



Db2 11.5.9 & Next Generation Warehousing

Les King

lking@ca.ibm.com

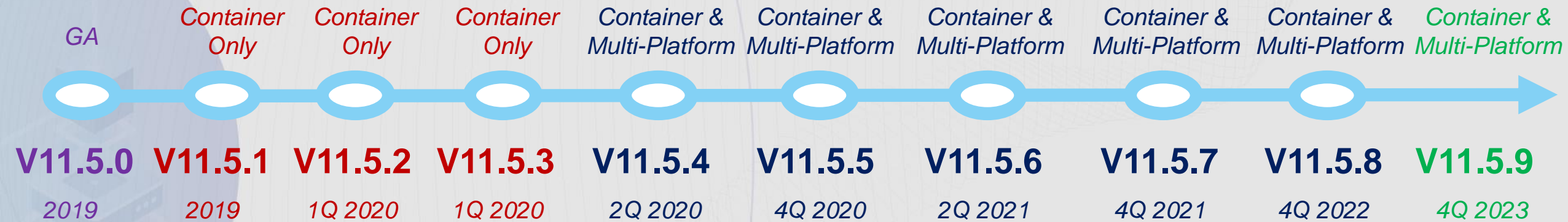
September 2023

Data Server Day – Stockholm, Sweden

Agenda

- Db2 11.5.8
- Db2 11.5.9
- Db2 11.5 Versioning
- Db2 Warehouse SaaS Gen 3
- Cloud Rack for Db2

Db2 V11.5 - Lifecycle



Regular fix/mod pack deliveries

- With overriding focus on stability for immediate production deployment
- Modifications will contain select functional enhancements (often off by default)

Db2 11.5.8 Highlights



Hybrid & Multi-cloud

Containerization:

- RHOS on AWS & Azure - ROSA and ARO
- Deploy on Amazon EKS
- Deploy on Azure AKS
- Integration with Velero for BAR
- **Note: most of above not tied to 11.5.8 but on similar 2H 2022 timeline (see roadmap for details)**
- **Operator driven Db2 native backup and restore**
- **Db2 W Scaling - vertical by adjusting resources per pod**
- **Db2 W Scaling - horizontal by scaling out number of pods**
- **HADR role-aware Kubernetes service for Db2U. Route clients to primary.**



Application Development

- Multi-row Fetch - new regvar to force the count to match Z behavior
- Federation Enhancements:
 - Perf improvements in: FFNR, Pushdown for Impala JDBC, CSE
 - MongoDB – SSL, authentication and multi-host
 - SSL Support for HANA w/ ODBC
 - Add MySQL8
- Client Enhancements:
 - JCC improved connection pool
 - db2dsdriver validation tool
 - Large result set query perf
 - .NET Visual Studio 2022 support
 - OpenJDK 17 runtime support



Columnar Enhancements

- **LOAD enhancements**
 - **Performance Improvements**
 - **Applies page level varchar compression techniques (if enabled)**
- **Logical Columnar Backup (aka COISBAR/Schema Backup):**
 - **Now available on linuxAMD/ppcle across db2 installs (containers and on-prem)**



Cognitive and Emerging

Db2 Graph, new tutorials plus:

- GUI Enhancements to Query Screen
- **Discovery mode to explore w/o writing gremlin queries**
- **History to revert and compare graphs models**
- **Autogenerate code to run queries**

RED = Shipped in 1Q 2023

Db2 11.5.8 Highlights



Availability

- **db2adutl to upload load copy files to TSM**
- LogReadAPI to map timestamp to LSN for (for logical rep)
- Log replay improvements (HADR/rollforward/crash recovery), and documentation for support recommended tuning parms
- GUI and silent installs now can install Pacemaker
- Improved deployment experience for Db2 HADR on AWS, Azure, GCP (automation and improved documentation)
- Pacemaker support for mutual failover on single node (aka single node active/passive w/ shared storage)



pureScale

- pureScale TCPIP on AWS – deployment and support (already available).
- **RHEL 8.6 support w/ RDMA on linuxamd+ppcle (via verbs)**
- **Simpler concurrent pureScale online modpack& fixpack update**
- DIRECT_IO as default in Z-linux
- Periodic collection of RDMA roundtrip network stats
- Avoid lock timeout due to EHL
- pureScale on z15 support (not be tied to modpack, will be later)

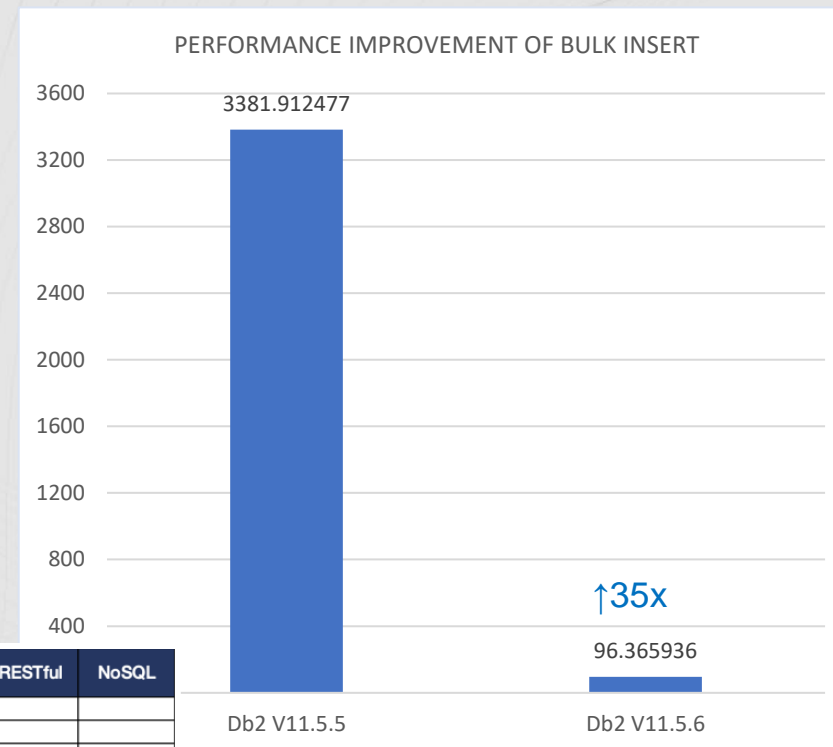


Core Engine

- Additional logging for XML data (optional) when changed data capture is used.. Can be used by logical replication to replay XML w/o calling back to server.
- New ADMIN_FMP_TERM SP to shutdown an fmp process
- External table support to read from “binary numeric” format
- Extend db2RemStgManager to use storage aliases
- **Support for TLS 1.3**
- **Force Load of GSKIT which comes with db2**
- **Monitor improvements: SSL info to be added to MON_GET_CONNECTION**
- Catalog And Process Model Serviceability improvements
- db2support to automatically gather text search logs
- GUI and silent installs now can install Pacemaker
- Java, TSA, GPFS and Pacemaker updates
- Windows 2022

Db2 V11.5.8 – Federation Enhancements

- Connectivity - Spark JDBC Connectivity Support
- Functionality - Column Length Auto-Extension for Code Page Conversion
- Functionality - Nickname Hidden Column Support
- Performance - Federation DRDA Bulk Insert for Db2 Family Data Sources
- Security - SSL Encryption Support for ODBC SAP HANA - Latest encryption capabilities for SAP HANA when connecting through ODBC



Category	Data Source	Native	ODBC	JDBC	RESTful	NoSQL
Relational	Db2 LUW	11.5		11.5		
	Db2 for IBM i	7.4				
	Oracle	19c	19c	19c		
	MS SQL Server	2019	2019	2019		
	Informix	12.1				
	Sybase	16.0				
Warehouse / Appliance	IIAS	-				
	Netezza		7.x	7.x		
	Teradata	16.0		16.0		
	SAP HANA		2.0	2.0		
	Greenplum		5.0	5.0		
Open Source	MySQL Community		8.0	8.0		
	MySQL Enterprise		8.0	8.0		
	PostgreSQL		12.x	11.x		
	MariaDB		10.x	10.x		
	Derby			10.8		
Hadoop	IBM Db2 BigSQL	7.1		7.1		
	Hive		5.13 & 3.1	3.1		
	Spark		2.1.1	2.1.1		
	Impala		2.7	2.7		
Files	Delimited	-				
	Excel	-				
	XML	1.1				
	JSON					-
Mainframe	CSV	-				
	Db2 for z/OS	12.x		12.x		
	IBM DVM for z/OS			1.1		

Category	Data Source	Native	ODBC	JDBC	RESTful	NoSQL
Message Queue	MQSeries	8.x				
Cloud	Db2 Warehouse	-		-		
	MS Azure SQL		-			
	Oracle Cloud		-			
	Amazon AWS Redshift			-		
	Google BigQuery			-		
	Amazon AWS Athena			-		
	Salesforce			-		
	Snowflake			-		
	MariaDB Xpand		Working			
	AWS Aurora PostgreSQL		Working			
NoSQL	Hyperledger Fabric					-
	MongoDB					3.4
	CouchDB					2.2.0
	Hbase HDFS					2.9
	Cassandra					
	Redis					
	Jira				-	
	Aha!				-	
	Git-Hub				-	
	HubSpot				-	
	TeamCity				-	
	api.spacepdata.com				-	
	earthquake.usgs.gov				-	
	Google Calendar API				-	
	groupkt.com				-	
	Yelp				-	

Supported Before v10.5
Supported In v11.1
Supported In v11.5 GA
Supported in v11.5.4
Supported in v11.5.5
Supported in v11.5.6
Supported in v11.5.8
Working / Planning

Db2 11.5.8 – Replication Enhancements

- **Multi-Model – Logical XML Replication** - For clients who are using our native XML capabilities, we will provide logging so that XML can be replayed without needing to get latest data from server

2023 & Db2 11.5.9 Calendar

Q2 2023

- Db2 Warehouse Cloud Rack for P10

Q3 2023

- Db2 Warehouse Gen3 on AWS w/ **Db2 11.5.9**
- **Db2 11.5.9** container software technical preview (Db2 W only)
- Db2 Warehouse Cloud Rack for x86

Q4 2023

- **Db2 11.5.9** full software release (aka 11.5.9 mod pack)

Db2 11.5.9 Highlights



Hybrid & Multi-cloud

Containerization:

- Operator driven Db2 native backup and restore
- Db2 W Scaling - vertical by adjusting resources per pod
- Db2 W Scaling - horizontal by scaling out number of pods
- HADR role-aware Kubernetes service for Db2U. Route clients to primary.
- Containerization support for Native Object Store
- Containerization support for ODF



Application Development

- Federation Enhancements
 - Additional push down for:
 - CTEs
 - numerics
 - datetime
- Client Enhancements:
 - Uplift of levels



Columnar Enhancements & Db2 Warehouse

- Support for tablespace storage to be on S3 (aka Remote Table spaces)
- Support for Open Data Formats: Iceberg, Parquet, AVRO, ORC
- Reference Architectures for Db2 Wh on Power and x86



Cognitive and Emerging

Db2 Graph:

- Discovery mode to explore w/o writing gremlin queries
- History to revert and compare graphs models
- Autogenerate code to run queries

* Delivered (1Q 2023) in latest containers

Db2 11.5.9 Highlights



Availability

HADR support for pS on Z



pureScale

- additional adapter support (Mellanox CX-6 RoCE)
- Z 16 support
- **pS on AWS Enhancements:**
 - Single click to deploy TWO pureScale in different AZ in same/different Region at the same time for HADR
 - More flexibility in specifying the number of file systems



Core Engine

- RHEL 9 support - still working out any limitations
- Amazon Linux 2 support - likely base only
- Stack uplifts
- Stored procedure to allow generation of formatted explain plans



