

THE FILLMORE GROUP

Relational Database Solutions



**IBM z Systems (Mainframes!) for IT Architects
and Data Scientists
The Fillmore Group – March 2016**

A Premier IBM Business Partner





Agenda

- ▶ Introduction
- ▶ Types of data on z Systems (z/OS)
 - ▶ DB2 for z/OS (relational database)
 - ▶ IMS – Information Management System (hierarchical database)
 - ▶ VSAM – Virtual Storage Access Method (indexed file)





Agenda

- ▶ Key Approaches to Unlocking the Data
 - ▶ HTAP – Hybrid Transaction / Analytics Processing
 - ▶ Replication to heterogeneous data stores
 - ▶ Virtualization of heterogeneous data stores
 - ▶ ETL – Extract, Transform, Load





History

The Fillmore Group, Inc.

- ▶ Founded in the US in Maryland, 1987
- ▶ IBM Business Partner since 1989
- ▶ Delivering IBM Education since 1994
- ▶ DB2 Gold Consultant since 1998
- ▶ IBM Champions since 2009





The Fillmore Group, Inc.

- ▶ IBM Analytics Technical Support and Consulting
- ▶ Authorized Training with IBM Global Training Partner
- ▶ IBM Analytics Software Reseller

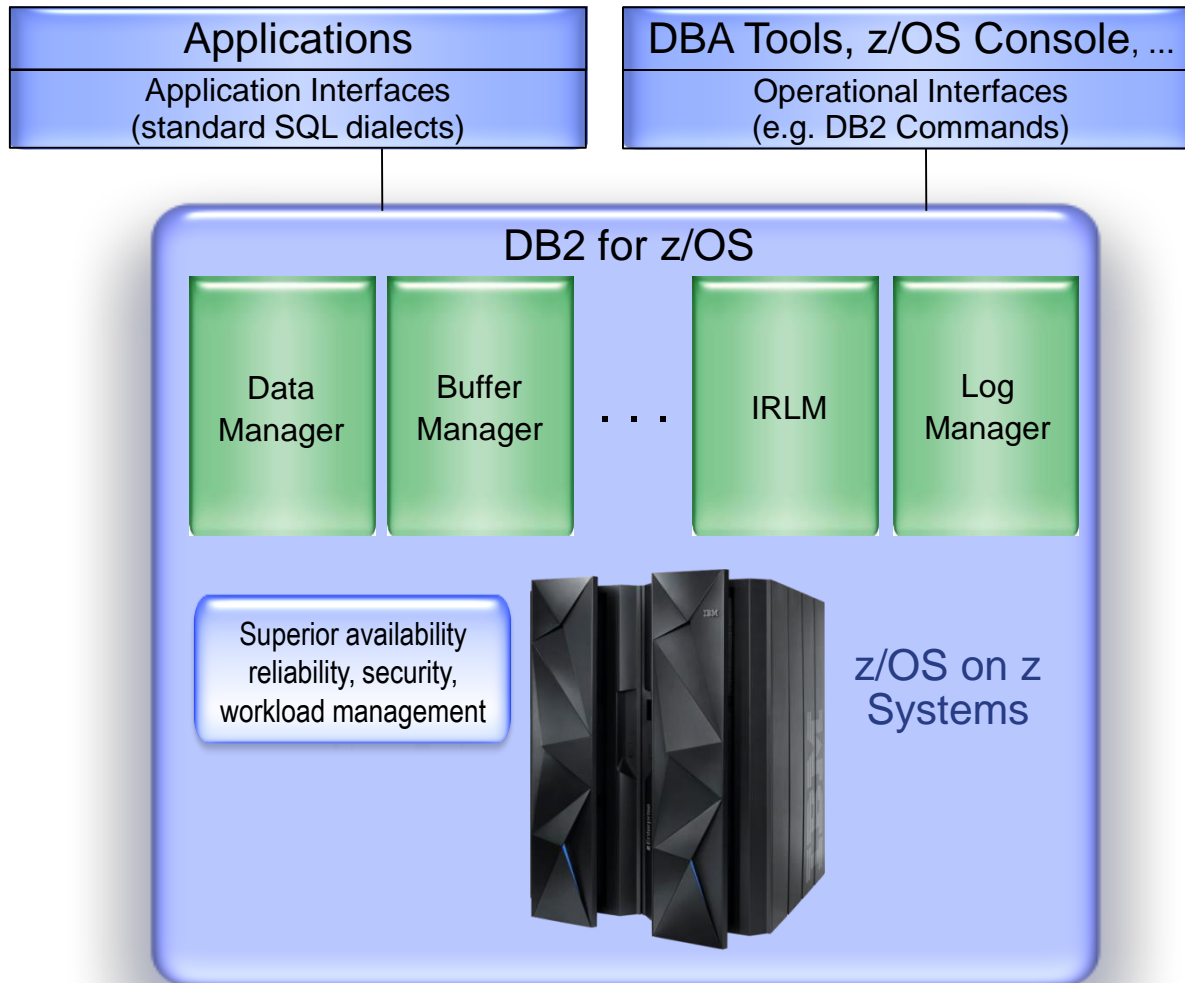




DB2 for z/OS v11

- ▶ Relational Database Management System (RDBMS)
- ▶ Primarily structured data formatted in tables using rows and columns
- ▶ Application access using Structured Query Language (SQL) ANSI X3.135-1999 (with extensions)
- ▶ B-tree indexes
- ▶ Cost-based Optimizer chooses access path







