

Les King

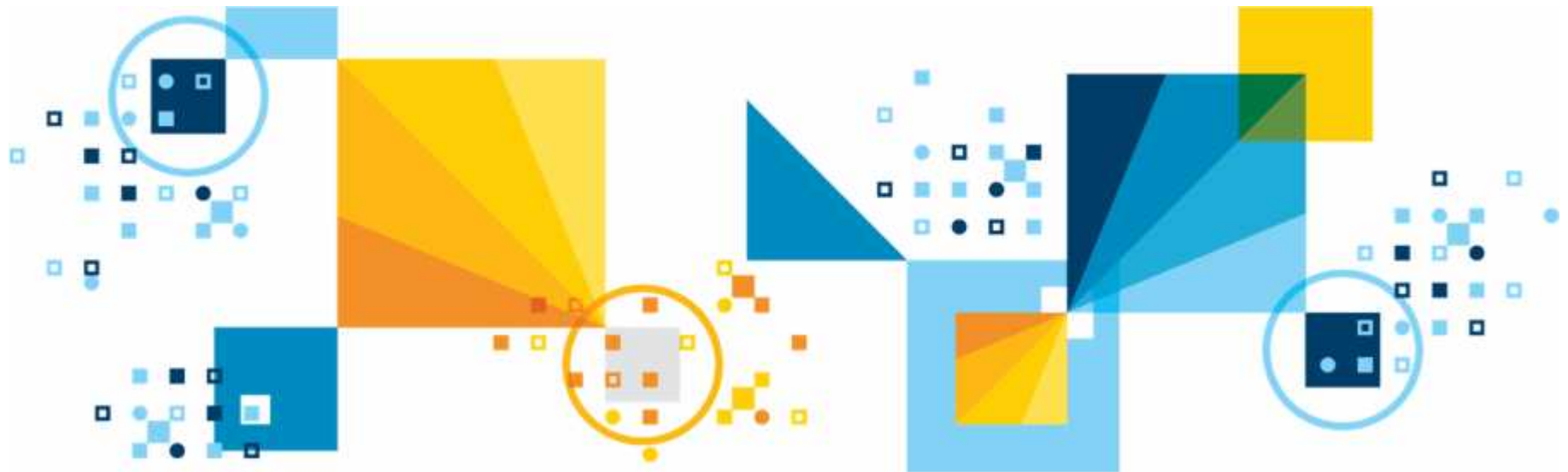
Director, Hybrid Data Management Solutions

September, 2018

lking@ca.ibm.com

ca.linkedin.com/pub/les-king/10/a68/426

Hybrid Data Management Strategy and New News !



IBM's Strategy is HYBRID

Its not about Cloud or On-Premises its about **Cloud** AND **On-Premises**

Its not about Traditional Relational or Open Source its about **Traditional Relational** AND **Open Source**

It's About Hybrid

Its not about SQL or NoSQL its about **SQL** AND **NoSQL**

Its not about Structured or Unstructured Data its about **Structured** AND **Unstructured** Data

Common SQL Engine – Business Value

A **COMMON SQL ENGINE** enabling true **HYBRID** data solutions for **ALL WORKLOAD** types

Systems of Record

Systems of Engagement

Systems of Insight

Event Processing



Investment Protection

WRITE ONCE, RUN ANYWHERE



PUBLIC CLOUD



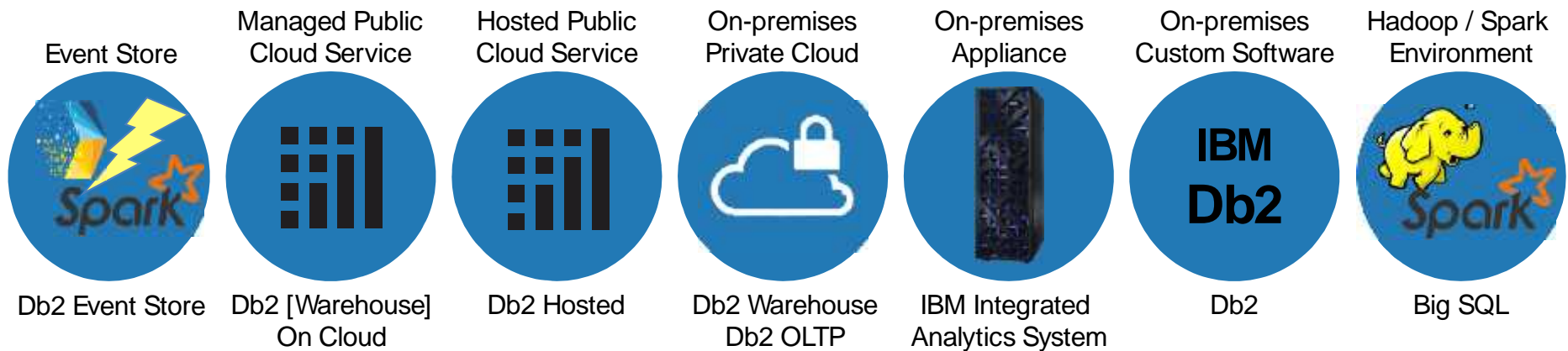
PRIVATE CLOUD



ON-PREMISES

Common SQL Engine – Consistent Technical Capabilities

A **COMMON SQL ENGINE** enabling true **HYBRID** data solutions for **ALL WORKLOAD** types



Foundation

- ✓ Full MPP scalability (GB-PB)
- ✓ High Concurrency
- ✓ Load and Go Simplicity
- ✓ Consistent Management and WLM
- ✓ HA, DR & Replication
- ✓ Integrated Security & Encryption

Application

- ✓ Built-in analytics (OLAP)
- ✓ Data Virtualization
- ✓ Application portability
- ✓ Hybrid by design
- ✓ Oracle Compatibility
- ✓ Netezza Compatibility

