

SmartData Fabric[®] (SDF) Hybrid Cloud 2.0

January 2022



Pros and cons of cloud(s)

- Great for on-demand, scale-up and scale-down, high performance computing
- Good for central management of both cloud and conventional hybrid cloud apps
- Expensive for data storage, I/O and other charges
- Typically, a lot of data copied:
 - a. EXTERNAL TO A CLOUD from on-premise, data centers, other clouds and third-parties to a single cloud for processing, some BI and analytics, e.g., a data lake, and latency introduced unless a (near) real-time data lake (unusual)
 - b. INTERNAL WITHIN A CLOUD, e.g., from a data lake, ETL to a data warehouse, on which queries used to populate data marts multiple copies made and increased latency
- Most enterprise operations still run off-cloud, on-premise, in data centers, SaaS, etc.
- Cloud-based data lakes and data warehouses tend to require ALL the data, but not all data can be copied or moved to a cloud or even a central location
- Difficult to feed back cloud-based analytics results to off-cloud-based enterprise operations, especially in (near) real-time as customers, employees, patients, etc. are interacting

What about conventional hybrid cloud (1.0)?

- Deploy compute remotely where data resides, e.g., on-premise, in data centers, on other clouds and (maybe) on SaaS and third-parties
- Centrally manage remote compute through a cloud-based system, e.g., hypervisors and VMs, Kubernetes (and variants) and containers
- Provision resources on remote systems to run remote compute/app virtualization
- Adopt conventional data virtualization
 - Addresses IT issues (to an extent) of access control (although data issues can impact)
 - Does NOT address remote compute provisioning and management, or data source and data issues
- Could adopt WhamTech *unconventional* data virtualization
 - DOES address data source and data issues in addition to IT issues of access control
 - Does NOT address remote compute provisioning and management (to an extent) issues



What is Hybrid Cloud 2.0?

- 1. Data is everywhere leave it where it is: On-premise, on mainframes, in data centers, cloud(s), SaaS, third-parties, Web, social media, etc.
- 2. Avoid deploying remote compute near or on data source system platforms, as is the case with Hybrid Cloud 1.0
- 3. Deploy SmartData Fabric® unconventional data virtualization 100% IN THE CLOUD, leveraging index-based, in addition to conventional, federated adapters for data-related pre-processing and query processing in and/or from the cloud no need to install and run anything elsewhere, as is the case with Hybrid Cloud 1.0
 - Establish index update process through changed data capture (CDC)
 - Multiple CDC options, including near real-time (NRT) updates
- **4.** Focus on data that needs processing for quality, standardization, security, relationship mapping and master data management (MDM) various options for the rest of the data
 - Enable data management fundamentals
 - Address data, data source and access control issues
- 5. Multiple configuration options, including (a) some data indexed and the rest stays in the source, (b) all data indexed and stored in indexes, and (c) no data indexed and all queries on data source, with other options in-between
- 6. Avoid incomplete or incorrect query results, query load and/or poor query performance of conventional data virtualization/federation, i.e., avoid dependence on data sources, data in sources or data source own access control
- 7. Immediate short-to-medium-term implementation
- 8. Facilitates medium-to-longer-term transition-migration to the cloud



SmartData Fabric[®] Hybrid Cloud 2.0 configuration and deployment options



SmartData Fabric[®] capabilities address issues





Smartphone CRM app invoking new BPM-based workflows with writeback to legacy systems through standard APIs/data services

- Patient-centric smartphone app interacts with legacy data sources through new workflows developed and orchestrated by BPM software
- BPM workflows interact with data source through standard FHIR REST APIs provided as data services
- BPM workflows both read and write back to legacy data sources





The End