Access to DB2 for z/OS from Windows/Linux (EBCDIC to ASCII/Unicode Protocol Conversion)

- ▶ DB2 Connect
- ▶ IBM Data Server Driver for JDBC and SQLJ
- ▶ IBM Data Server Driver for ODBC and CLI

- ▶ DB2 Adapter for z/OS Connect
 - ▶ Mobile and Cloud connectivity
 - ▶ Packaged with WAS, CICS and IMS





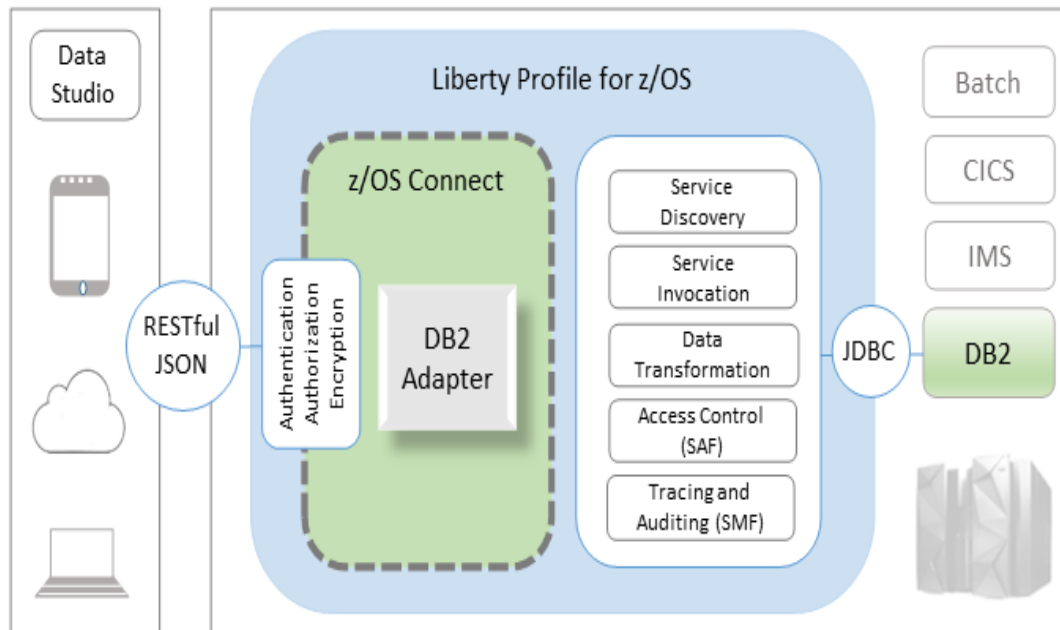
Access to DB2 for z/OS from Windows/Linux

- ▶ Scripting languages
 - ▶ PHP, Ruby on Rails, Python, Perl, Node.js
- ▶ Hibernate/iBatis Accelerators
- ▶ JavaScript Object Notation (JSON)
- ▶ Microsoft
 - ▶ .NET Data Provider, Visual Studio
- ▶ Application servers
 - ▶ WebSphere, Apache