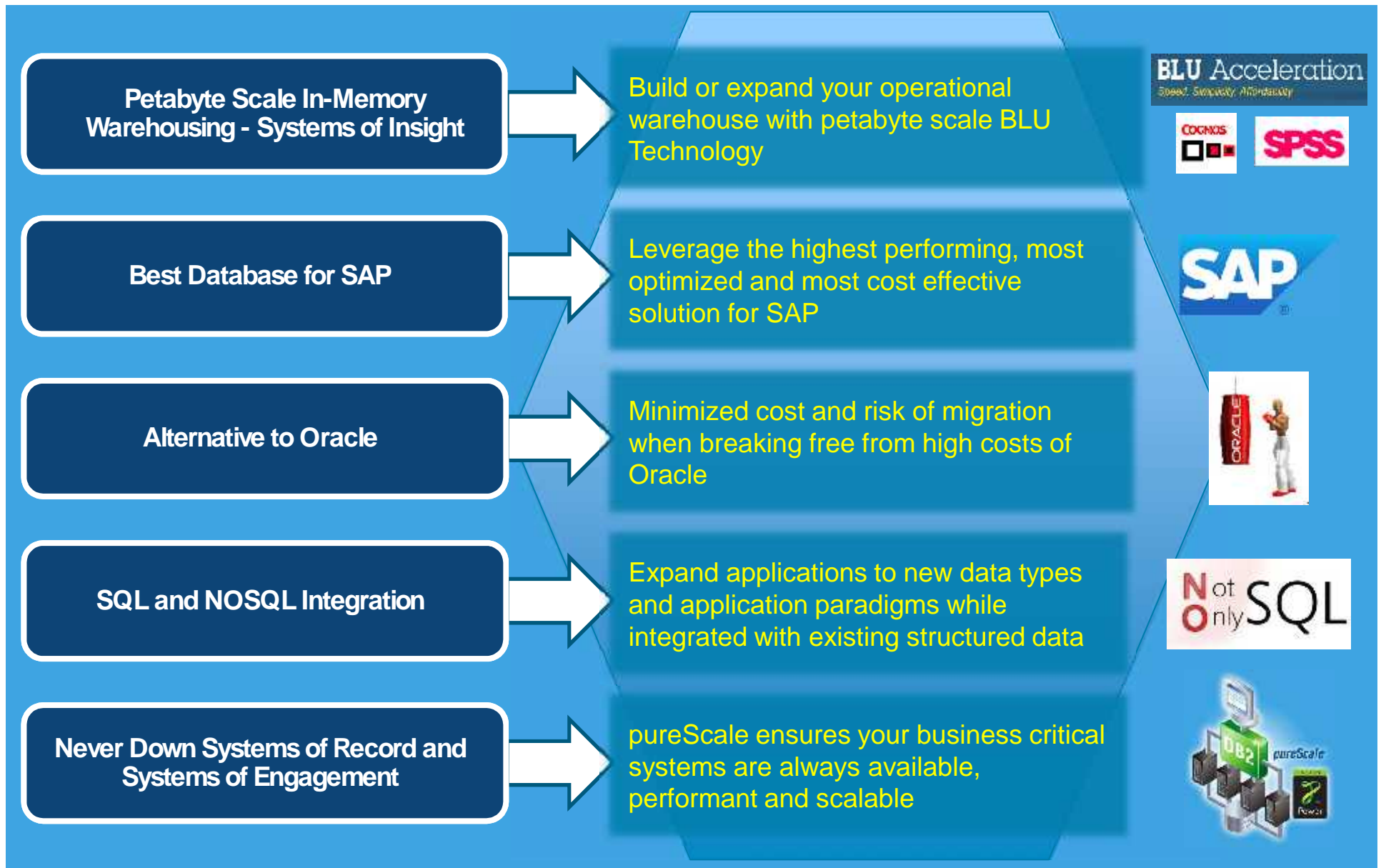
New Growth Trends

- ✓ Spark Integration
- ✓ HTAP Support
- ✓ SQL & NOSQL Capabilities
- ✓ Native JSON Support
- ✓ R Language Support
- ✓ Structured & Unstructured Data



Db2 V11.1.3.3

DB2 - Highlights and Strategic Investment Areas



Db2 Version 11.1.3.3 Highlights

Higher Availability and Core Capabilities



- Faster Rollback of very large transactions
- WLM – Improve deadlock detection
- HADR Resilience and SSL Encryption
- Db2iupdt – ADD/DROP CFs on-line
- pureScale – on-line CREATE INDEX w/R/W access to table
- pureScale – faster member crash recovery

Column-Organized (BLU) Tables

- UDF Caching for BLU
- BLU Memory Usage enhancements
- Temporal Query Support
- Index Support



Data Virtualization

- MariaDB Connectivity Support
- Db2 iSeries 7.2&7.3 Connectivity Support
- Teradata 16 Connectivity Support
- JSON over RESTful Service (MongoDB)
- Boolean, Binary/Varbinary Data Type Mapping Enhancement
- Pushdown Improvement for Hadoop Datasource
- Function Mapping Pushdown Enhancement

Additional Operating System Support

Solaris Support – 11.3+

Packaging Changes

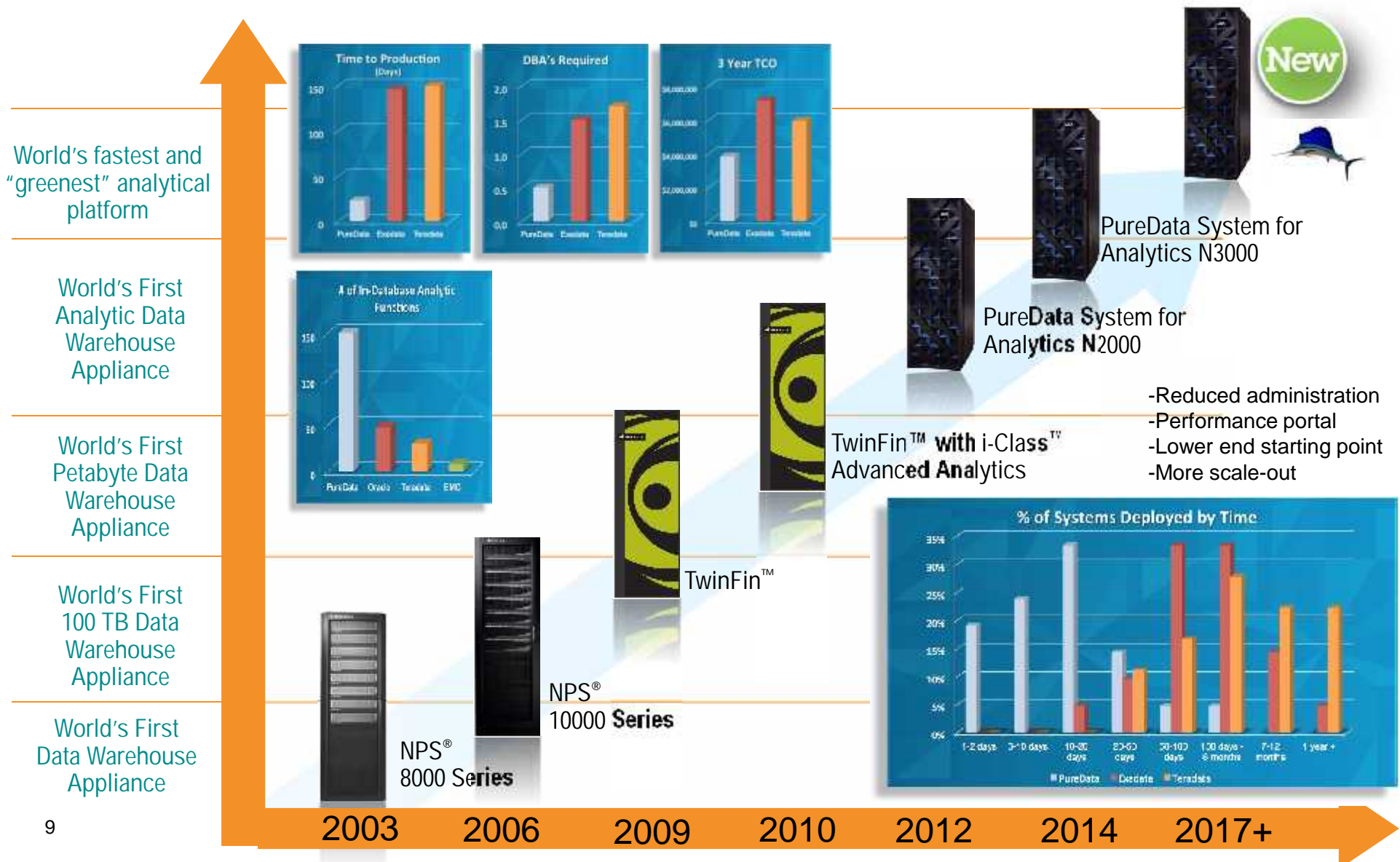
- Hybrid Data Management Packaging



NEW

IBM Integrated Analytics System

Next Generation Analytics Appliance – Maintain Core Values



IBM Integrated Analytics System

Next Generation Hybrid Data Warehouse

Optimized for **high performance** to support the broadest array of workload options for structured and unstructured data in your **hybrid data management** infrastructures

