

**THE FILLMORE GROUP**  
Relational Database Solutions

**Mainframe Modernization  
Field Notes  
The Fillmore Group – September 2020**

**A Premier IBM Business Partner**





## Agenda

- ▶ Introduction
- ▶ System z Platform
  - ▶ Operating Systems
  - ▶ Processor Primer
  - ▶ HiperSockets
- ▶ z/OS Data Sources
  - ▶ Information Management System (IMS)
  - ▶ Virtual Storage Access Method (VSAM)
  - ▶ Other



## Agenda (contd.)

- ▶ Data Provisioning
  - ▶ Extract, Transform, Load (ETL)
  - ▶ Replication
  - ▶ Federation
- ▶ Data Source Considerations
  - ▶ Single file - multiple formats
  - ▶ Data encoding
    - ▶ Extended Binary Coded Decimal Interchange Code (EBCDIC)
    - ▶ American Standard Code for Information Interchange (ASCII)
    - ▶ Unicode



## Agenda (contd.)

- ▶ Putting it all together
- ▶ Contacts and Resources



## Introduction

### The Fillmore Group, Inc.

- ▶ Founded in the US in Maryland, 1987
  - ▶ IBM Business Partner since 1989
  - ▶ Delivering IBM authorized education since 1994
  - ▶ IBM Gold Consultant since 1998
- <https://www.ibm.com/analytics/ibm-gold-consultants>
- ▶ IBM Champions since 2009





# System z Platform





## Operating Systems

- ▶ z/OS - Legacy mainframe operating system
- ▶ z/VM - Hypervisor
  - ▶ z/VM, z/OS, z/VSE, z/Linux
  - ▶ Conversational Monitor System (CMS)
  - ▶ Red Hat OpenShift (RHOS)
    - ▶ Container application platform based on the Kubernetes
- ▶ z/Linux
  - ▶ Red Hat Enterprise Linux (RHEL)
  - ▶ SUSE Linux Enterprise Server (SLES)
  - ▶ Ubuntu



## Processor Primer

- ▶ Central Processor (CP)
  - ▶ z/OS workloads
  - ▶ Capacity determines software costs
- ▶ Integrated Facility for Linux (IFL)
  - ▶ z/Linux workloads
- ▶ z Application Assist Processor (zAAP)
  - ▶ Java workloads
- ▶ z Integrated Information Processor (zIIP)
  - ▶ Eligible Db2 database workloads