DB2 Adapter for z/OS Connect





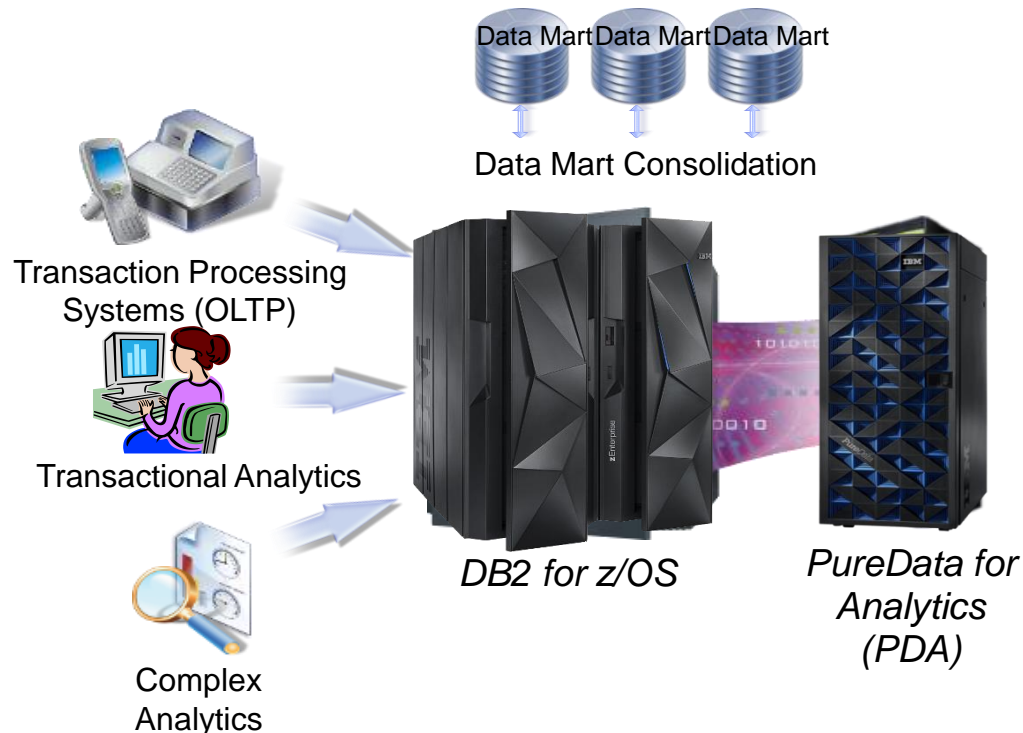
Real-time DB2 for z/OS Analytics Challenges

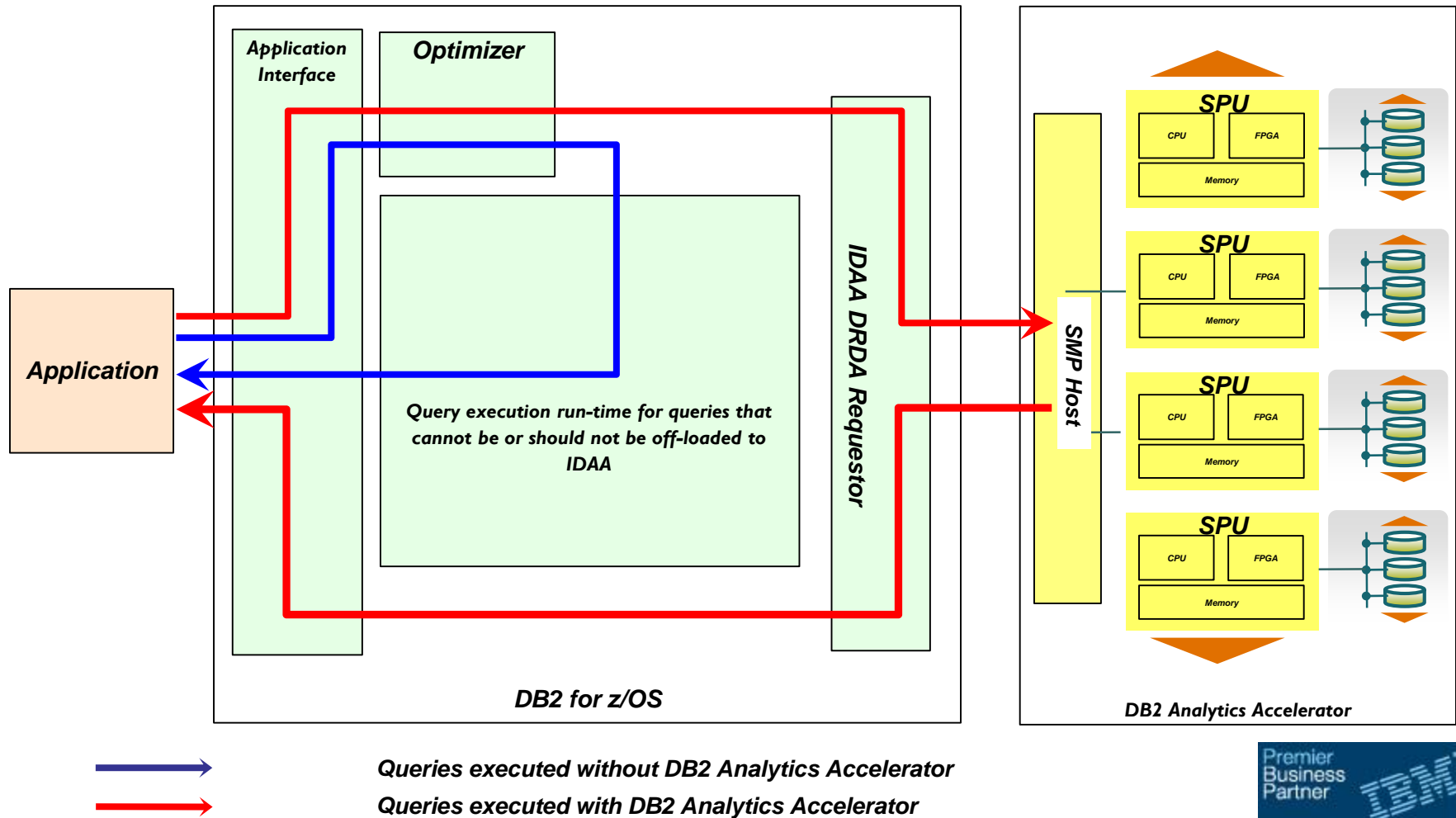
- ▶ Might interfere with online transaction processing (OLTP)
- ▶ B-tree index access degrades over billions of rows
- ▶ z Systems disk storage cost