Security

- Support for Audit logs to COS
- Advanced Data Masking

* 11.5.8 support announce in 2Q

LAKEHOUSE Table – Market Dynamics

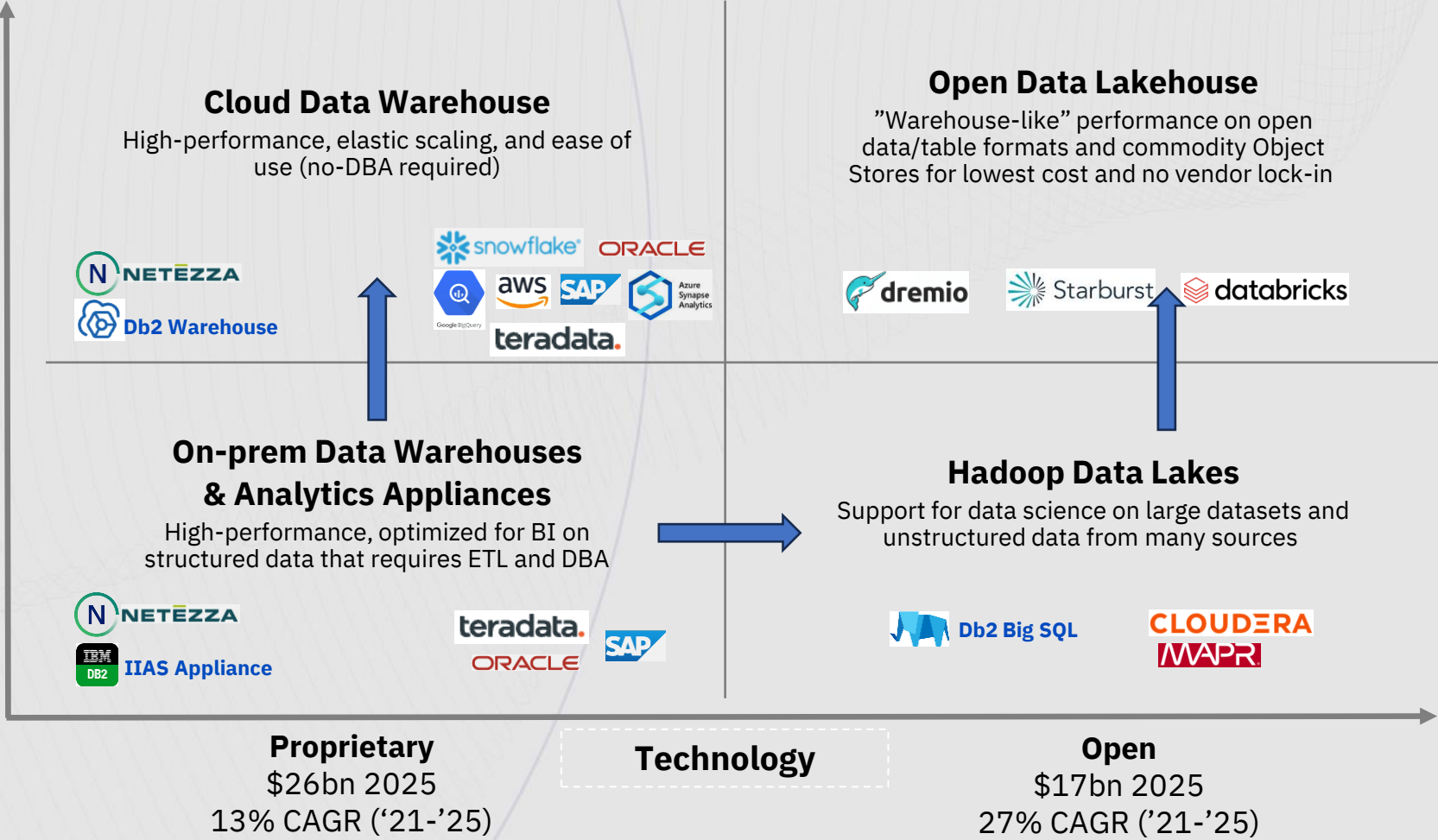
Major disruptions are driving the growth in the analytics repositories market **from on-prem to SaaS** and **from proprietary to open technologies**

Analytics Repositories Market Landscape

SaaS
\$31bn 2025
27% CAGR ('21-'25)

Deployment

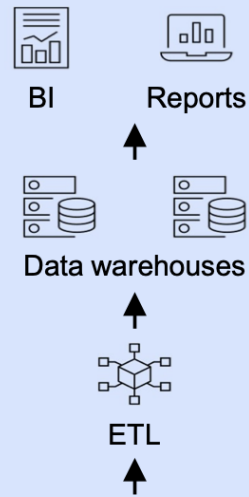
On-prem
\$12bn 2025
2% CAGR ('21-'25)



Sources: IDC Data Management Forecast (November 2021), IDC BDA Forecast (June 2021), MI modeling

LAKEHOUSE Table – Market at a Crossroads

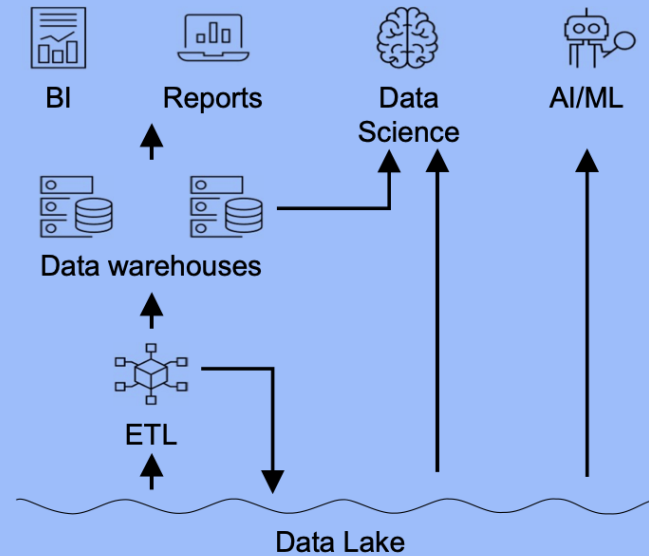
Data Warehouse



Warehouses have become too rigid and expensive for modern workloads

- Too **expensive** to store large datasets
- **Inflexible** due to their tight integration with storage
- **Vendor lock-in** via proprietary data formats

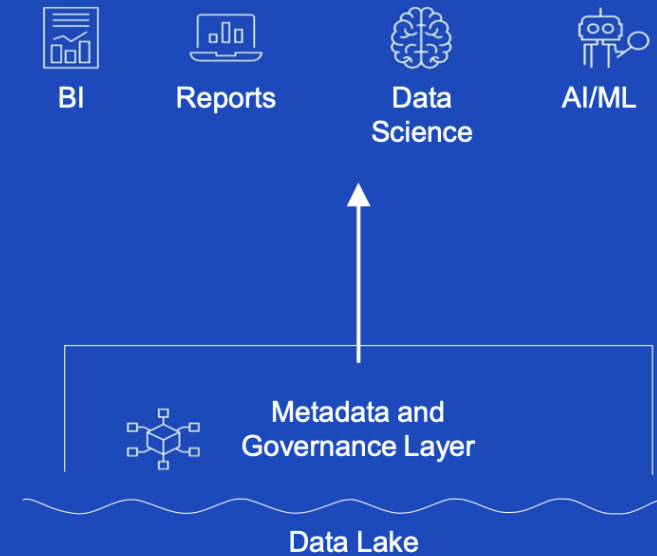
Data Lake



Data Lakes are too complex to implement

- “**Less than 20%** of Hadoop data lake implementations are **successful**” Gartner
- **Feature poor** – ACID, Time travel

Data Lakehouse



The 2 never worked well together

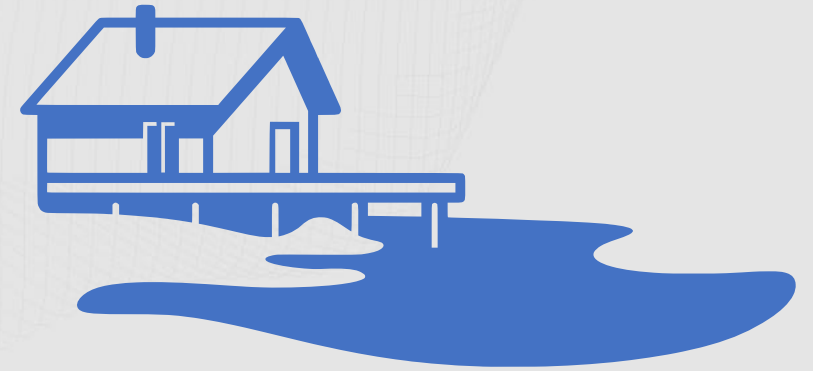
- **Too much cost / effort / risk to have both**

A new approach is needed

LAKEHOUSE Table - The Data Lakehouse

The Data Lakehouse implements the **data structures and management features** of a data warehouses on the **low cost, reliable & scalable** object storage within a new architectural approach that leverages open-source technology.

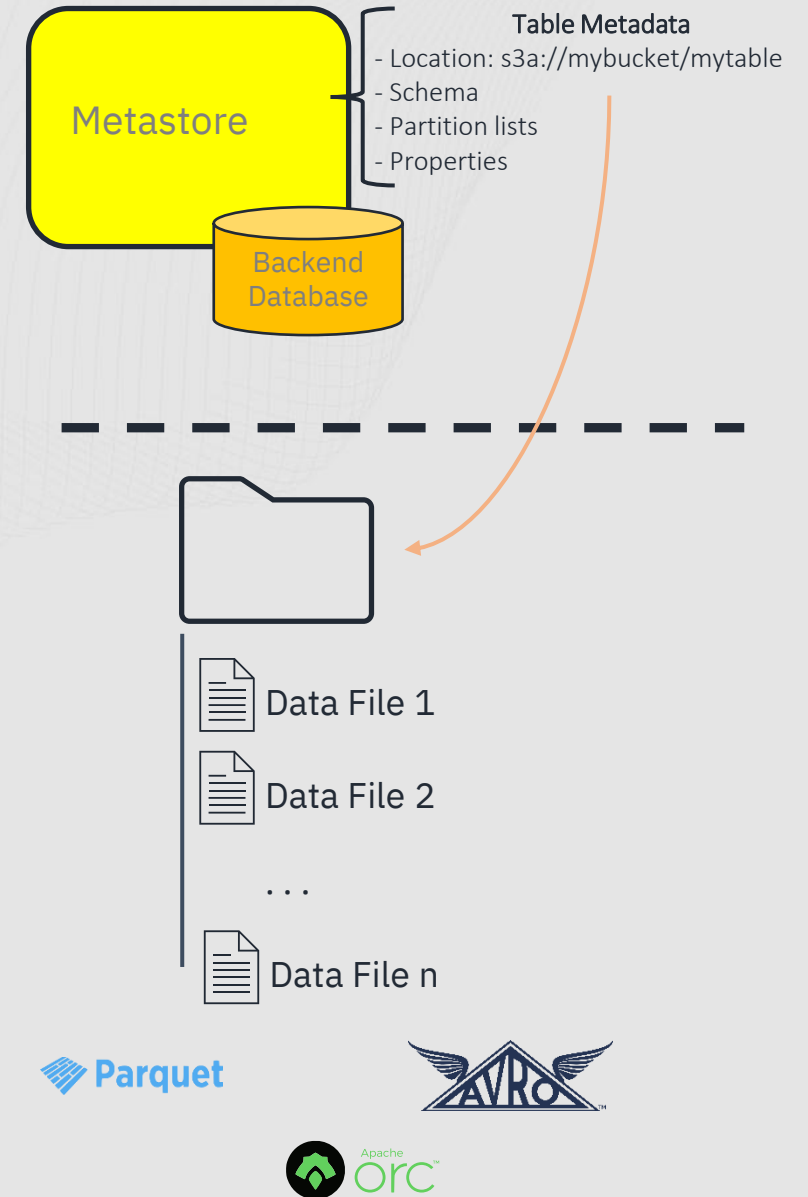
It enables organizations to manage their data in an **open, flexible, cost-effective, feature rich and scalable way**, enabling Business Intelligence and Machine Learning on all data.



