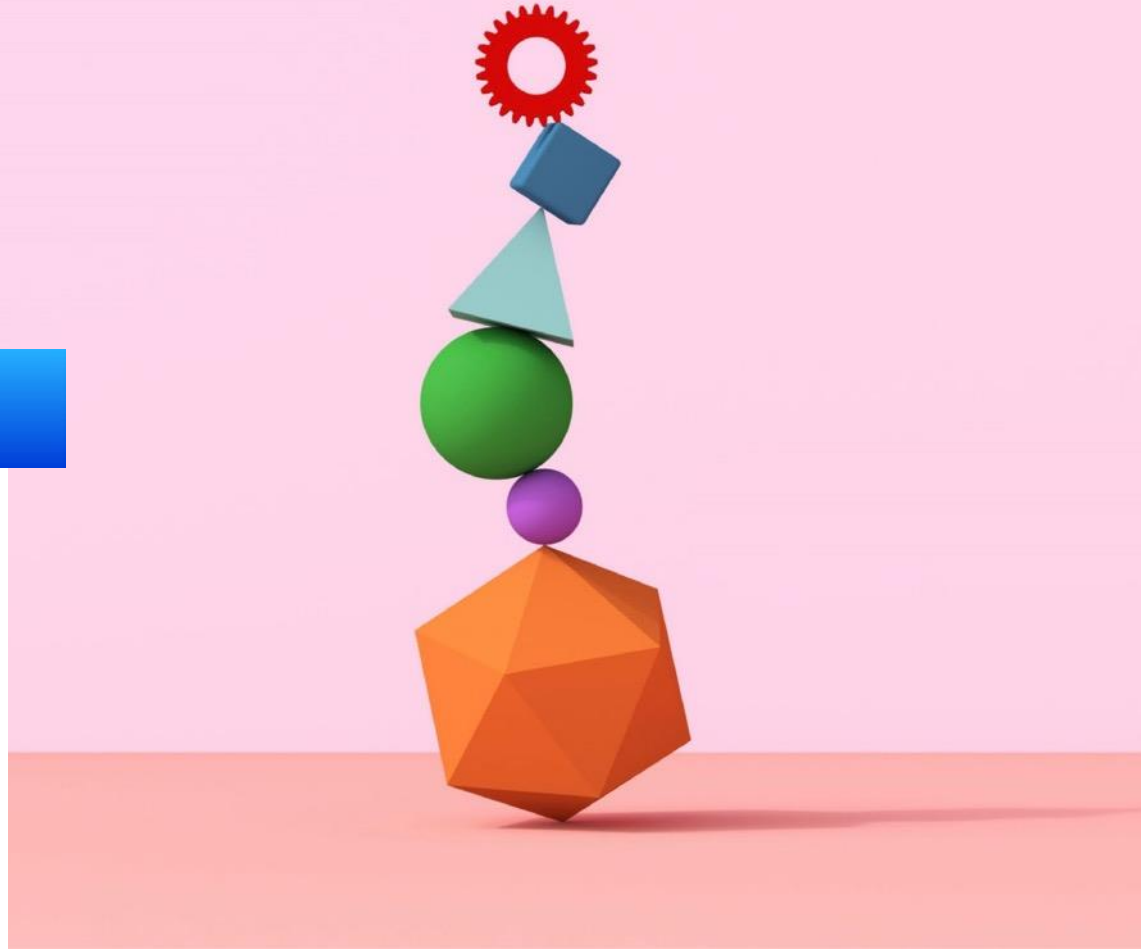
The SAP HANA Spark Controller is an application that sits on top of Spark that allows Spark to expose its development artifacts to HANA through remote source connections. HANA connects to the Spark system using an SDA connection with a native adapter called 'sparksql'. Once the connection has been established, the data can be exposed as virtual tables, to be used in development. There is also a DLM component that can be used to flag and archive cold data in the Hadoop system.

- **The Spark Controller:**
  - **Is assembled, installed, and configured on a Hadoop cluster.**
  - **Facilitates data transfer between SAP HANA and Hadoop.**
  - **Facilitates query execution and enables SAP HANA to fetch data in a compressed columnar format.**
  - **Supports SAP HANA-specific query optimizations and secure communication.**



Courtesy of SAP

Wrap Up



Where to find more information

- Balancing Risk with Innovation in the Cloud
  https://sapblog.protiviti.com/2022/03/16/a-balancing-act-sap-in-the-cloud/
- Balancing Innovation and Security in Cloud Deployments
  https://sapinsider.org/videos/balancing-innovation-and-security-in-cloud-deployments/
- Video – SAP Workloads in the Cloud: Opportunities and Risks
  https://sapinsider.org/videos/video-sap-workloads-in-the-cloud-opportunities-and-risks/
- Managing Risk Along Your SAP S/4HANA Journey
  https://www.protiviti.com/US-en/insights/internal-audit-role-sap-hana

Key Points to Take Home

- Cloud is **not a one size fits all**. Know who you are as an SAP customer.
- Know **where you produce most of your data** to avoid egress charges!
- Don’t just consider price when looking at a hyperscaler. Free or cheap storage or services can often **become the most expensive solution**
- SAP has lots of great options. Don’t lose sight of the fact that integration **is really expensive work**. SAP handles much of that for you in their cloud apps/platform
- Cloud services are here to stay: **Why build something when you can buy it?**

Thank you! Any Questions?

Don Loden
Don.Loden@protiviti.com
@donloden

Please remember to complete your session evaluation.

SAPinsider

SAPinsider.org

PO Box 982Hampstead, NH 03841

SAPinsider comprises the largest and fastest growing SAP membership group worldwide, with more than 600,000 members across 205 countries.