Real time analytics with **machine learning** that accelerates decision making, bringing new opportunities to the business – ready for **business analysts** and **data scientists**

Cloud-ready to support multiple workload deployment options



Reliable, elastic and flexible system that reduces and **simplifies management** resources

Leverages a **Common SQL Engine** for workload portability and skill sharing across public and private cloud

Built-in **IBM Data Science Experience** to collaboratively analyze data

IIAS - Addressing Top Customer Requirements

Broader set of workloads

- Combination of reporting, analytics, operational analytics and data stores

Higher Concurrency

- Expand number of business analytics and machine learning activities within a single system

In-Place Expansion

- Independently scale both compute and storage as needed while protecting existing investments

Richer Availability Solutions

- High Availability, Disaster Recovery and replication solutions



Optimized Analytics Performance

Next Generation In-Memory

In-memory columnar processing with dynamic movement of data from storage



Analyze Compressed Data

Patented compression technique that preserves order so data can be used without decompressing



Embedded Spark

Spark As an Analytics Engine



Spark/R, Spark/ML, Rest API, Object Store ETL, Complex Transformations (ELT), Streaming

Integrated Flash Storage

Hardware Accelerated architecture enabling faster insights with extreme performance, 99.999% reliability and operational efficiency



CPU Acceleration

Multi-core and SIMD parallelism (Single Instruction Multiple Data)



Data Skipping

Skips unnecessary processing of irrelevant data



Powered by Hardware

Designed for Deep Complex Analytics



4X Threads per core
4X Memory Bandwidth
4X More cache at Lower Latency

Multi Temperature



Most frequently accessed data on "hot" storage tier
Less frequently accessed data on "cold" storage tier

Flexible - Expansion Capabilities

Non-disruptive in-place incremental expansion

- Reduce disruptions to your analytics systems as you scale out

Cloud-ready

- Tools to move workloads seamlessly to the cloud based on your requirements

Non-disruptive in-place tiered storage expansion

- Independently scale storage for cost effective capacity growth

Cost efficient multi-temperature storage

- Most frequently accessed data (“hot”) on faster flash storage
- Less frequently accessed data (“colder”) on cost efficient enterprise storage systems



IBM Integrated Analytics System - Configurations



- IBM Power 8 S822L 24 core server 3.02GHz
IBM FlashSystem 900
- In-place Expansion Tiered storage
- Mellanox 10G Ethernet switches
Brocade SAN switches

	M4001-003 1/3 Rack	M4001-006 2/3 Rack	M4001-010 Full Rack	M4001-020 2 Racks	M4001-040 4 Racks	M4001-080 8 Racks
Servers	3	5	7	14	28	56
Cores	72	120	168	336	672	1344
Memory	1.5 TB	2.5 TB	3.5 TB	7 TB	14 TB	28 TB
User capacity (Assumes 4x compression)	64 TB	128 TB	192 TB	384	768	1536
Tiered storage (Optional)	TBD—GA 1H 2018					
2 Racks + Tiered Storage targeted for 1H 2018; In place expansion targeted for 2H 2018						



NEW

IBM Queryplex



BETA – TECHNOLOGY PREVIEW

Analytics Today...



- Costly and Complex
- High Latency to copy and synchronize
- Available compute resources under-utilized
- Error prone and difficult to retain data integrity

IBM Queryplex

An emerging technology now in beta trial

1

Query anything, anywhere.

Query **many diverse data sources** across cloud, on-premise and mobile with advanced analytics using the most popular languages and tool

SQL, Spark, R, Notebooks, Python, Data Science Experience (DSX), Cognos Analytics, common Analytics tools



2

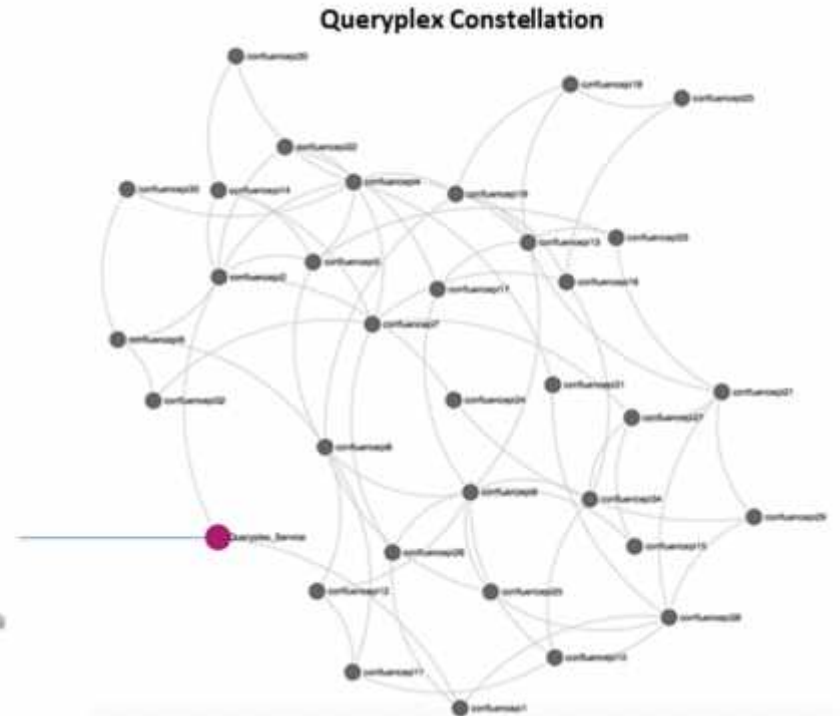
Query many sources as one with extreme simplicity.

Connect **few to many devices and data stores** into a single self balancing constellation. Avoid the complexity of centralized copies. Data only persists at the source.

3

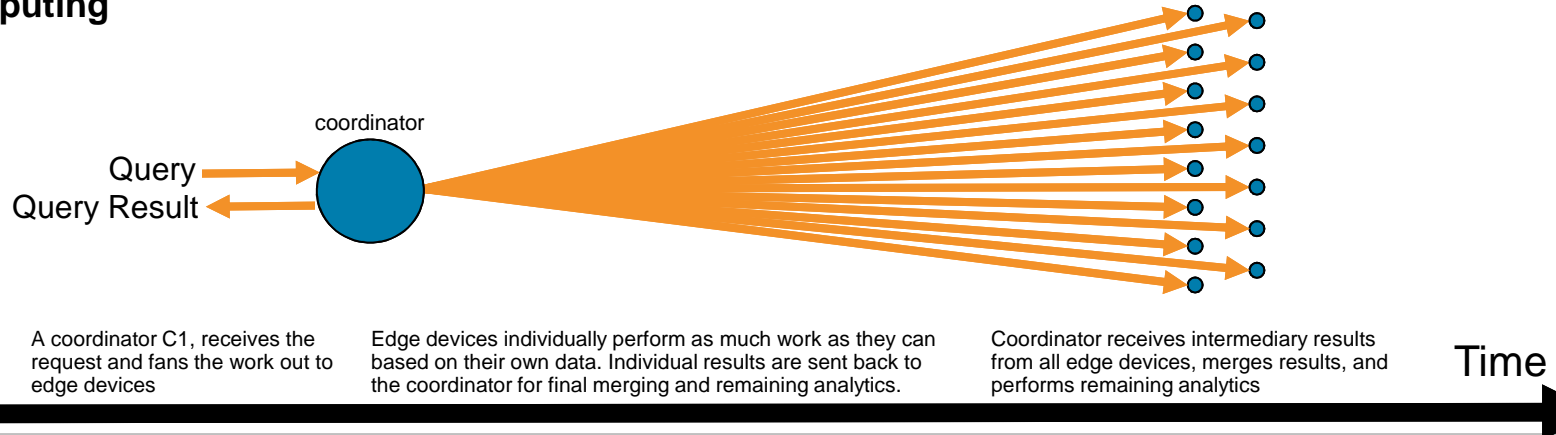
Massive speedup.