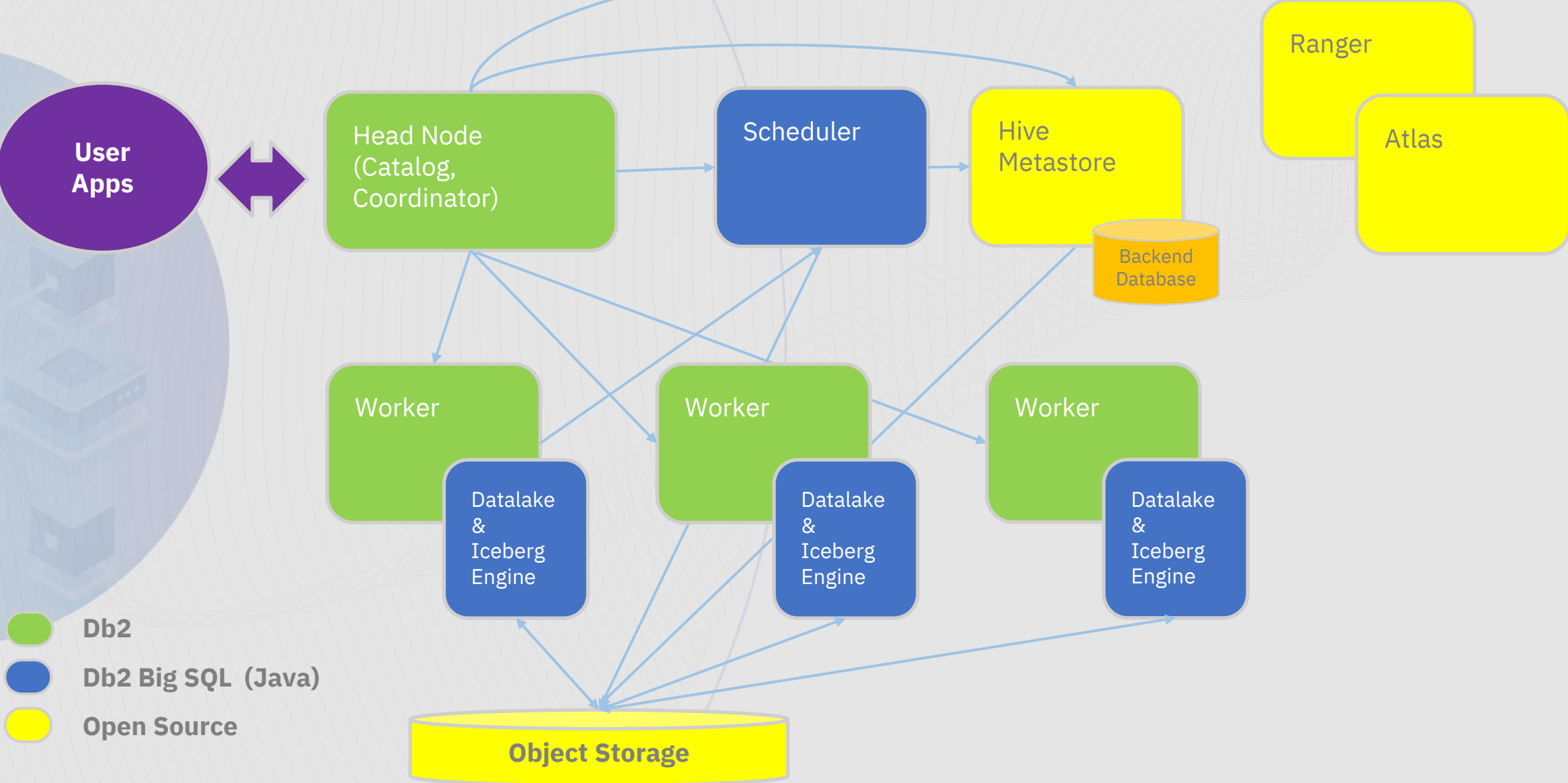
Data *Lake*
+
Ware*house*

LAKEHOUSE Table - Datalake Tables

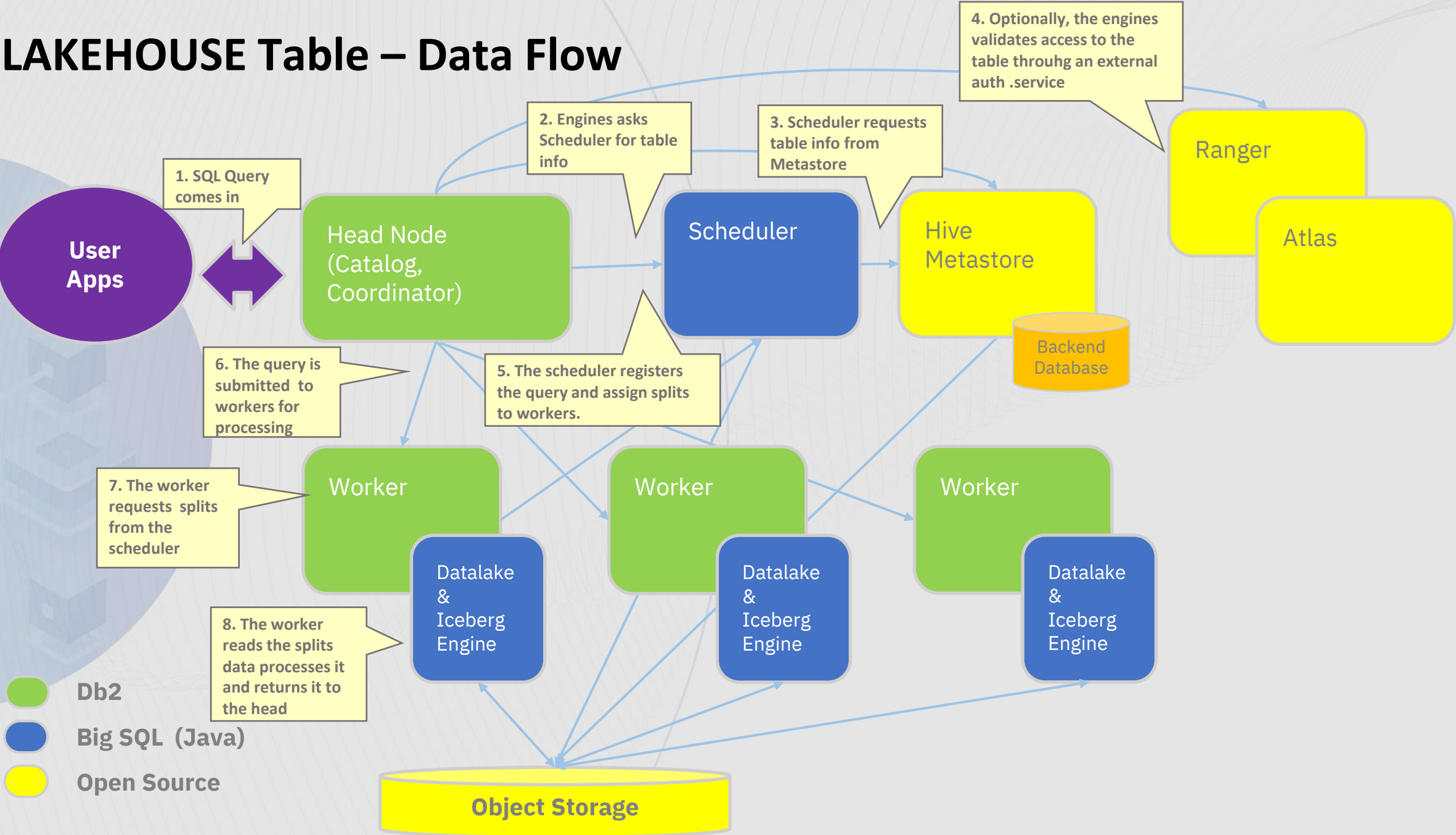
- A Data Lake “Table” is a collection of files serialized using an **Open Data File (ODF)** format (CSV, ORC, Avro, Parquet ...) stored on remote storage (HDFS, S3, COS, ...)
- The **metadata** of the table is stored in a Metastore server
 - Location
 - Schema
 - Partition lists
- An engine querying the table must query the metadata first and can proceed to read the data from remote storage
- Benefits
 - Interoperability of open data formats
 - Ease of use



LAKEHOUSE Table - IBM Db2 Warehouse - Datalake Table Support

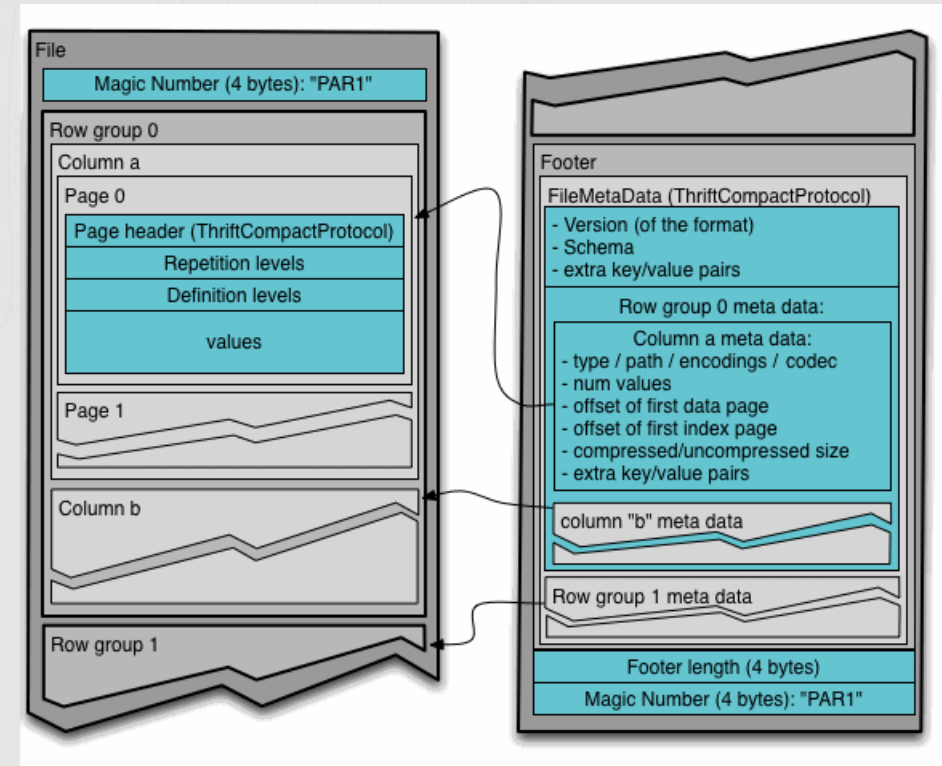


LAKEHOUSE Table – Data Flow



LAKEHOUSE Table – Limitations

- A Data Lake “Table” is a collection of files serialized following an Open Data File format
- Passive data structures – serialized set of data records
 - No notions of their **state** or **history**
 - No concurrency control between applications
 - **No ACID**, even less transactions
- Separate metadata
 - Need for a “**Catalog**”
 - No awareness of catalogs – it’s an external system