Hybrid Transaction / Analytic Processing (HTAP) IBM DB2 Analytics Accelerator (IDAA)



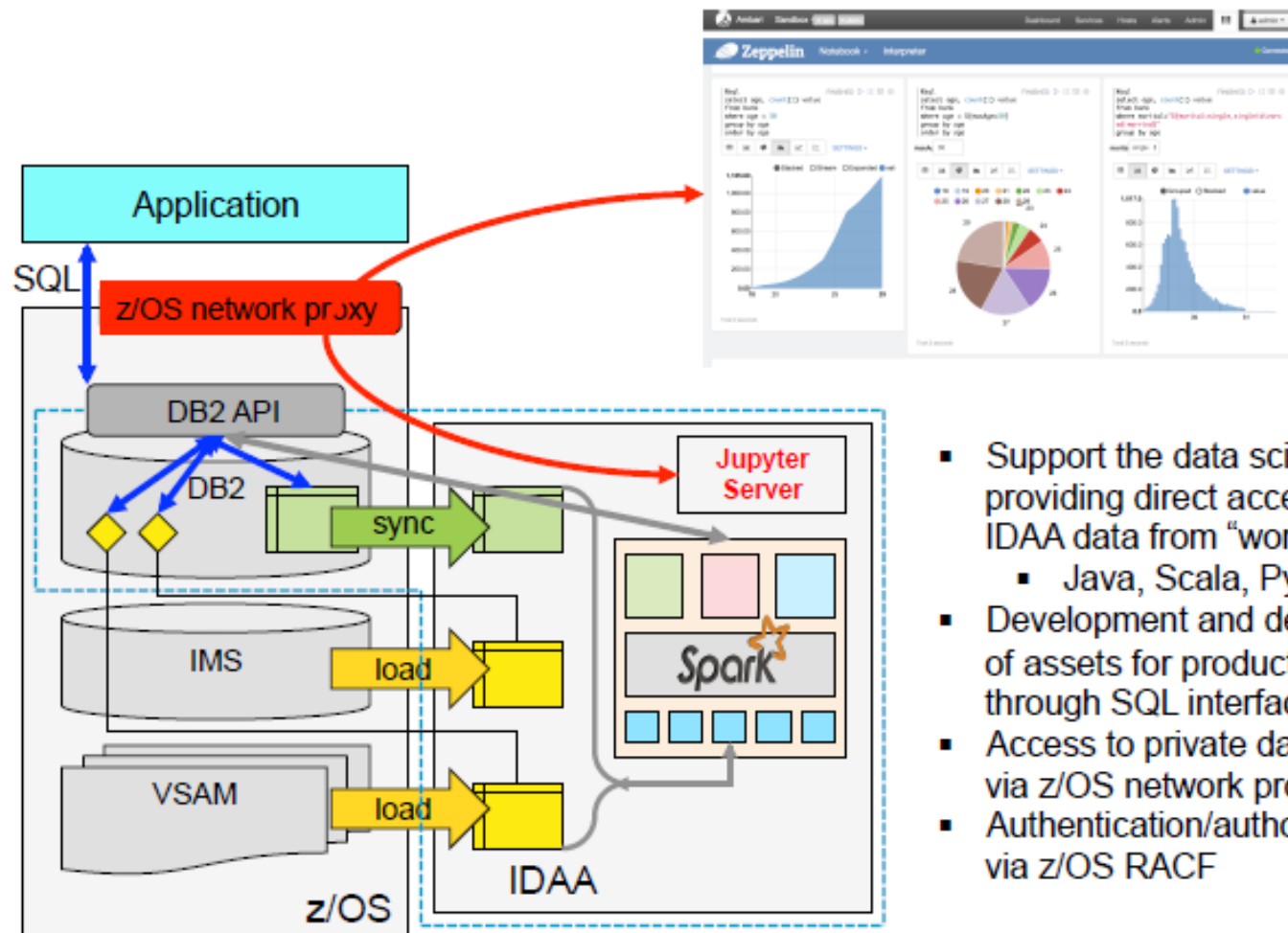




				DB2 Only		DB2 with IDAA		Times Faster
Query	Total Rows Reviewed	Total Qualifying Rows	Total Rows Returned	Hours	Sec(s)	Hours	Sec(s)	
Query 1	591,941,065	2,813,571	853,320	2:39	9,540	0.0	5	1,908
Query 2	591,941,065	2,813,571	585,780	2:16	8,220	0.0	5	1,644
Query 3	813,343,052	8,260,214	274	1:16	4,560	0.0	6	760
Query 4	283,105,125	2,813,571	601,197	1:08	4,080	0.0	5	816
Query 5	591,941,089	3,422,765	508	0:57	4,080	0.0	70	58
Query 6	813,343,052	4,290,648	165	0:53	3,180	0.0	6	530
Query 7	591,941,065	361,521	58,236	0:51	3,120	0.0	4	780
Query 8	813,343,052	3,425,292	724	0:44	2,640	0.0	2	1,320
Query 9	813,343,052	4,130,107	137	0:42	2,520	0.1	193	13



Non-SQL Access to Spark Running in IDAA