Many times acceleration using the power of every device.

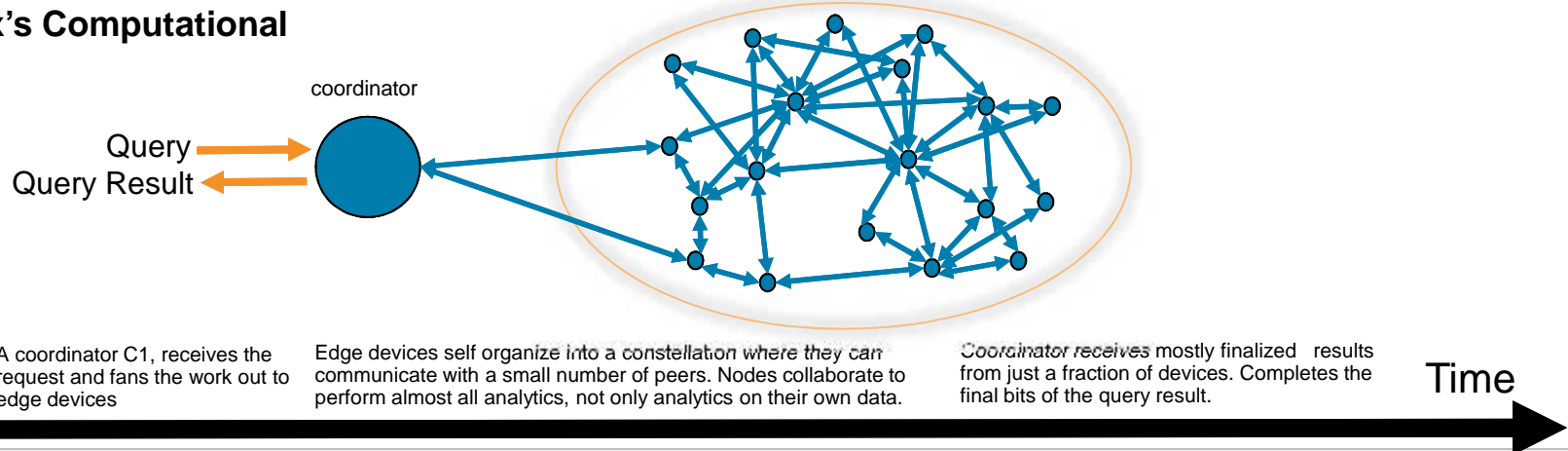


IBM Queryplex's Computational Mesh

Edge Computing



Queryplex's Computational Mesh



IBM Queryplex - Supported Languages & Data Sources

Query Languages	
SQL (ANSI)	✓
SQL (Oracle)	✓
SQL (DB2)	✓
SQL (PostgreSQL, Netezza)	✓
Scala	✓
PL/SQL	<i>Future</i>
SQL PL	<i>Future</i>
PySpark	✓
Python	✓
R & SparkR	✓

Mix Any Combination of Data Sources			
Oracle	✓	Excel	✓
DB2	✓	CSV (delimited text)	✓
Netezza	✓	MongoDB	✓
PostgreSQL	✓	Accumulo	<i>Future</i>
Informix	✓	Redis	<i>Future</i>
MySQL	✓	Cloudant	<i>Future</i>
SQLServer	✓		
DerbyDB	✓		

IBM Queryplex – Interested in hearing more ?

IBM Queryplex
The power of many together

<http://queryplex.com>



Db2 Big SQL 5.0.3

IBM Big Data High Value with Hortonworks

IBM's Offerings Unlock the value of Hadoop Data

- IBM BigIntegrate / BigQuality / BigMatch**
- Large scale data ingest & transformation
 - Data analysis, cleansing, & monitoring
 - Accurate linkage of customer data

#1 Data Science Platform: DSX

- Community and social features to provide collaboration
- The best of open source and IBM value-add to create state-of-the-art data products
- Built-in learning and advanced tutorials

IBM Information Governance Catalog

- Understand, Curate, and Govern
- Business level glossary and Catalog
- Comprehensive data lineage and tool impact analysis

#1 SQL Engine for Hadoop: Big SQL

- Data virtualization layer
- Large data volume, extremely complex query support
- Supports low latency, high concurrency workloads

Cognos, Watson Analytics

- Self service analytics capabilities
- Guided Analytics Discovery
- Natural Language Dialogue

IBM Big Replicate / IBM Data Replication

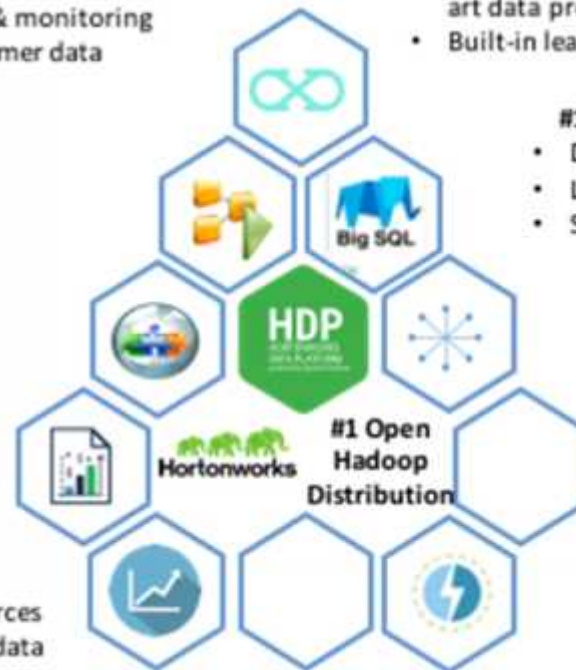
- Multiple Hadoop distributions to Hadoop
- Source Application to Hadoop Replication
- Provides HA/DR, with virtually zero RTO/RPO
- On-Prem to Cloud and Cloud to On-Prem