LAKEHOUSE Table – Introducing APACHE ICEBERG

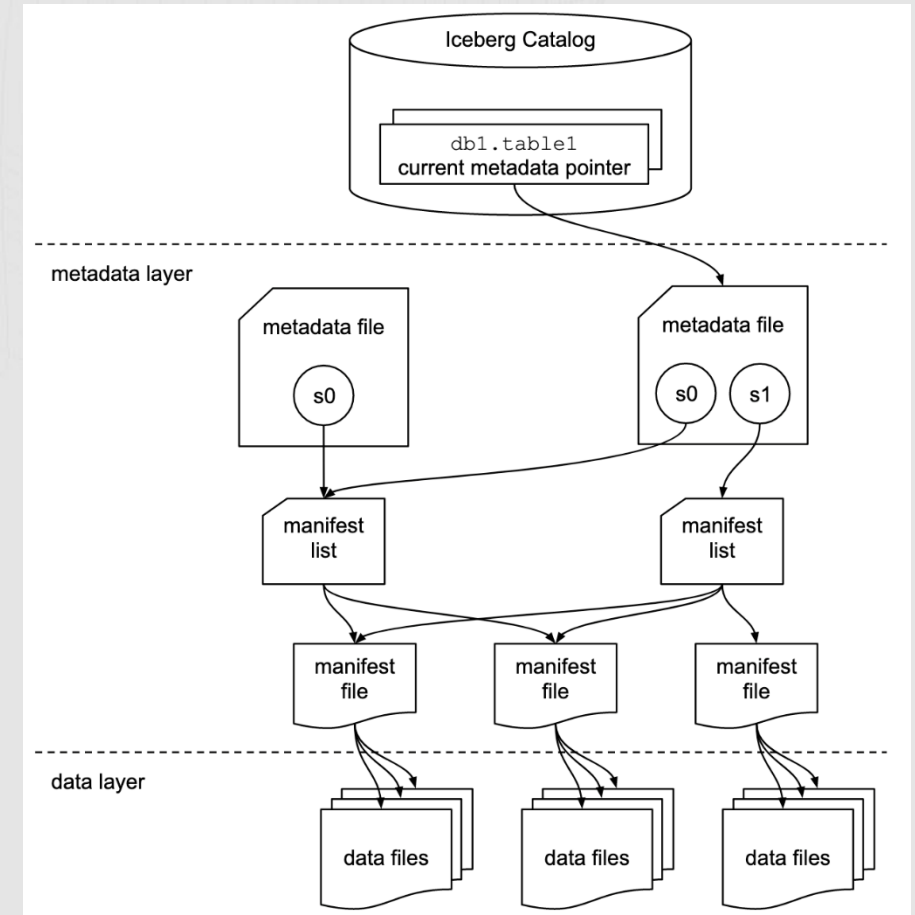
Apache Iceberg

An Open Data Table format for the Lakehouse



Full **open-source, Open Data Table format**, quickly becoming an **industry standard**

Relies on Open Data File formats for storage, but provides an additional layer of metadata for enhanced capabilities

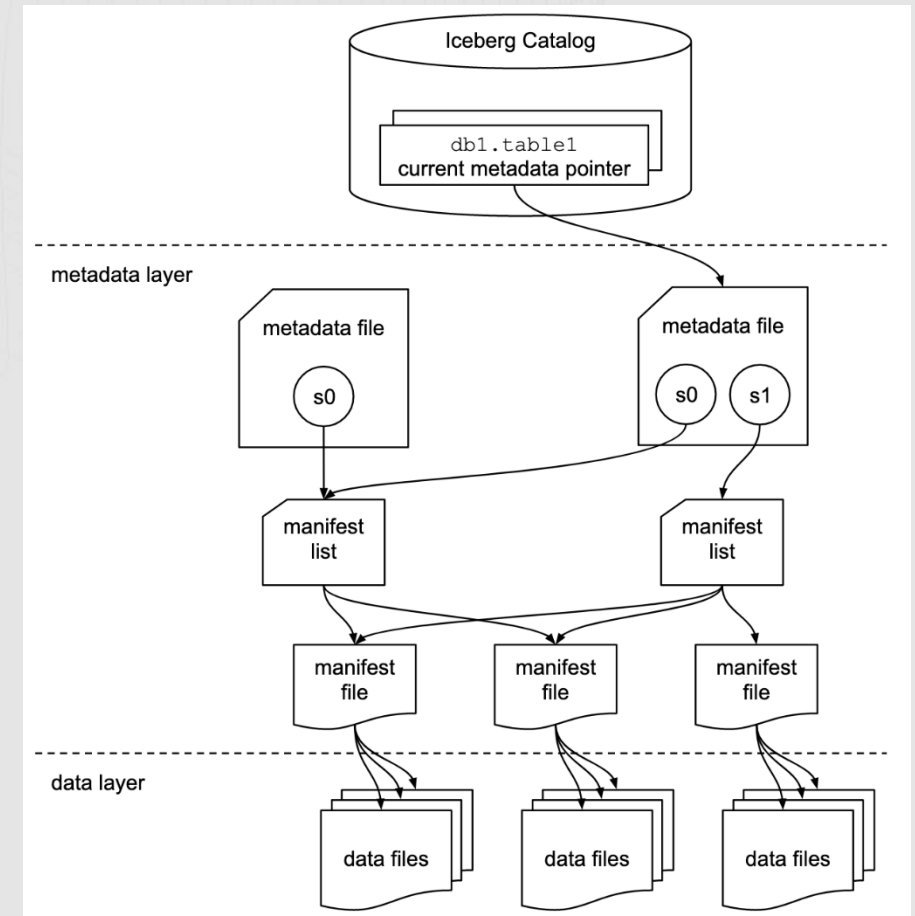


LAKEHOUSE Table – Introducing APACHE ICEBERG

Apache Iceberg Support

Restrictions in Db2 11.5.9

- Support for CREATE, SELECT & INSERT
 - Includes partitioning support
 - No UPDATE / DELETE
- No schema evolution
- No time travel yet
- Smaller restrictions related to Iceberg / Db2 type compatibility (nested types, etc ...)



Lakehouse & Warehouse - Designed to work together

Any number of these *runtimes* actively running inside Lakehouse instance

Any number of these *runtimes* actively interoperating with 1 or more Lakehouse instances

IBM Db2 Warehouse

Performance: +++
Price: \$\$\$

Netezza

Performance: +++
Price: \$\$\$

Performance: +.. +++
Price: \$.. \$\$\$

IBM Lakehouse

APACHE **Spark**

presto

Warehouse data is stored in proprietary format in block storage

IBM Db2

Block Storage

IBM Netezza

Block Storage

Lakehouse Metadata Service

Data in the Lakehouse is shared through common metadata stored in Lakehouse *Metastore* Service (HMS).

Warehouse extended to store proprietary format in object storage leveraging disk cache to optimize performance

Warehouse can access open data formats including Iceberg tables in object store

Lakehouse can "promote" data to warehouse. Warehouses may "offload" data to the lakehouse.

IBM Db2

Object Storage

Object Storage

Iceberg Parquet AVRO ORC JSON CSV

hadoop

Legacy Hadoop Deployments

Db2 Warehouse (SaaS) – Gen 3 – Based on Db2 11.5.9



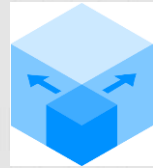
Fully managed / SaaS

Focus on the analytics, we'll take care of the rest



Blazing-fast

Columnar-organized, memory-optimized data warehouse



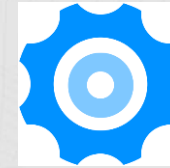
Scalable & elastic

Independently scale and manage compute & storage



Continuously available

Managed compute, highly available storage, cross-cloud replication



Reliable

Double protection with disaster recovery & self-service backup/restore



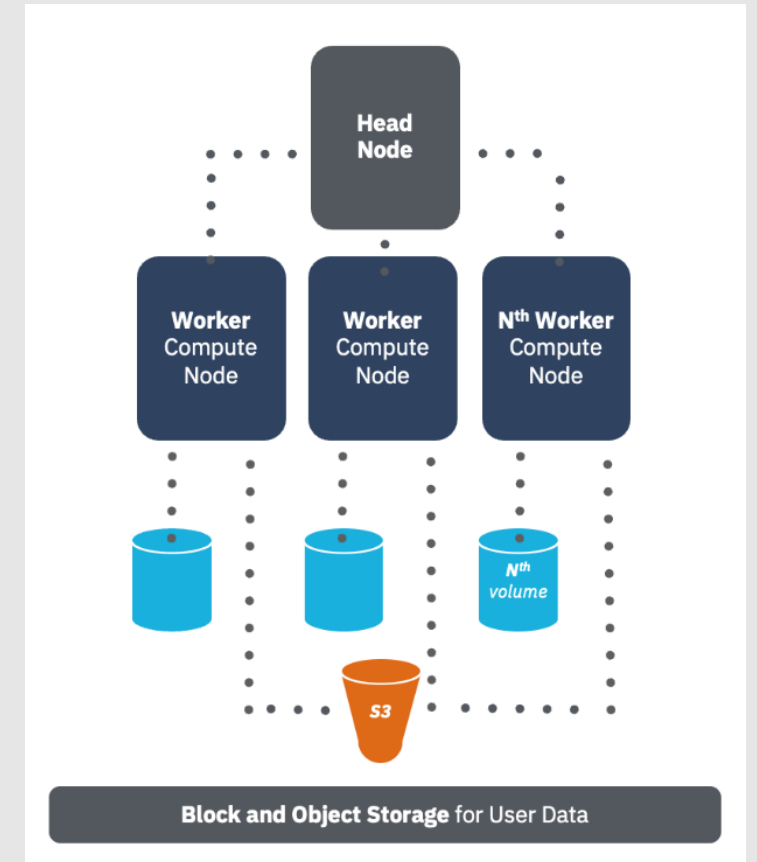
Secure

Encrypted at the storage level by the vendor and customer can protect data using their Key Protect keys.

Deploy as fully-managed DBaaS on **IBM Cloud** and **Amazon Web Services**

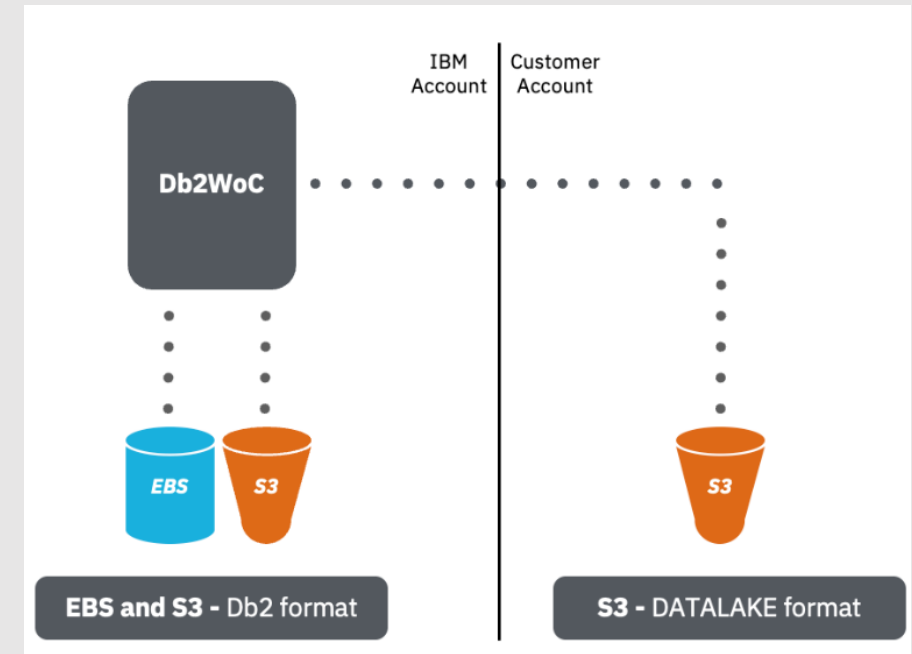
Db2 Warehouse (SaaS) – Gen 3 – Cloud Object Store

- Db2WoC now supports Amazon S3 object storage for database table storage, where customer data resides within the database
- Customer saves cost by using object storage instead of block storage
 - Allows customers to store data across a mix of block storage and object storage, based on business or technical requirements
 - Db2WoC uses different mechanisms in order to facilitate reads and writes to object storage
- No applications and workload changes necessary in order to use this feature
 - Db2WoC handles all the necessary interfacing to object storage, thus existing applications and warehouse workloads do not have to be changed in order to make use of this
 - Specific Db2 tablespace available backed by S3 for customer use
 - Insert, Update, Delete data as needed into and out of tables within object storage
 - Move and copy data to and from column-organized tables residing in block storage and object storage
 - Query data seamlessly no matter where it resides (in block or object storage), in isolation or in combination with each other



Db2 Warehouse (SaaS) – Gen 3 – LAKEHOUSE Table

- Db2WoC now also supports Open Data Formats as DATALAKE tables, allowing for seamless access to other data within the enterprise for integrated workloads.
- Leverage existing compute resources dedicated to the warehouse
 - Facilitate data use to and from the Db2 Warehouse to quickly access a variety of enterprise data
 - Leverage the high performance Db2 engine for queries against enterprise data
- Db2WoC provides interfaces for customers to leverage their enterprise data residing in object storage as DATALAKE tables
 - Supports both regular and Iceberg DATALAKE table types, based on existing data formats or for business/technical requirements such as ACID compliance
 - Browse, explore, and query enterprise data in both Db2 and DATALAKE formats, using either the web-based UI, or through SQL
 - Access data in place within DATALAKE tables, joining as necessary with Db2 based data for queries
 - Access data within DATALAKE tables and import into Db2 formatted tables
 - Create new DATALAKE tables in S3 and export from Db2 formatted tables