- Support the data scientist by providing direct access to IDAA data from “workbench”
 - Java, Scala, Python,...
- Development and deployment of assets for production use through SQL interfaces
- Access to private data network via z/OS network proxy only
- Authentication/authorization via z/OS RACF



IMS – Information Management System v14

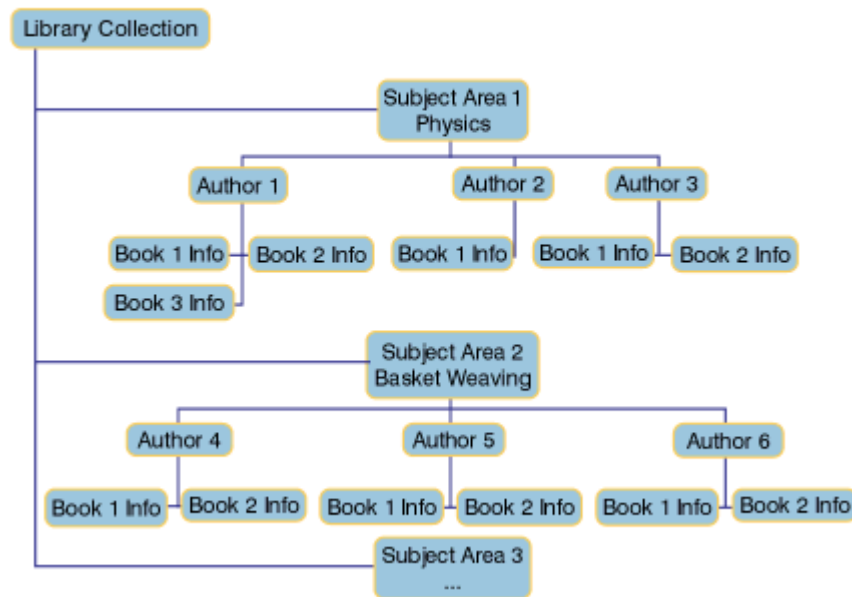
- ▶ Hierarchical data model
- ▶ Typically accessed via 3GL (e.g. Cobol, PL/I)
 - ▶ Imbedded Data Language/Interface (DL/I)
- ▶ Multiple fields comprise a “segment” or record type
- ▶ Applications navigate between segments using imbedded pointers





IMS – Information Management System

Hypothetical Hierarchical Database Model





Access to IMS

- ▶ Extract natively on z/OS
 - ▶ Custom (e.g. Cobol)
 - ▶ Utility (e.g. File-Aid)
- ▶ IMS Enterprise Suite
 - ▶ Microsoft .NET
 - ▶ Java
 - ▶ SOAP
- ▶ z/OS Connect