SPSS

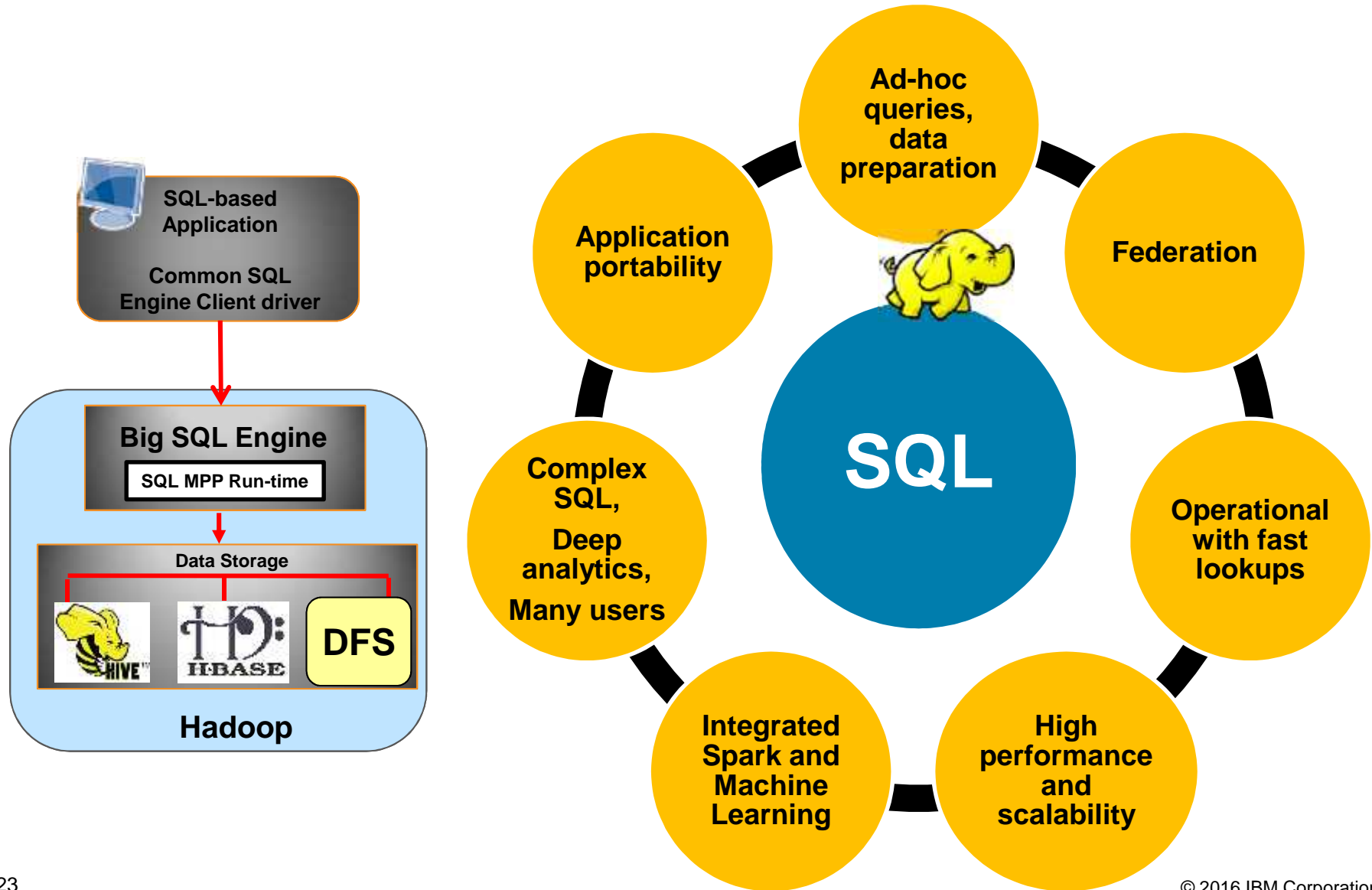
- Further embrace and extend Open source
- Integrate with other IBM offerings & data sources
- Energize your Analytics (text analytics for Big data on System-T)

IBM Streams

- Built-in streaming analytics
- Open architecture. Built for Speed
- Integrated Dev Environment



Db2 Big SQL – For all WH Needs in Hadoop



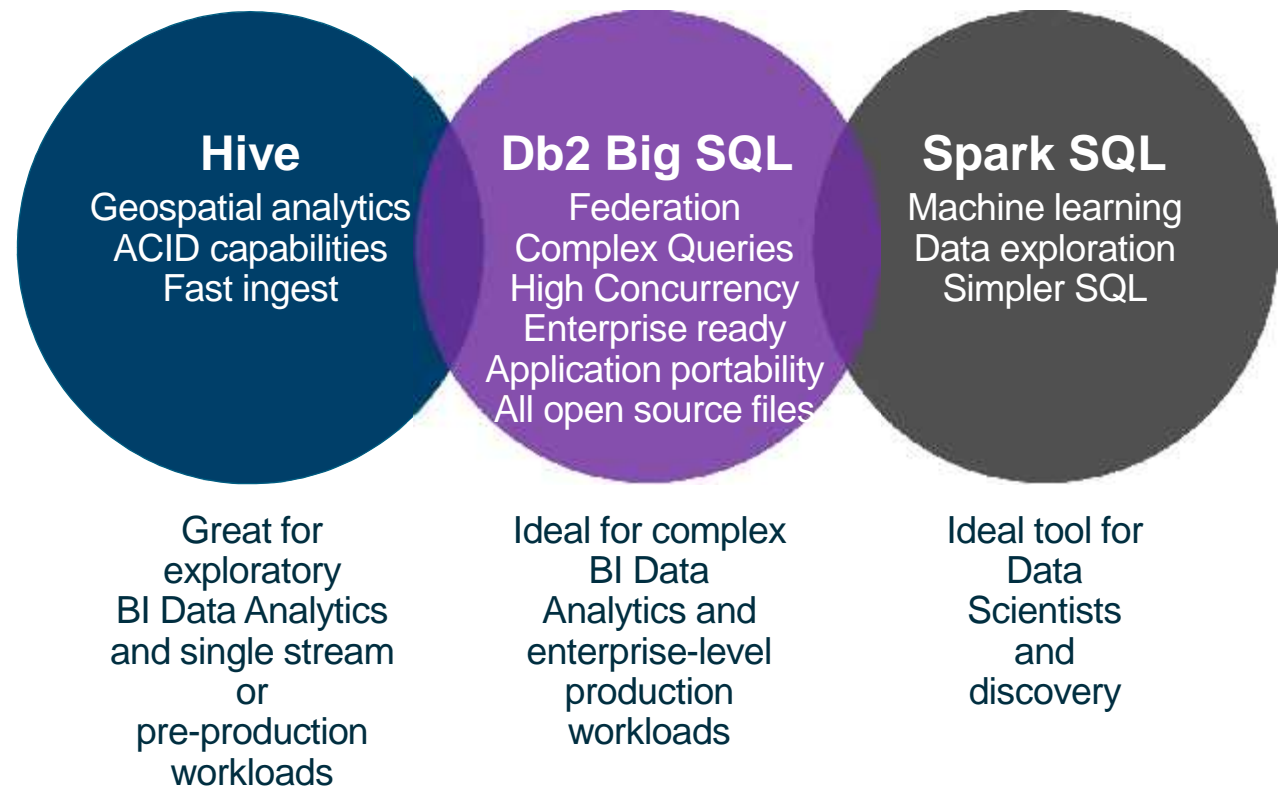
Db2 Big SQL V5.0.3

Applications	<ul style="list-style-type: none"> •ETL •Reporting •Data mining •Deep analytics 	<ul style="list-style-type: none"> •Reporting •Complex queries •BI Tools: Cognos, Tableau, etc 	<ul style="list-style-type: none"> •Query EDW •Join data •Use ML 	<ul style="list-style-type: none"> •Reuse applications •Reuse skills 	<ul style="list-style-type: none"> •Ad-hoc, exploratory •BI tools: Cognos, Tableau, etc
	Batch SQL (minutes to hours)	Interactive SQL (seconds to minutes)	Data augmentation (Spark integration)	Application portability	Self-service / Interactive BI (Sub-second)
Capabilities	SQL compatibility – Db2, Oracle, Netezza	SQL and NoSQL Structured & Unstructured	DSM, Ambari	MQTs	Ranger
	Advanced cost-based optimizer	Federation	Automatic memory management	Elastic boost – logical worker nodes	Roles
	Comprehensive ANSI SQL coverage	Spark Integration	Automatic workload management WLM	Query rewrite for optimized execution	SQL based RBAC
Core	Core SQL Engine	Integration	Administration	Performance	Security

www.tpc.org – check out TPC-H and TPC-DS – Big SQL vs Impala vs Hive
 Db2 Big SQL 5.0 is **2X** faster than Hive LLAP with Tez – and much more functional
 Db2 Big SQL 5.0 is **3X** faster than Spark SQL 2.1