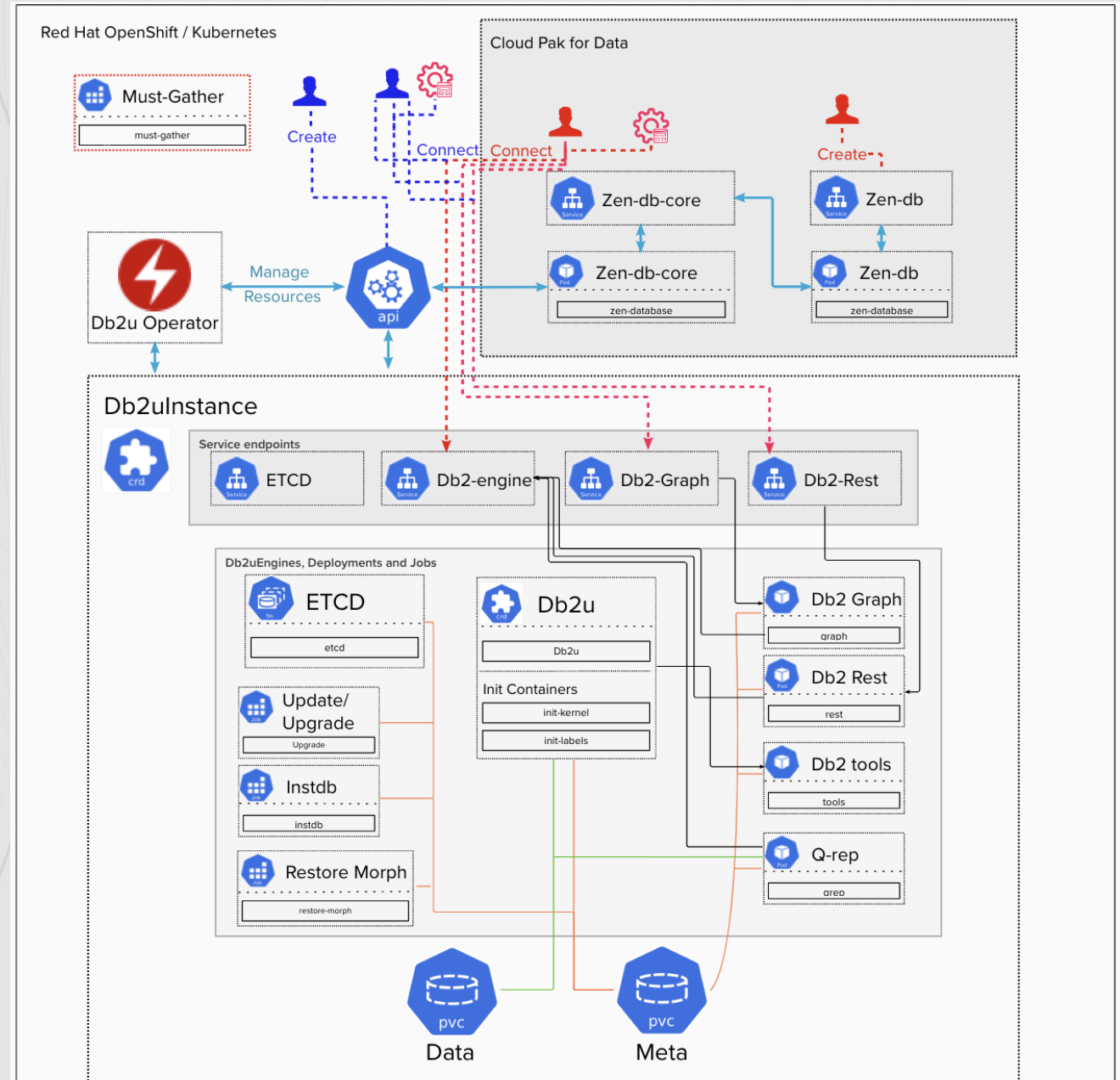


Db2 Warehouse (SaaS) – Gen 3 – Product Tiers

	Starter Model	Middle-tier model	Large-scale Model
	Storage Optimized An MPP data warehouse ideal for cost-effective data warehousing, dev/test environment, infrequently accessed data. Support for Object Storage + DataLake tables	Compute Optimized An MPP data warehouse ideal for high-performance and production workloads Support for Object Storage + DataLake tables	Compute Optimized An MPP data warehouse ideal for high-performance and production workloads and very large systems Support for Object Storage + DataLake tables
Architecture	MPP (Massively Parallel Processing)	MPP (Massively Parallel Processing)	MPP (Massively Parallel Processing)
Configuration	Dedicated compute & storage	Dedicated compute & storage	Dedicated compute & storage
Cloud Platform Support	Amazon Web Services	Amazon Web Services	Amazon Web Services
Relative Cost	\$\$\$	\$\$\$\$	\$\$\$\$\$
Independently scale storage & compute	Yes	Yes	Yes
Storage for Table Data	Up to 80 TB (Block + Object Storage)	Up to 240 TB (Block + Object Storage)	Up to 600 TB (Block + Object Storage)
Compute Range	16 – 128 cores	48 – 576 cores	288– 2880 cores
Private Endpoint	Available	Available	Available

Db2 Warehouse (SaaS) – Gen 3 – Based on Db2U

- Managed via k8s Custom Resource (CR)
 - Kind: Db2uInstance
 - Short name: db2
- Underlying Kubernetes resource model:
 - Db2 Engine Pod lifecycle managed using a NEW Db2uEngine CR, which manages a collection of Db2 Pods.
 - Onetime tasks managed via a Job resource
 - In-pod HA to recover Db2 failures, avoiding a pod lifecycle event. This built-in HA leverages ETCD for state information
 - Lifecycle of (stateless) Add-Ons (REST, Graph, Q-rep, etc.) managed via Deployment resources



Db2 Warehouse (SaaS) – Gen 3 – Based on Db2U

- DB2u containers are the base.
- Every customer deployment maps to a **Formation**. A formation is realized by the Operator model and deploys a complete cluster
 - Gets storage provisioned
 - Gets N/W components created
 - Creates container/pod configurations based on formation specs
 - Db2u process sets up a db2 database with the storages and sets up tablespaces with them
 - Customer gets two endpoints, one to the Console and the other to the database.
- **Recipes**
 - Used for all Operations like scale, backup, addition of users etc.

Db2 Warehouse (SaaS) – Gen 3 – Highly Scalable

- Scale using the Console or APIs
 - The platform provides mechanics to initiate a scale (we create a scale recipe)
 - In the scale recipe, we provide the new core and storage values.
- The operator framework will
 - Gracefully shutdown the engine and scale down the formation
 - Bring it up in the new spec (appropriately setting the correct MLN count per pod)
 - As the new spec is brought up, new nodes are automatically added by the Cluster Auto Scaler
 - If it is a shrink operation, unused nodes are automatically returned by the Cluster Auto Scaler
- Compute Expansion / Shrink incurs a Downtime
- Storage scaling works the exact same way (recipe mechanism)
 - Storage scale up is a totally online operation
 - Block Storage shrink (support date TBD) will be offline in nature

Db2 Warehouse (SaaS) – Gen 3 – Improved HA

- **Multi prong, provided by**
 - Wolverine (our HA component) - Core Db2 container
 - Kubernetes - non Db2 containers
- **Wolverine responsible for and provides**
 - System recovery (when the pods fail)
 - Overall System health and status recording in ETCD
 - Device Manager framework
 - HA loop (for monitoring and handling HA for processes within core Db2 Container)
 - API endpoints
- **Kubernetes**
 - Pod level HA based on liveness probes
- **When a Db2 node goes down,**
 - Kubernetes will automatically schedule the pod to run on a free node from the pool
 - Wolverine will wait for the pod to come back up and take the engine through a recovery process
 - Engine is back running in full capacity within a short time.
- **When a Db2 pod goes down**
 - Kubernetes will restart the pod
 - Wolverine will wait for the pod to come back and take the engine through a recovery process.

Db2 Warehouse (SaaS) – Gen 3 – More Granular Backup/Restore

- Backups - a combination of Snapshot Backups and S3 backup
 - Scheduled every 24 hour period or on demand
 - Scheduled or on-demand both internally create a Recipe
 - The Recipe will go through the following phases
 - Copy s3 metadata to block storage location
 - Do a db2 write suspend (including S3 deletes off)
 - Create volume snapshots
 - Initiate S3 backup process
 - Backup metadata
 - Write resume (but leave S3 delete on)
 - When AWS S3 backup completes (tag backups and turn on S3 deletes)

Restore process, On-demand, internally creates a Recipe.

The recipe will go through the following phases

- Stop Db2
- Initiate meta data restore
- Initiate S3 restore
- Initiate volume snapshot restore
- When S3 restore is ready to use, scales all pods up

Moving Forward from PDOA/IIAS



IBM Integrated
Analytics System
(IIAS)



IBM PureData System
for Operational Analytics
(PDOA)

Migrate to...



1. Db2 Warehouse SaaS

Fully managed, high performance, cloud-elastic data warehouse

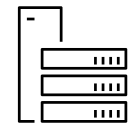
- ↳ Runs on IBM Cloud, AWS
- ↳ Available in US, EU, UK, JP, AP commercial regions



2. Db2 Warehouse Self-managed

Self-managed, data warehouse on cloud provider managed Kubernetes or OpenShift infrastructure.

- ↳ Runs on Amazon EKS, Amazon ROSA, Azure AKS, Azure ARO
- ↳ Available worldwide



3. Db2 Warehouse Cloud Rack

On-premises, OpenShift high performance, data warehouse deployable on Power 10 or X86 infrastructure

- ↳ Runs on POWER10 or x86 reference architecture

Moving Forward from PDOA/IIAS – Cloud – Fully Managed



IBM Integrated
Analytics System
(IIAS)



IBM PureData System
for Operational Analytics
(PDOA)

Migrate to...



1. Db2 Warehouse SaaS

Fully managed, high performance, cloud-elastic data warehouse

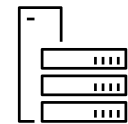
- ↳ Runs on IBM Cloud, AWS
- ↳ Available in US, EU, UK, JP, AP commercial regions



2. Db2 Warehouse Self-managed

Self-managed, data warehouse on cloud provider managed Kubernetes or OpenShift infrastructure.

- ↳ Runs on Amazon EKS, Amazon ROSA, Azure AKS, Azure ARO
- ↳ Available worldwide

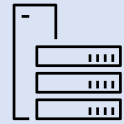


3. Db2 Warehouse Cloud Rack

On-premises, OpenShift high performance, data warehouse deployable on Power 10 or X86 infrastructure

- ↳ Runs on POWER10 or x86 reference architecture

Moving Forward from PDOA/IIAS – Cloud – Self Managed



IBM Integrated
Analytics System
(IIAS)



IBM PureData System
for Operational Analytics
(PDOA)

Migrate to...



1. Db2 Warehouse SaaS

Fully managed, high performance, cloud-elastic data warehouse

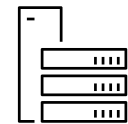
- ↳ Runs on IBM Cloud, AWS
- ↳ Available in US, EU, UK, JP, AP commercial regions



2. Db2 Warehouse Self-managed

Self-managed, data warehouse on cloud provider managed Kubernetes or OpenShift infrastructure.

- ↳ Runs on Amazon EKS, Amazon ROSA, Azure AKS, Azure ARO
- ↳ Available worldwide



3. Db2 Warehouse Cloud Rack

On-premises, OpenShift high performance, data warehouse deployable on Power 10 or X86 infrastructure

- ↳ Runs on POWER10 or x86 reference architecture

Moving Forward from PDOA/IIAS – Cloud – Self Managed

Db2 Warehouse Software

Container, Operator and Reference Architecture

1

Container: Launch a standardized Db2 container across your entire deployment.