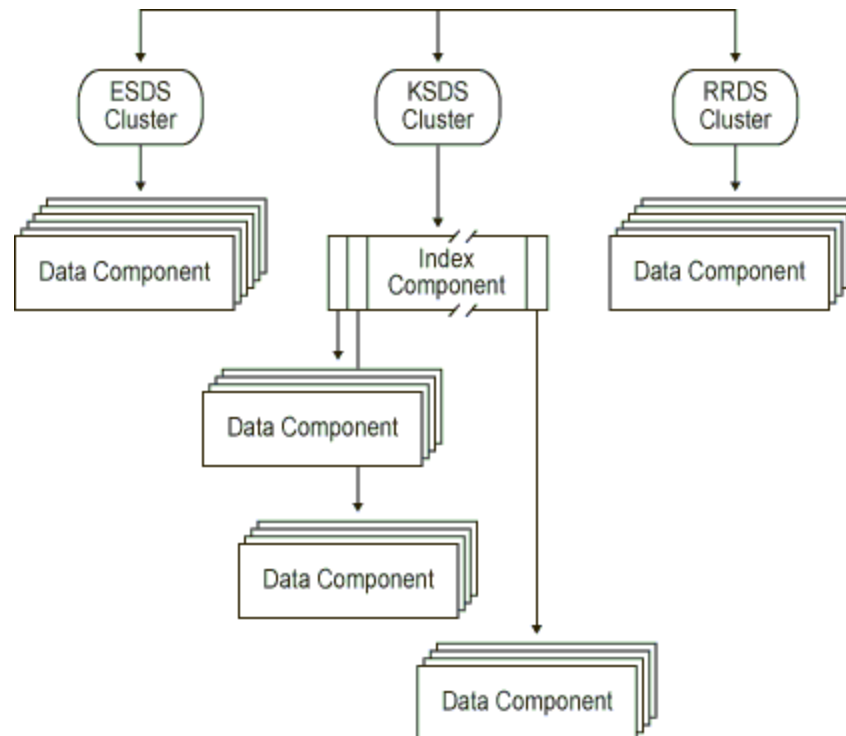
VSAM – Virtual Storage Access Method

- ▶ **Types**
 - ▶ Entry-sequenced Data Set (ESDS)
 - ▶ Key-sequenced Data Set (KSDS) - indexed
 - ▶ Relative-record Data Set (RRDS)
- ▶ Typically accessed via 3GL (e.g. Cobol, PL/I)
- ▶ Schema imbedded in application program
- ▶ Single file might have multiple schemas
- ▶ “Dirty” data





VSAM – Virtual Storage Access Method





Access to VSAM

- ▶ Extract natively on z/OS
 - ▶ Custom (e.g. Cobol)
 - ▶ Utility (e.g. IDCAMS REPRO)





Access to IMS and VSAM

- ▶ Replication
 - ▶ IBM InfoSphere Data Replication (IIDR)
- ▶ Virtualization
 - ▶ InfoSphere Classic Federation for z/OS (SQL)
 - ▶ Rocket Data Virtualization (SQL, NoSQL)
- ▶ HTAP - IDAA
 - ▶ DB2 Analytics Accelerator Loader v2.1