Combining Hadoop Technologies

Not Mutually Exclusive.
 Hive, Db2 Big SQL &
 Spark SQL can co-exist
 and complement each
 other in a cluster





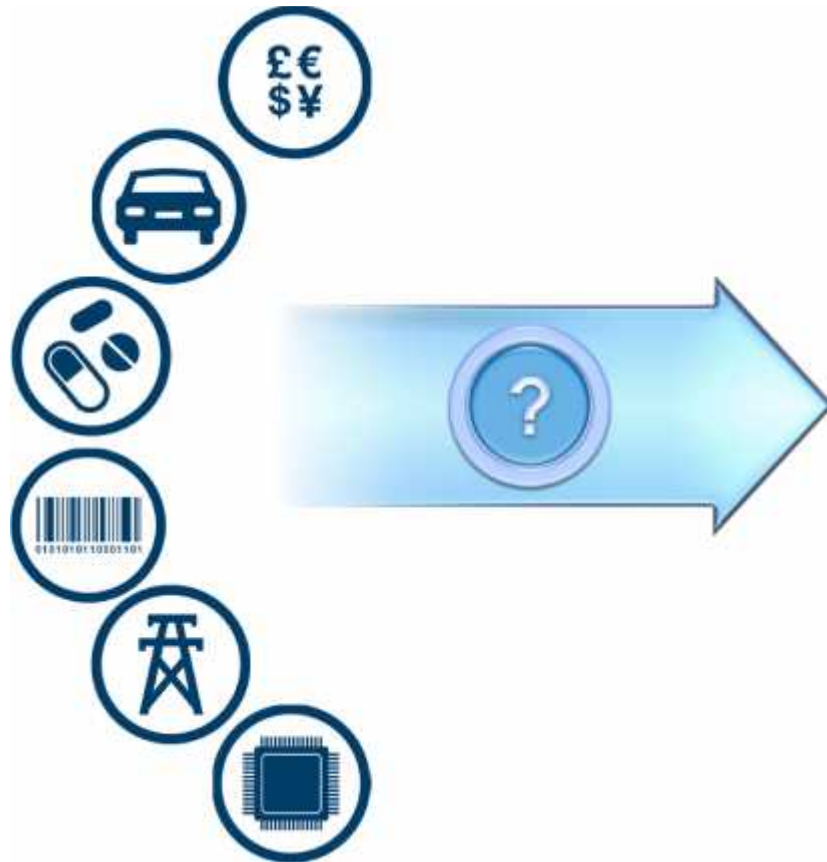
NEW

Db2 Event Store



ON-PREMISES TODAY
CLOUD COMING

Event-Driven Systems Span Many Industries



- 

Multi-channel customer sentiment and experience analysis
- 

Detect life-threatening conditions at hospitals in time to intervene
- 

Predict weather patterns to plan optimal wind turbine usage, and optimize capital expenditure on asset placement
- 

Make risk decisions based on real-time transactional data
- 

Identify criminals and threats from disparate video, audio, and data feeds

Event Processing Workloads

A unified offering for Fast Data which delivers...



1 Lightning Fast Ingest

- 1 Million inserts per second per node
- Ingest scales linearly with added nodes
- Data ingested quickly, then refined and enriched



2 Real-time Analytics

- Real-time analytics over ALL ingested data
- Super-fast lookups and intelligent scans
- Integrated machine learning capabilities



3 Integrated and Highly Available

- Packaged and integrated with IBM Data Science experience; available Streams sink
- Remains available on node failure
- Architected to scale to very large clusters



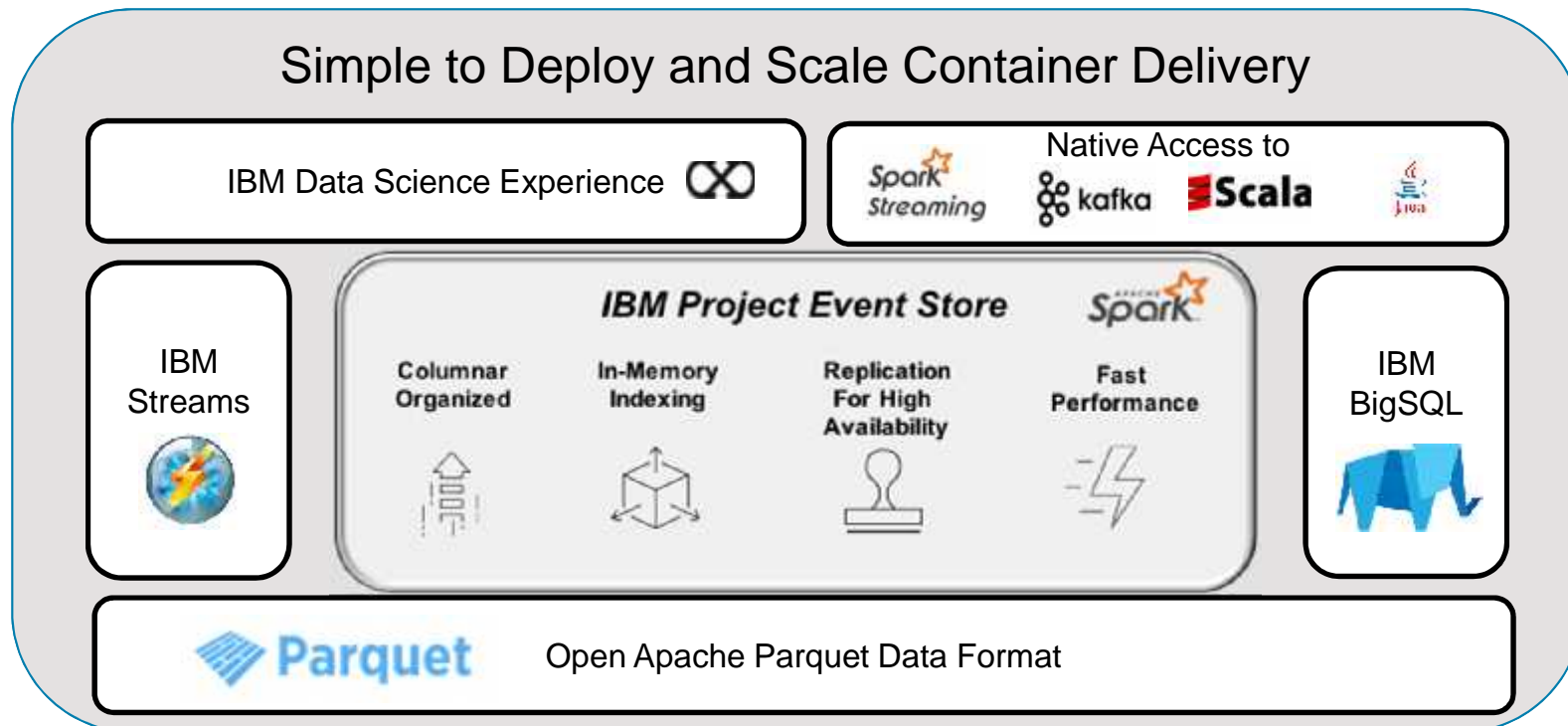
4 Built for Data Sharing and Efficiency

- Writes to shared storage in Parquet format
- Able to leverage low-cost object storage
- Single copy of the data