2

Operator: Easily manage the lifecycle of that container and various database functions.

3

Reference architecture: Detailed guide on how to deploy Db2 Data Warehouse on various topologies.

Step-by-step Tutorials for deployment

[Db2 on AWS Marketplace](#)



[Db2 Warehouse on AWS EKS](#)



[Db2 Warehouse on AWS ROSA](#)

[Db2 on Azure Marketplace](#)



[Db2 Warehouse on Azure AKS](#)



[Db2 Warehouse on Azure ARO](#)

Moving Forward from PDOA/IIAS – On-Premises



IBM Integrated Analytics System (IIAS)



IBM PureData System for Operational Analytics (PDOA)

Migrate to...



1. Db2 Warehouse SaaS

Fully managed, high performance, cloud-elastic data warehouse

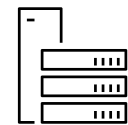
- ↳ Runs on IBM Cloud, AWS
- ↳ Available in US, EU, UK, JP, AP commercial regions



2. Db2 Warehouse Self-managed

Self-managed, data warehouse on cloud provider managed Kubernetes or OpenShift infrastructure.

- ↳ Runs on Amazon EKS, Amazon ROSA, Azure AKS, Azure ARO
- ↳ Available worldwide



3. Db2 Warehouse Cloud Rack

On-premises, OpenShift high performance, data warehouse deployable on Power 10 or X86 infrastructure

- ↳ Runs on POWER10 or x86 reference architecture

IBM Power 10 Cloud Rack For Db2 Warehouse

Complete Data Warehouse Solution

*Simplify and accelerate your data warehouse deployment in days with our Power 10 based pre-configured solution that includes everything you need to **securely** analyze your data on your premises – including Servers, Storage, Switches, and Services!*



IBM Db2



• Simple

- Ease into a Db2 data warehouse with a **pre-configured, all-in-one design** that removes the guess work

• Fast

- Get your on-premises data warehouse up and running in days vs. weeks or months to **start transforming your business faster**

• Efficient

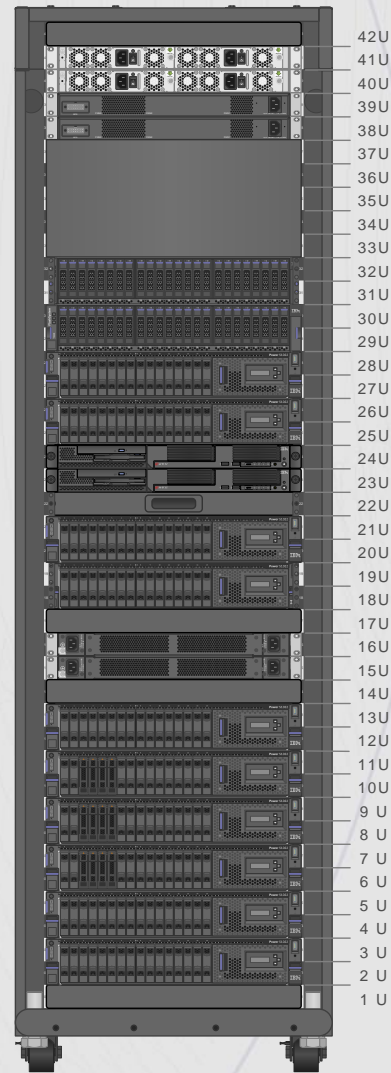
- **Integrate** traditional environments (AIX & IBM i) and new cloud-native applications with Red Hat OpenShift **synergistically**, managed in an on-premises private cloud

• Flexible

- Build the data warehouse needed for today and **easily scale and adapt** as your future needs change

Enable a scalable Power 10 Db2 Data Warehouse that grows as your data increases!

- **Experience:** A balanced resilient architecture for maximum reliability and performance
- **Flexibility:** The Data Warehouse starts small, but grows as you do
- **Efficiency:** E-config enables the full stack pre-selected and configured as a “solution edition” set-up
- **Storage:** IBM FlashSystem adds simplicity and consolidation, capable of delivering more work with fewer drives and lower cost



On-Premise HW Stack

Servers:

- OpenShift Management: 3x Power 1022s
- Workers: From 3 to 80 - IBM Power 1022

Storage:

- IBM FlashSystem 7300 for Tier-1
- IBM FlashSystem 5200 for Tier-2

High Speed Networking

- IBM 100 GB Ethernet Switching
- IBM Storage Networking SAN

Management:

- IBM HMC

Warehouse Software Stack

- RHEL OpenShift
- Db2 Warehouse on OpenShift
- IBM PowerVC
- IBM Spectrum Scale Data Management Edition
- PowerVM Enterprise

IBM Services

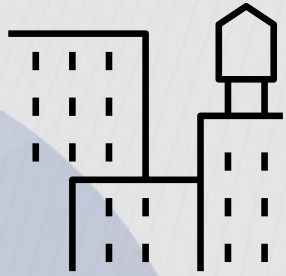
Pre-sales:

- Size an appropriate configuration for customer's existing workload.
- Collaboration with business partners.

Install:

- Build the infrastructure hardware stack and pre-load the OCP/SS software stack in IBM MFG
- Gold-glove installation of the infrastructure hardware/software stack at customer's premises
- Remote installation/customization of Db2 Warehouse Software on the infrastructure at customer's premises
- Migration Assessment session with the customer

Db2 Warehouse *Reference Architecture Solutions*



On-premises offerings that provide an “appliance-like” experience and a smooth upgrade path from PDOA/IIAS (Sailfish) appliances¹.

- Maximizes simplicity, minimizes time to value, and ensures fully supported environments
- *Optimized for performance* and scalability
- **Prescriptive** configuration of IBM hardware, Db2 Warehouse software and Technology Expert Lab services
- IBM built and lab tested
- Includes *Gold-glove Installation* and *Solution Support Services* with a single point of entry.

¹As opposed to Appliances where Db2 Warehouse licenses are an integrated part of the solution, in a Reference Architecture Db2 License can be separated from the Solution and moved to another infrastructure.

¹Clients can also procure the prescribed hardware, software, and Expert Lab services a-la-carte to create a roll-your-own (RYO) or Do-It-Yourself (DIY) self-managed on-premises deployment of a Reference Architecture. However, for RYO/DIY self-managed deployment of a Reference Architecture Solution, IBM will provide only component level installation and support services and **Gold-glove Installation** and **Solution Support Services will not be available to RYO/DIY Reference Architecture Solutions.**

PDOA/IIAS vs. P10 Cloud Rack for Db2 Warehouse

Hardware View

• IBM Integrated Analytics System (IIAS / Sailfish)

- Power 8 2U Servers
- FS900 Flash Modules
- 4x10 Gb bonded network
- 16 Gb HBAs
- Db2 Warehouse



• P10 Cloud Rack for Db2 Warehouse

- Power 10 2U Servers
- FS7300 Flash Modules
- 100 Gb bonded network
- 32 Gb HBAs
- Modernized Db2 Warehouse on OpenShift



Customer Value

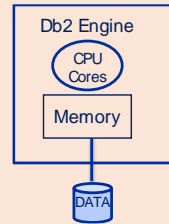
- Substantially improved performance and capacity for the same footprint
- Modernized Db2 Warehouse deployment
- Risk-free ease of migration with options for modernization
- Single-point-of-Contact for all support issues
- Long Solution life and less costly

PDOA/IIAS vs. P10 Cloud Rack for Db2 Warehouse

Multiple Logical Node (MLN) / Hash Partition View

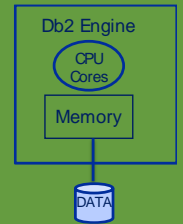
- IBM Integrated Analytic System (IIAS / Sailfish)

- 8 Data MLNs per Server
- 3 Power8 Cores per MLN
- 1.6TB Usable Tier 1 space per MLN
- 51GB Memory Per MLN



- P10 Cloud Rack for Db2 Warehouse

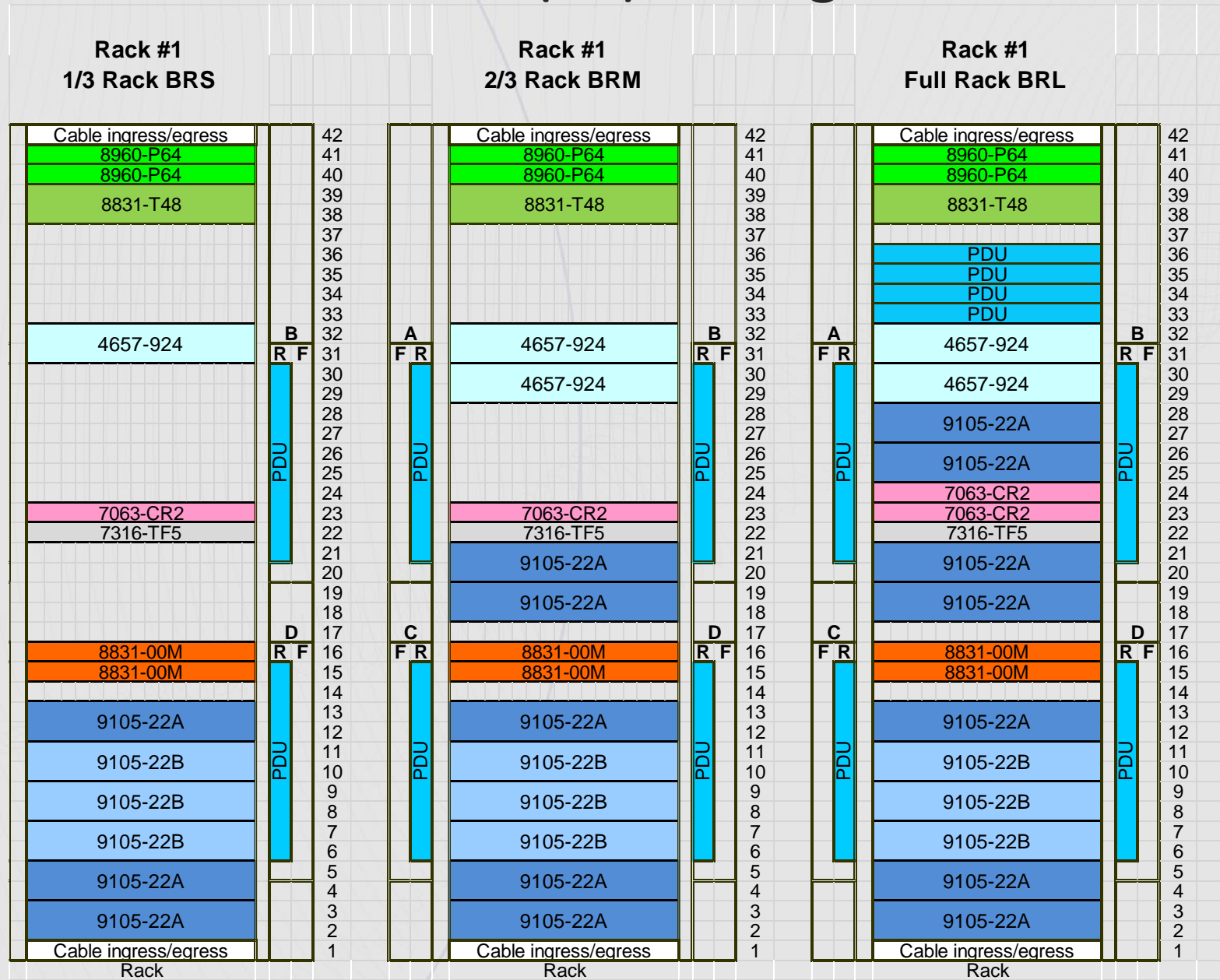
- 8 Data MLNs per Server
- 4 Power10 Cores Per MLN
- 3TB Usable Tier 1 space per MLN
- 100GB Memory Per MLN