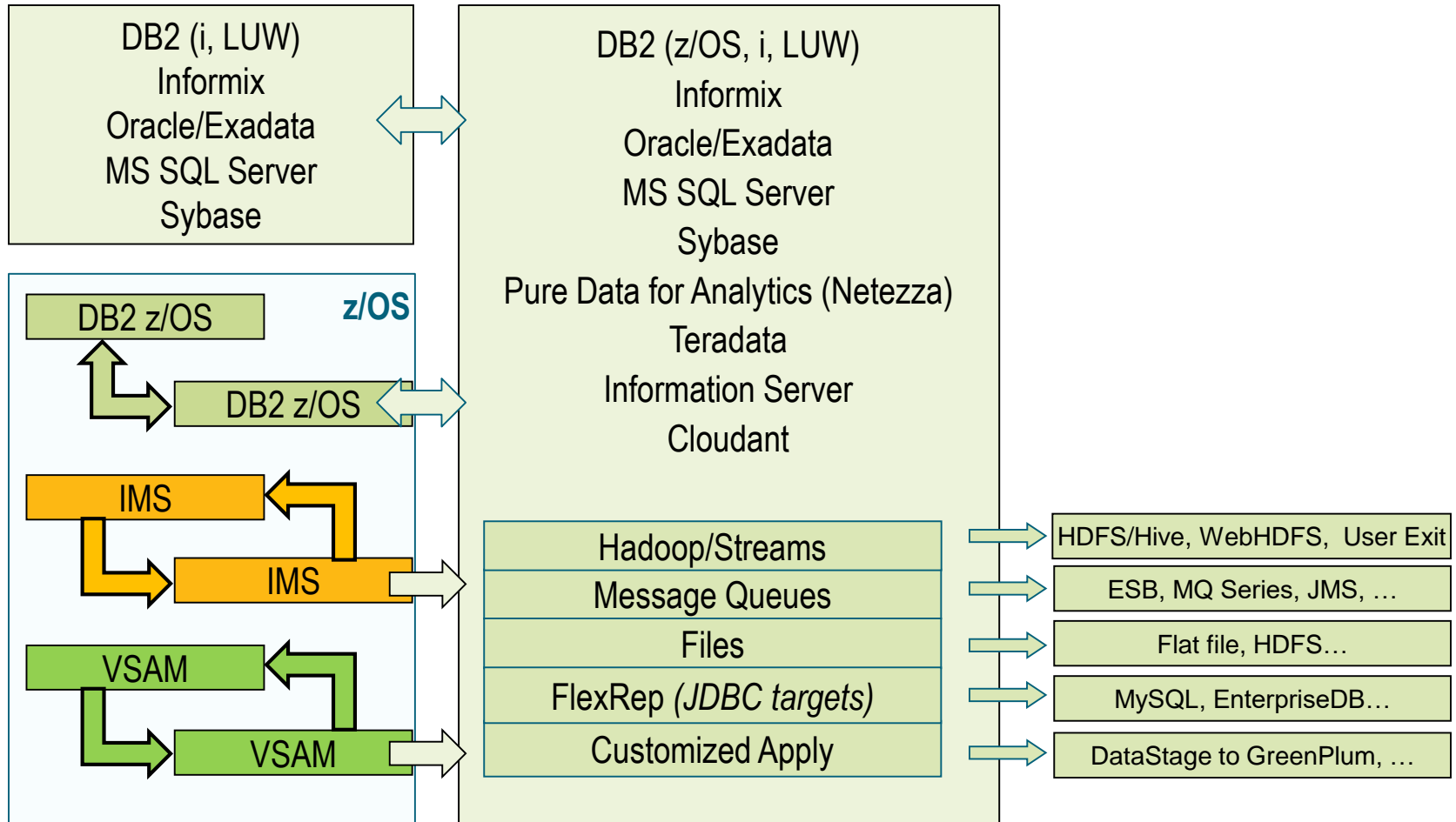
Access to IMS and VSAM

- ▶ Extract, Transform, Load (ETL)
 - ▶ InfoSphere Classic Connector for z/OS and InfoSphere Information Server (Datastage)
 - ▶ Third-party





IBM InfoSphere Data Replication (IIDR)



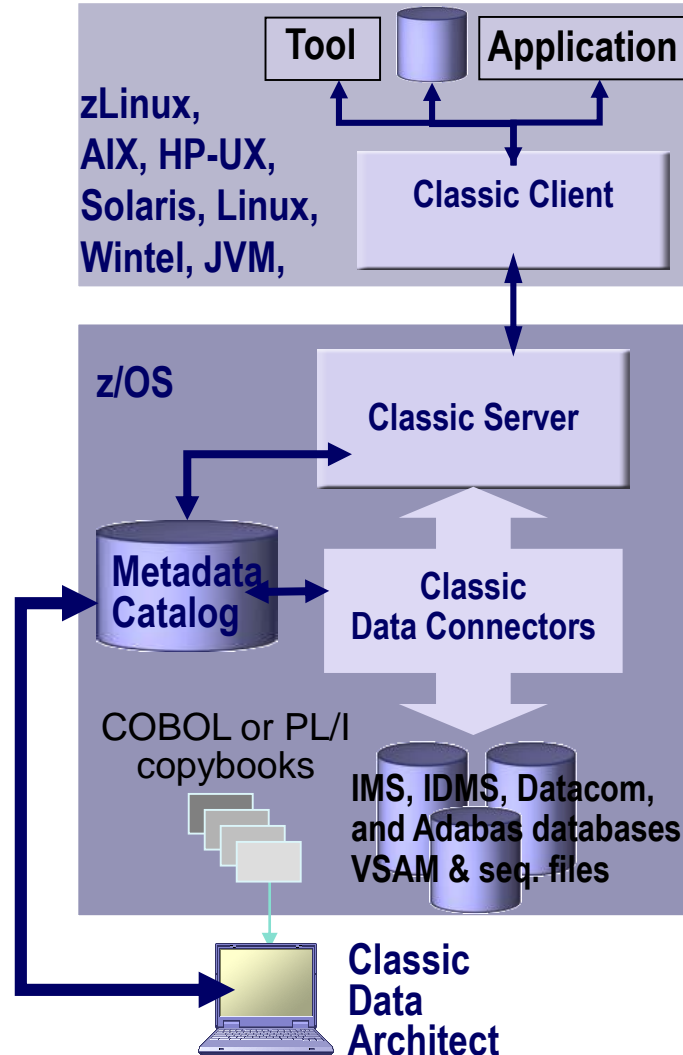


Replication

- ▶ Strengths
 - ▶ Typically log-based Capture
 - ▶ Isolate analytics from OLTP on a different platform
- ▶ Weaknesses
 - ▶ Data conversion issues (depending on the target)
 - ▶ EBCDIC to ASCII/Unicode protocol conversion
 - ▶ NULLs
 - ▶ Latency – data might be stale by seconds or minutes
 - ▶ Limited data transformations



InfoSphere Classic Federation for z/OS (Virtualization)