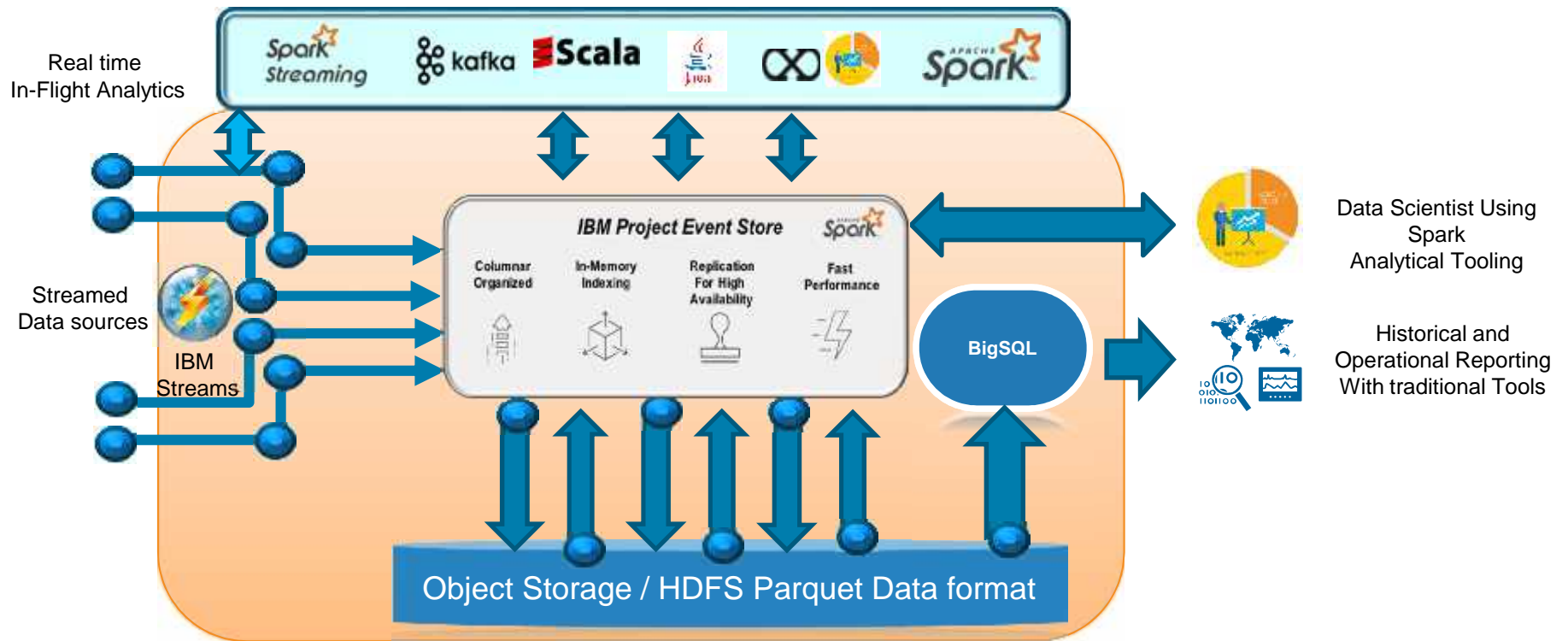


Db2 Event Store

Integrated System for Managing Events



Db2 Event Store











IBM Cloud Private



















Db2 and the Cloud

Provisioning
& Db2 Setup

Management

Maintenance

 <p>“Bring Your Own License”</p>	<ul style="list-style-type: none"> • Custom-deployable software on your own infrastructure or private cloud or public cloud • Fully customizable for any type of workload • Complete flexibility including DPF and pureScale * • Customer managed
 <p>Db2 Hosted</p>	<ul style="list-style-type: none"> • Hosted database-as-a-service • Pre-defined hardware configurations • Fully customizable for any type of workload • Available on SoftLayer and AWS • Customer managed
 <p>Db2 on Cloud</p>	<ul style="list-style-type: none"> • Fully managed database-as-a-service • Pre-defined and flexible hardware configurations optimized for transactional and general purpose workloads • Available on Bluemix public cloud
 <p>Db2 Warehouse on Cloud</p>	<ul style="list-style-type: none"> • Fully managed database-as-a-service • Pre-defined hardware configurations optimized for analytics workloads • In-database analytics • Available on Bluemix and AWS public cloud
 <p>Db2 Warehouse</p>	<ul style="list-style-type: none"> • Deploy on your own infrastructure or private cloud • Docker container technology for fast and simple deployment • Optimized for analytic workloads • Scalable, elastic • Customer managed
 <p>Db2 OLTP</p>	<ul style="list-style-type: none"> • Deploy on your own infrastructure or private cloud • Docker container technology for fast and simple deployment • Optimized for operational and OLTP workloads • Scalable, elastic • Customer managed


		
		
		
		
		
		

Db2 and the Cloud

Provisioning
& Db2 Setup


Management

Maintenance




“Bring Your Own License”

- Custom-deployable software on your own infrastructure or [private cloud](#) or [public cloud](#)
- Fully customizable for any type of workload
- Complete flexibility including DPF and pureScale *
- [Customer managed](#)



Db2 Hosted

- [Hosted](#) database-as-a-service
- Pre-defined hardware configurations
- Fully customizable for any type of workload
- Available on SoftLayer and AWS
- [Customer managed](#)




Db2 on Cloud

- [Fully managed](#) database-as-a-service
- Pre-defined and flexible hardware configurations optimized for transactional and general purpose workloads
- Available on Bluemix [public cloud](#)




Db2 Warehouse on Cloud

- [Fully managed](#) database-as-a-service
- Pre-defined hardware configurations optimized for analytics workloads
- In-database analytics
- Available on Bluemix and AWS [public cloud](#)





















Db2 Warehouse

- Deploy on your own infrastructure or [private cloud](#)
- Docker container technology for fast and simple deployment
- Optimized for analytic workloads
- Scalable, elastic
- [Customer managed](#)



Db2 OLTP

- Deploy on your own infrastructure or [private cloud](#)
- Docker container technology for fast and simple deployment
- Optimized for operational and OLTP workloads
- Scalable, elastic
- [Customer managed](#)

Introducing IBM Cloud Private



Innovation

Kubernetes-based container platform

Cloud Foundry for prescribed container-based application development and deployment and life cycle management

Integrated DevOps toolchain



Integration

Catalog of integration services

API availability and management to integrate applications and data across environments



Investment Protection

Prescriptive guidance on where to run and how to architect your critical workloads

Next generation versions of industry leading IBM Middleware and Analytics
(MQ, Db2, Data Science, Cognos, Blockchain, IIB)



Management and Compliance

Core operational services, including monitoring, log mgmt, and security

Integration with existing systems and operations management solutions

Analytics Roadmap : Offerings / Capabilities on ICp

Available on ICp

- Db2 OLTP
- Db2 Warehouse
- Db2 Warehouse MPP
- Db2 Big SQL
- Db2 Event Store
- Data Server Manager

And also

- Data Science Experience
- Data Stage
- InfoSphere Governance
Catalog

And more



IBM Cloud Private for Data

In Summary – Why Analytics on IBM Cloud Private

True Hybrid Solution - consistency between public cloud and private cloud

No vendor lock-in. **Open Platform as a Service** (PaaS) for maximum integration ability

Container-based platform with very fast time to value (hours instead of weeks)