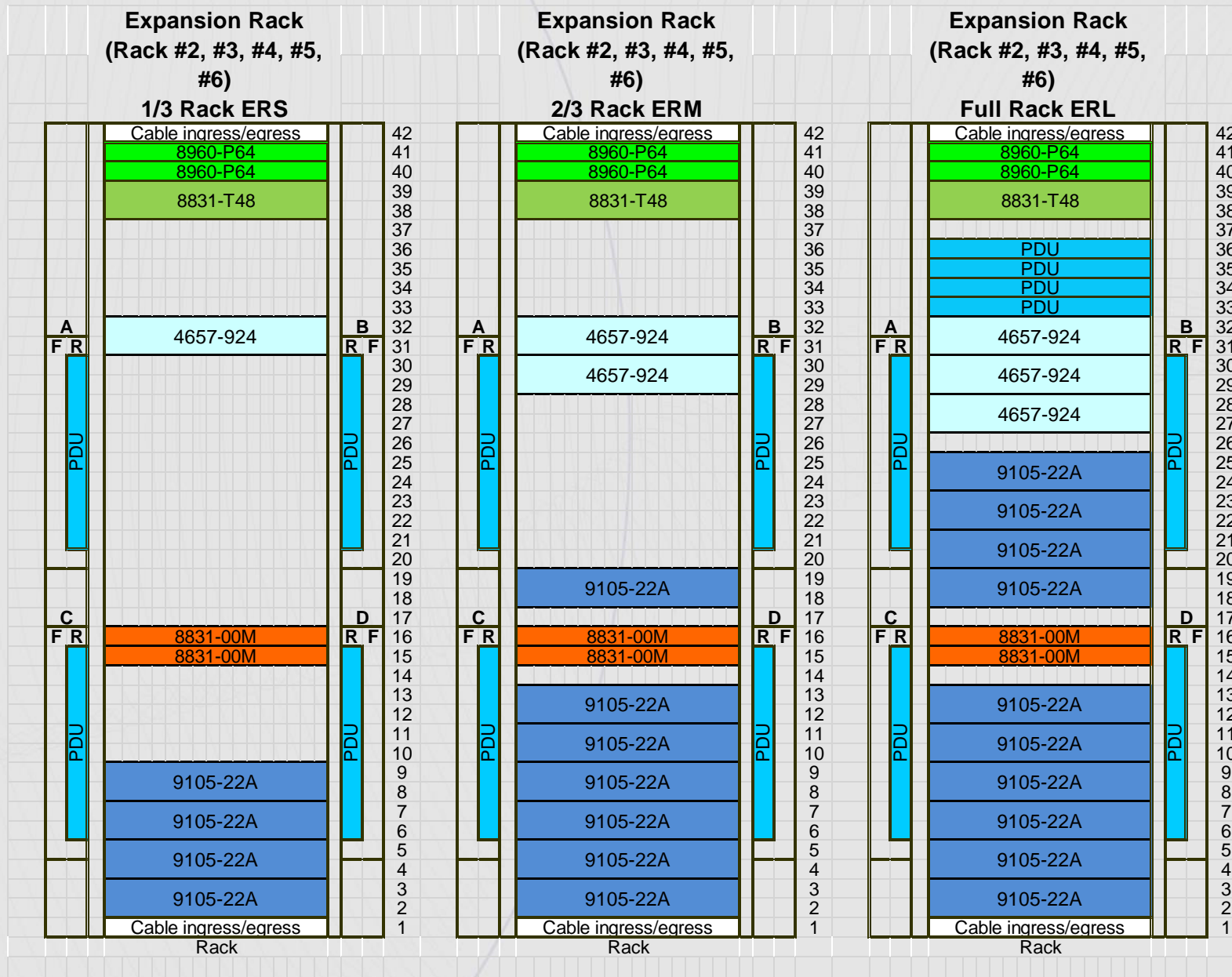
Addition Cloud Rack for Db2 Benefits

- Db2 on RedHat OpenShift embraces Micro Services Multiple Containers vs. One Large Container, **More Granular Upgrades** reducing upgrade time.
- Failover Server per Rack: **Consistent Performance** while in Failover Mode.
- Kubernetes deployment: **Simplifying Initial deployment, Expansion, High Availability, etc.**
- Private Cloud deployment: **Simplifying movement to other Cloud Form Factors**

Cloud Rack for Db2W Base Rack (BR) Configurations



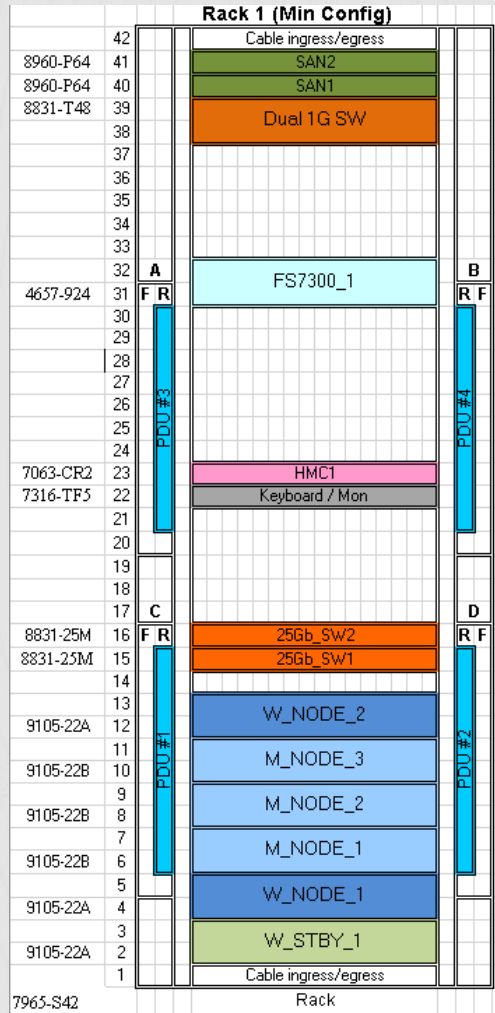
Expansion Rack (ER) for Db2W Base Rack Configurations



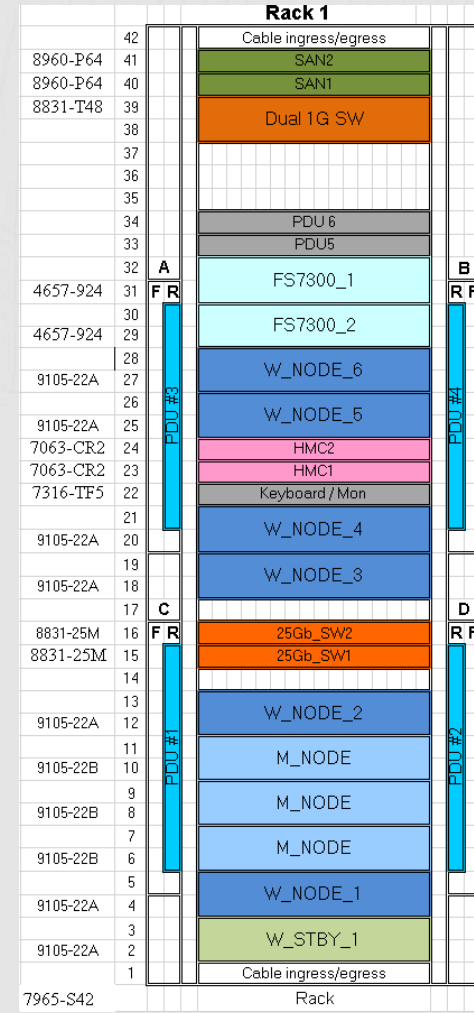
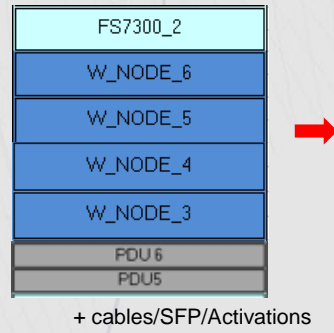
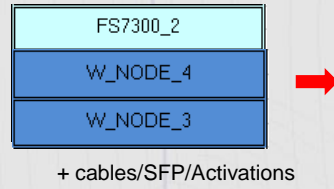
Db2 Warehouse P10 Cloud Rack Solution

Minimum Configuration and Future Expansion Options

and
e Out

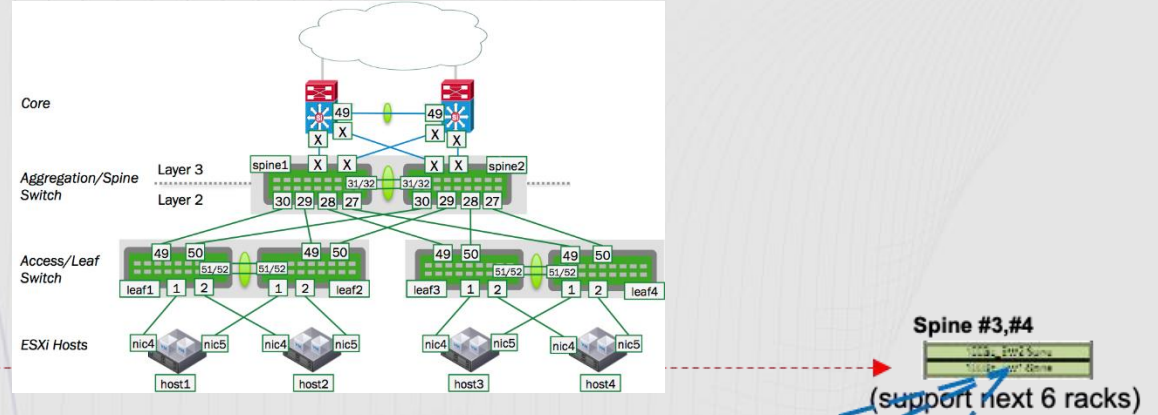


Expansion Options:



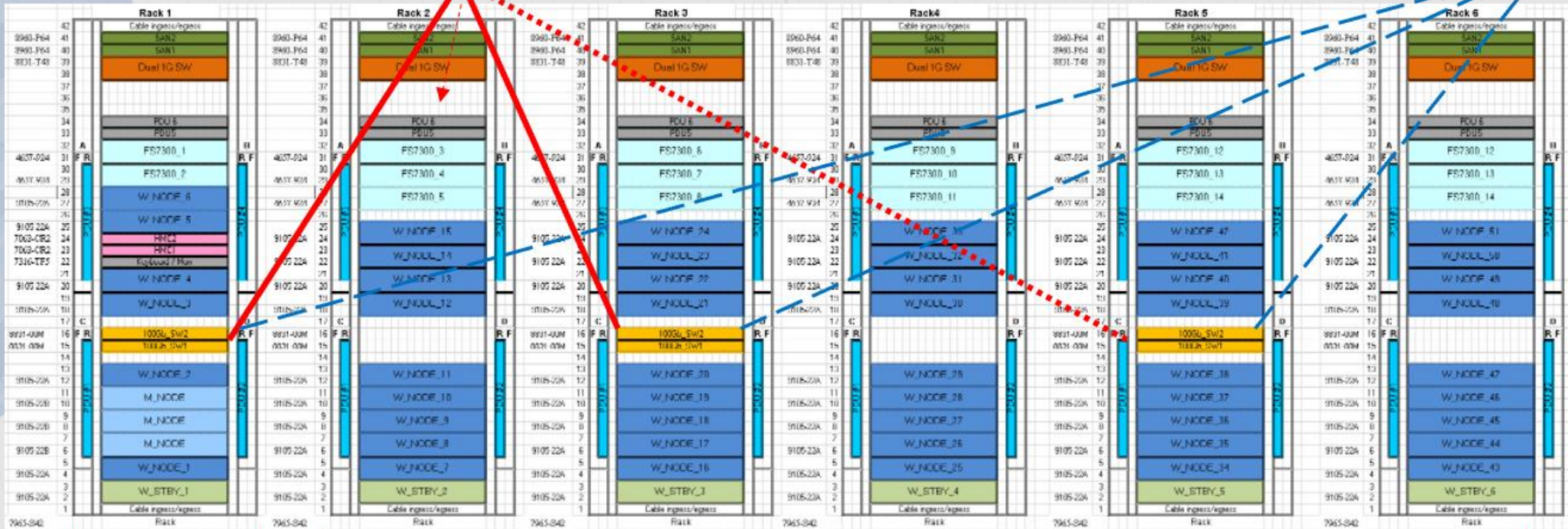
Db2 Warehouse P10 Cloud Rack Solution

A Multi-Rack Configuration with Network



Spine #1,#2
 1005L_Sw2 Spine
 1005L_Sw1 Spine (support 6 racks)

Spine #3,#4
 1005L_Sw2 Spine
 1005L_Sw1 Spine (support next 6 racks)



Foundation Rack

Expansion Racks

Gold-glove Installation – P10 Reference Architectures

Hardware and OS

- Pre-installation assistance to the client
 - Answer installation/shipping/delivery questions
 - Assist in preparing facilities for the system, e.g., power/cooling compliance, and power/network infrastructure
- On site installation
 - Rack placement
 - Cabling
 - Hardware verification
 - Power on and inspection
 - Upgrades of firmware and OCP/ODF/Spectrum Scale to bring them up to rev
 - network connectivity
 - Sanity check (diagnostics)
 - Stage DB2 SW on system storage
- Post-installation assistance
 - Collect logs

Db2 Software

- Remote installation
 - Install db2u operators
 - Update the configuration of the worker node to match Db2u recommended setting
 - Install db2u entitlement key
 - Create the db2 instance and database (run YAML file)
 - Db2 instances are now Table Ready
 - Hand over the system to the client for ingesting data
- All P10 configurations will be shipped with OCP/Spectrum Scale software installed at the Factory.
- OCP Licenses are included with Db2 Warehouse Cartridges.
- Spectrum Scale licenses are included in eConfig Quote.
- Db2 Warehouse Cartridge Licenses are included in DSW Quote
- OCP/SS/Db2 licenses are enabled during gold-glove installation.