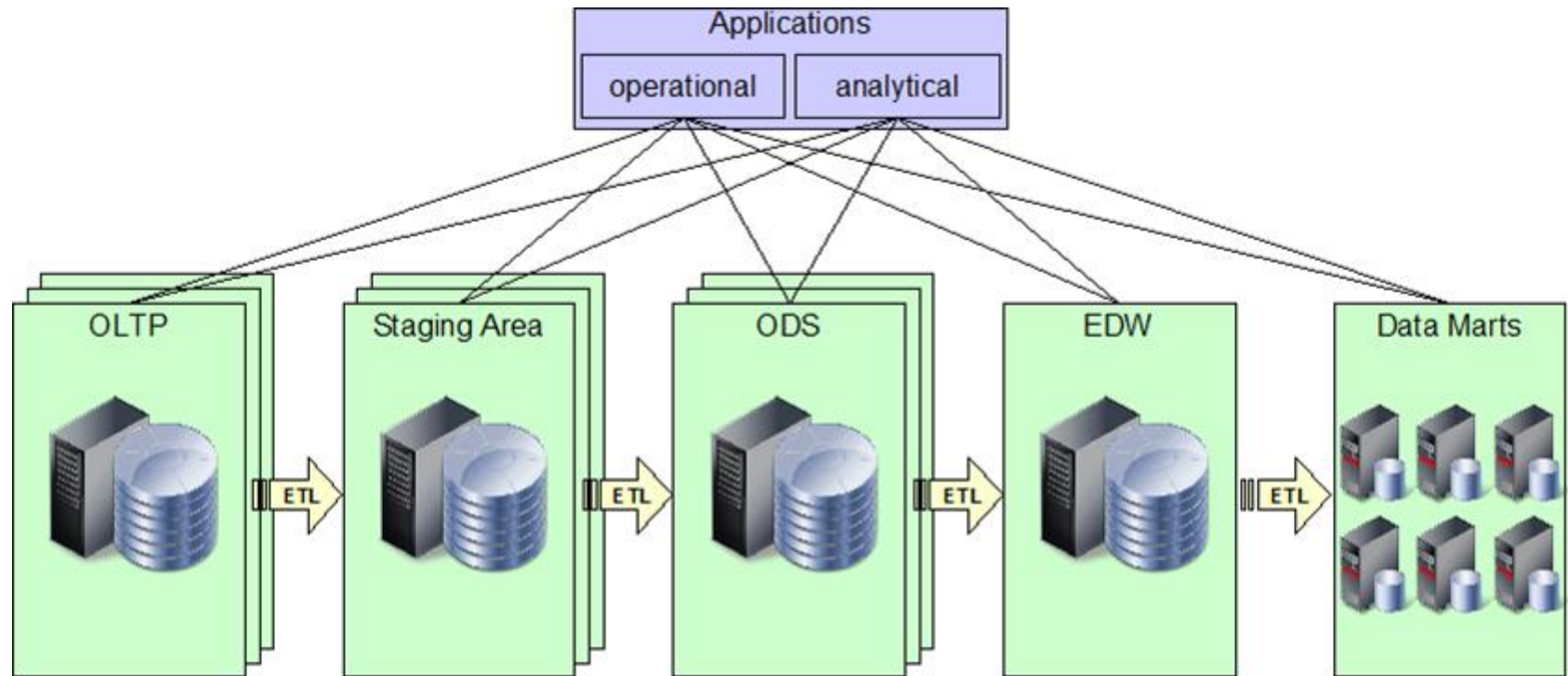
Virtualization (i.e. Federation)

- ▶ Strengths
 - ▶ Real-time access to system-of-record (no latency)
- ▶ Weaknesses
 - ▶ Might interfere with OLTP
 - ▶ Data conversion issues (but not “landing” the data)
 - ▶ Limited data transformations (SQL-based access)





ETL – Extract, Transform, and Load





ETL – Extract, Transform, Load

- ▶ **Strengths**
 - ▶ Unlimited data transformations (e.g. star schema)
 - ▶ Isolate analytics from OLTP on a different platform
- ▶ **Weaknesses**
 - ▶ Cost
 - ▶ Latency – data might be stale by hours or days
 - ▶ Complexity risk
 - ▶ Data provenance
 - ▶ Security





HTAP – Hybrid Transaction / Analytic Processing

- ▶ Reduce
 - ▶ z/OS CPU utilization
 - ▶ Analytics latency
 - ▶ Data modeling
 - ▶ Complexity risk
 - ▶ Integration costs
 - ▶ Storage costs for archival and historical data



Next Steps

- ▶ Determine application requirements in terms of:
 - ▶ Latency (e.g. in-transaction analytics)
 - ▶ Volume of data
 - ▶ Granularity
 - ▶ Retention
 - ▶ Other data to be used in combination z/OS data
 - ▶ Structured
 - ▶ Schema-on-read
 - ▶ Preferred analytical tools



Next Steps

- ▶ Whiteboarding session
- ▶ Hands-on Workshop
- ▶ Proof-of-concept (POC)
- ▶ Financial analysis



Attributions

- ▶ Karen Durward, IBM
- ▶ Namik Hrle, IBM Fellow
- ▶ Jørn Thyssen, IBM
- ▶ Paul Wirth, IBM





Contacts

▶ Kim May

- ▶ kim.may@thefillmoregroup.com
- ▶ twitter.com/KimMayTFG
- ▶ <https://www.linkedin.com/in/kimmaytfg>

▶ Frank Fillmore

- ▶ frank.fillmore@thefillmoregroup.com
- ▶ twitter.com/ffillmorejr
- ▶ <https://www.linkedin.com/in/frank-fillmore-9a65976>
- ▶ tinyurl.com/ChannelDB2
- ▶ Flipboard for iPad, iPhone, Android: “BigData”