Extensive service-oriented analytic and machine learning capabilities ready for **Data Scientists** and **Business Analysts**

Optimized and secure **Data Management Services** for SQL, NoSQL, structured, semi-structured and unstructured data

Secure, governed and compliant platform for integration with any data source

IBM Cloud Private – Next Steps



Learn more

- ✓ ICP Product page: <http://ibm.biz/IBMCloudPrivate>
- ✓ ICP on Power DeveloperWorks page: <http://ibm.biz/ICP-Power-TechnicalCommunity>
- ✓ Linux on Power Development portal: <https://developer.ibm.com/linuxonpower>
- ✓ ICP Technical Community: <http://ibm.biz/ICP-TechnicalCommunity>
- ✓ ICP Knowledge Center: <http://ibm.biz/ICP-KnowledgeCenter>
- ✓ Introduction to ICP (video): https://www.youtube.com/watch?v=UL_jXJoRPdY
- ✓ ICP Overview (video): <https://www.youtube.com/watch?v=yzXA3qhfaq0>
- ✓ ICP on IBM Power (video): <https://www.youtube.com/watch?v=73LpA1Cmqcc>

See it in action

- ✓ Try ICP Community Edition for free: <http://ibm.biz/ICP-SignUp>
- ✓ YouTube ICP tutorials play list: <http://ibm.biz/ICP-YouTubeTutorials>

Get help

- ✓ Join the #ibm-cloud-private public Slack channel: <http://ibm.biz/ICP-Slack>
- ✓ ICP on Stack Overflow: <https://stackoverflow.com/questions/tagged/ibm-cloud-private>

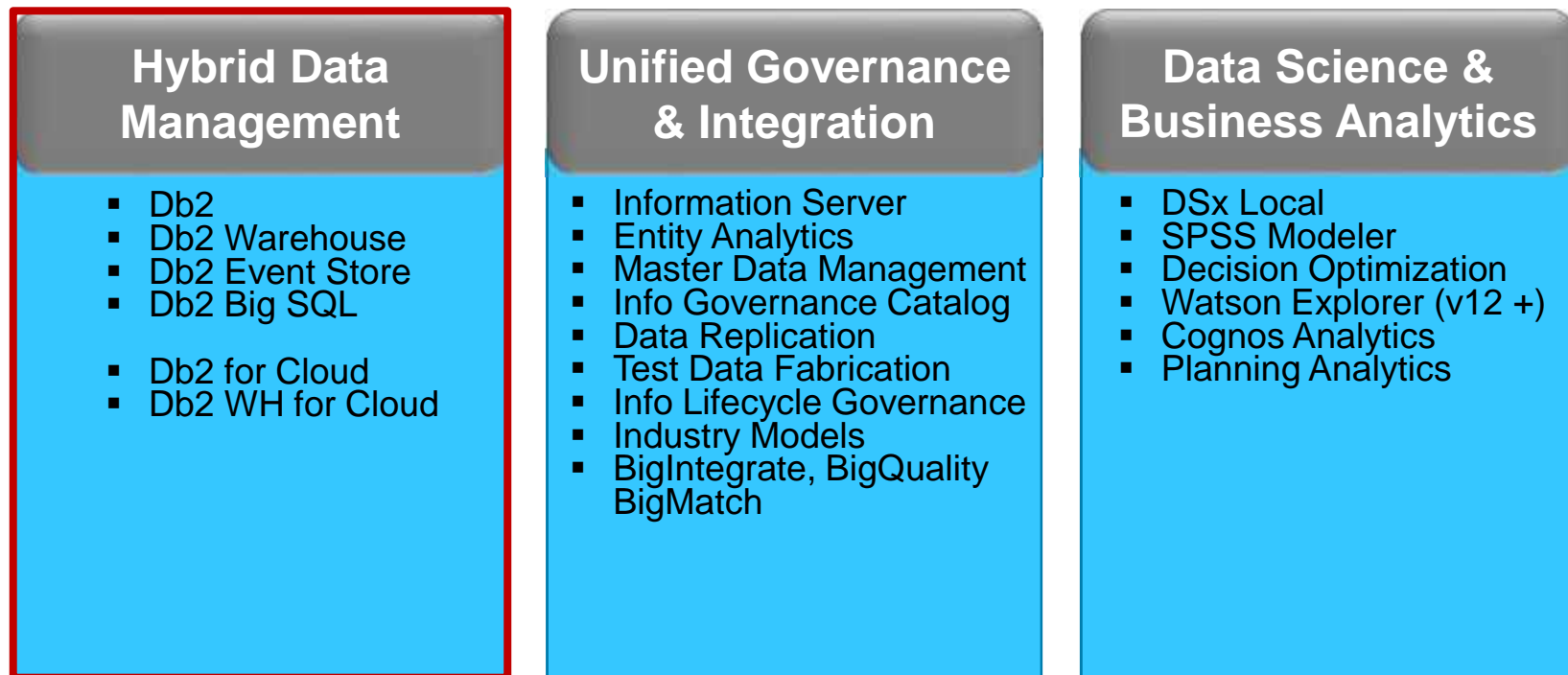


NEW

FlexPoints & HDM Offering

Portfolio Simplification:

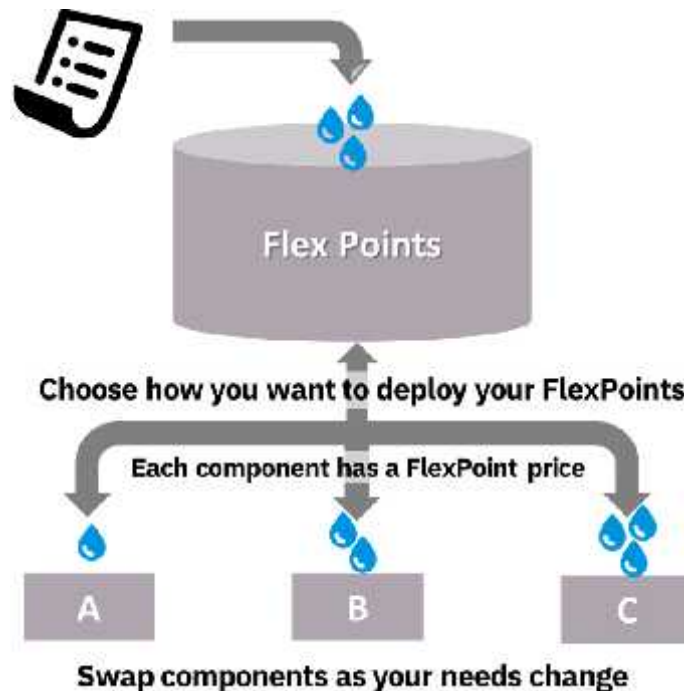
Three new bundles



We will now focus on Hybrid Data Management

FlexPoints: How It Works

Buy FlexPoint licenses for the “Platform of your Choice”



Platform Offerings deliver integrated capabilities – now offered as flex bundles to simplify planning for adoption and growth at the lowest cost

Available for Our 3 Platform Offerings:

- Hybrid Data Management
 - Db2
 - Db2 Warehouse
 - Db2 Event Store
 - Db2 Big SQL
- Unified Governance & Integration
- Data Science & Business Analytics

FlexPoints CANNOT be used across PLATFORMS

As an example, Data Science and Business Analytics FlexPoints are NOT valid for Hybrid Data Management

Les King

Director, Hybrid Data Management Solutions

September, 2018

lking@ca.ibm.com

ca.linkedin.com/pub/les-king/10/a68/426

Hybrid Data Management Strategy and New News !