Expertise Connect Technical Account Manager - Ensure Success



• Deliverables

- **Proactive** guidance for preventing issues
- **Remediate** issues
- Achieve application project **milestones**
- **Change management** recommendations
- **Technical** reviews
- **Transformation**
- **Use case** assistance



• Custom Technical Guidance

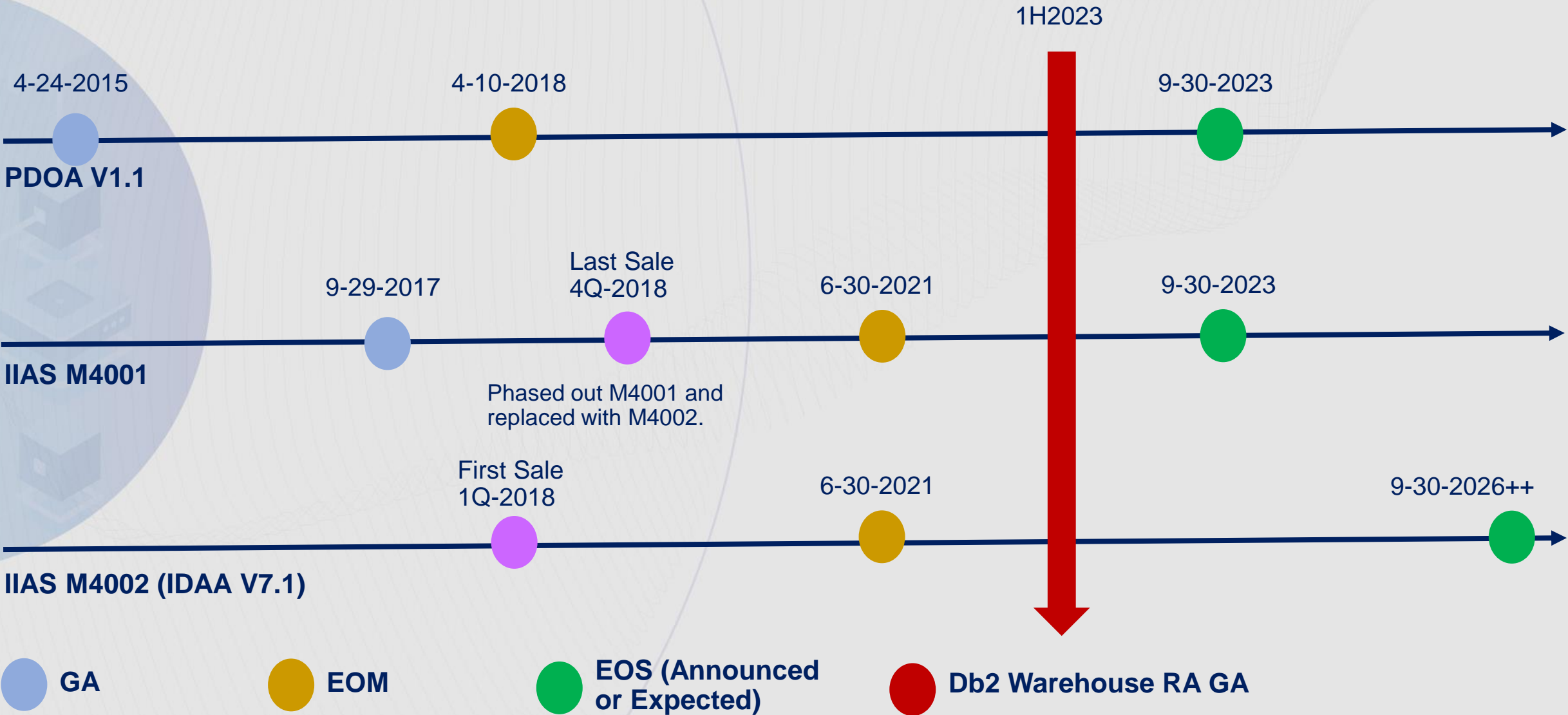
- Validate your design with **interoperability and security** concerns
- Performance, stability and upgrade readiness **assessments**
- Architecture & health **reviews**
- **Assess skill levels** & identify education options to cover gaps



• Knowledge Sharing

- **Best practices**
- Access to IBM's **product experts**
- Build **your team's competency** on making the most of IBM software
- Information **contextual** to your specific needs
- **Custom tech talks** on key topics

Currently Installed IBM Db2 Warehouse Appliances



Db2 11.5 – What's Up with the Numbering ?

Db2 11.5 follows a continuous delivery model

(same as 11.1)

If a delivery vehicle has:

- new function -- will increment Mod level

(same as 11.1)

- But it will also reset the FP level to 0

(*NEW*)

→ 11.5.M.0

- ONLY fixes, we've tried 2 options:

- Fixpack - increment just the FP level
 - only done once in life of 11.5 (11.5.5.1)

(same as 11.1)

- CSB, no numbering change, it is a cumulative SB level
 - Maintained continually on 11.5.4 and later

Db2 11.5 – Cumulative Special Builds vs “Past” Special Builds

How are they alike:

- both are how customers get APARs fixes “asap”
- both are identified by a special build ID (versus change in mod or FP level)

How do they differ (aka why is the CSB much better).

- C stands for Cumulative / Continuous
 - The CSB is all APARs (including security) fixed since the last mod pack.
- Any/all customers pick up the same stream, the previous special build model resulted in “1-of” builds for customers
- Security SB is just a CSB w/ the latest security fixes-- they are part of the cumulative stream.
- Being a continuous stream, it gets continuous testing including pipeline, functional, performance and stress; previous SB had minimal testing (decided per fix).

If a customer is installing a new Db2 instance or moving up to 11.5; it is recommended that they pick up the latest CSB stream (will include the latest).

Note we still have ability to do 1-ofs (and they are call unofficial SB) but used almost exclusively for debug code or partial fixes not meant or ready for everyone’s use.

Db2 11.5 – And the winner is

Rely on CSB stream, no fix packs.

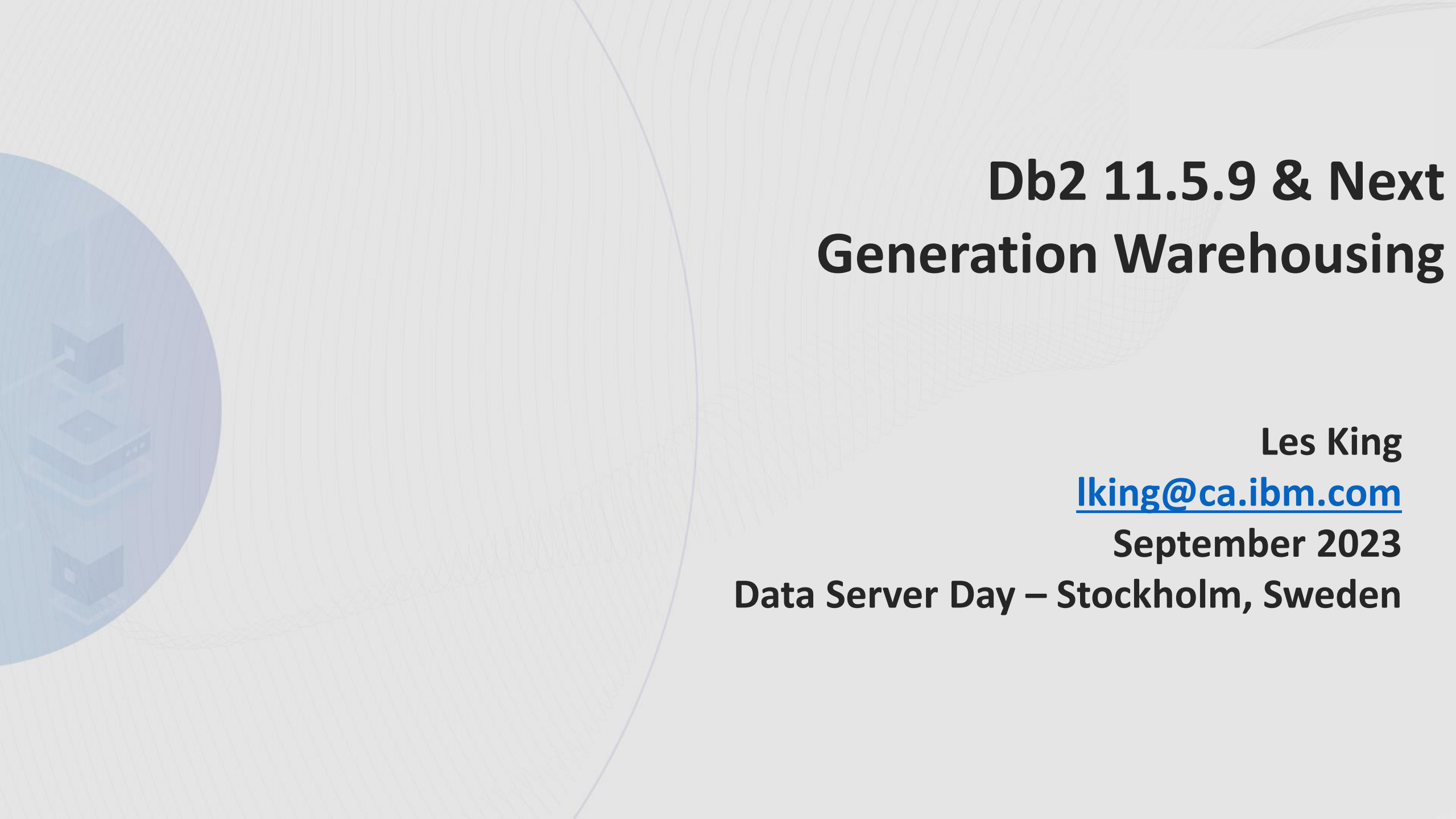
Fix packs simply package many APARs together.

Good:

- as a periodic encouragement to move up; “look here’s a bunch of great fixes”
- to give a target level/date to help support/justify a plan to update (to mmgt, other teams, etc)

But:

- often dates don’t align with when customers want to do a new install, or upgrade or even pre-planned updates.
- it is stale the moment it ships, let alone by the time the next one comes out
 - e.g mod 11.5.8 released it’s first CSB with ~50 additional APAR fixes the same day
- same issue tends to be hit multiple times till next FP out (which can months)



Db2 11.5.9 & Next Generation Warehousing

Les King

lking@ca.ibm.com

September 2023

Data Server Day – Stockholm, Sweden