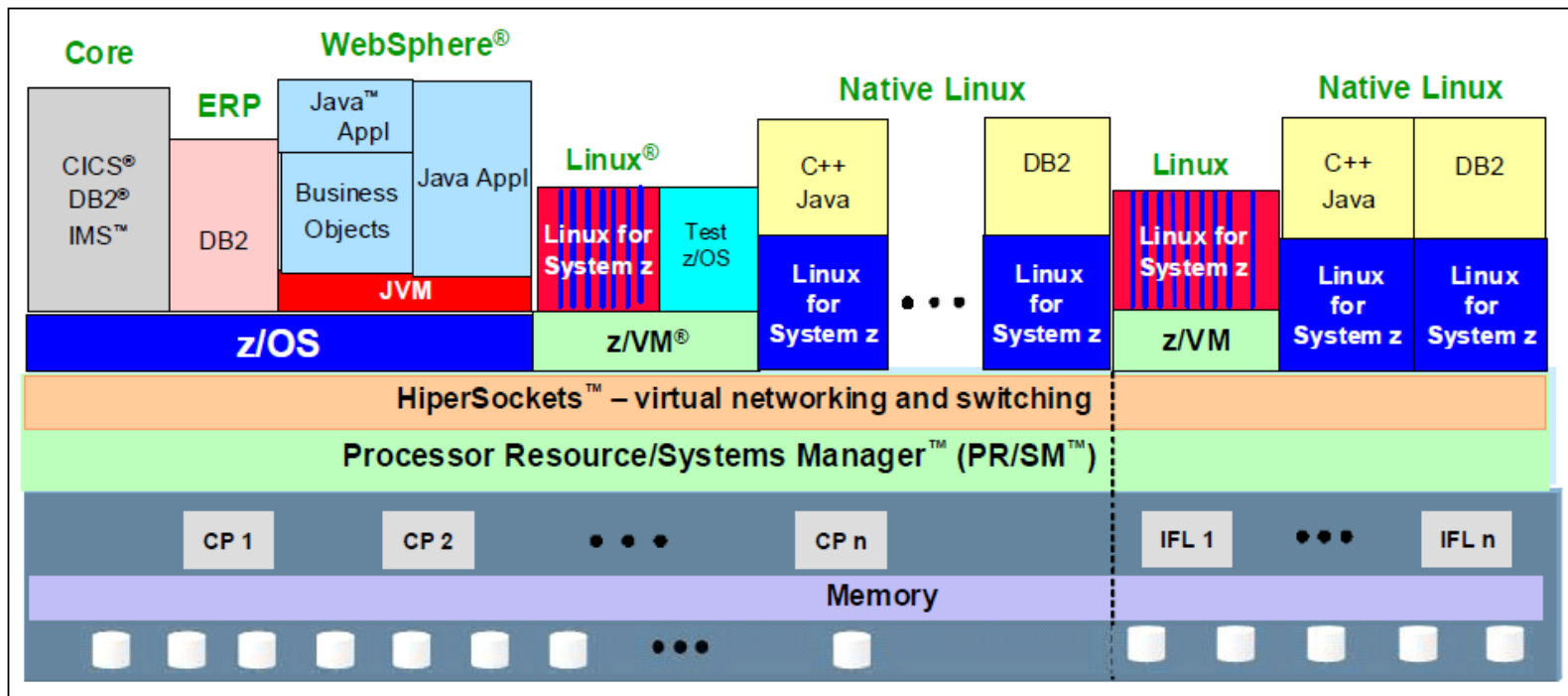


## HiperSockets

- ▶ internal Queued Direct Input/Output (iQDIO, internal QDIO)
- ▶ “Skinny” TCP/IP communication between consolidated
  - ▶ z/OS
  - ▶ z/VM
  - ▶ z/VSE
  - ▶ z/Linux
- ▶ Efficient, secure internal network for all System z workloads to communicate with one another
- ▶ Cross-address space memory transfer using the memory bus



# Operating Systems + Processors + HiperSockets





# z/OS Data Sources

Premier  
Business  
Partner





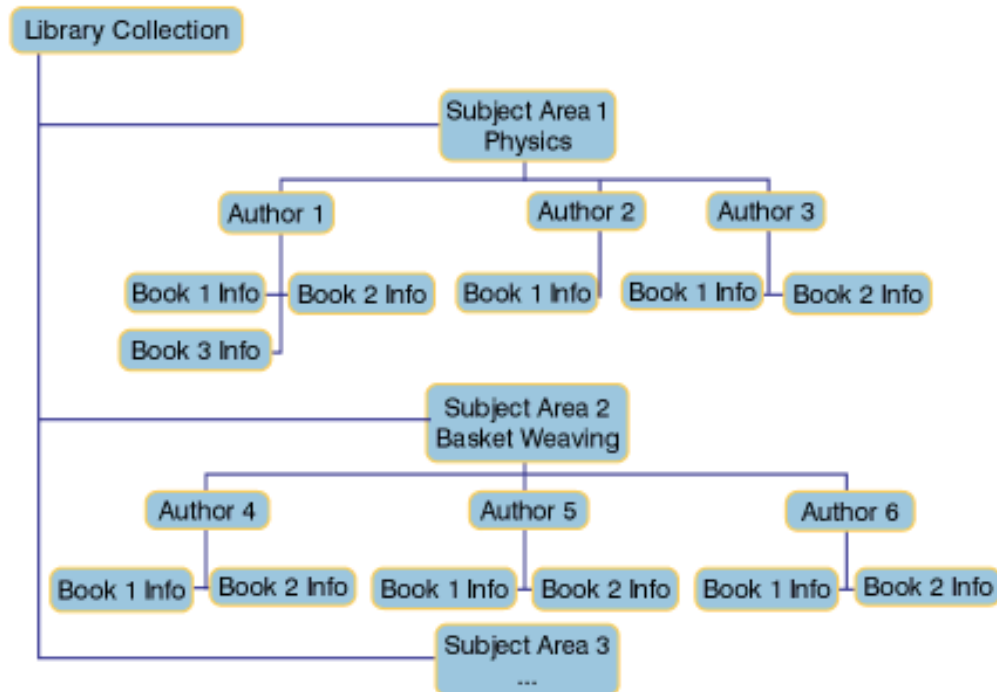
## **IMS – Information Management System v15**

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



## IMS (contd.)

Hypothetical Hierarchical Database Model



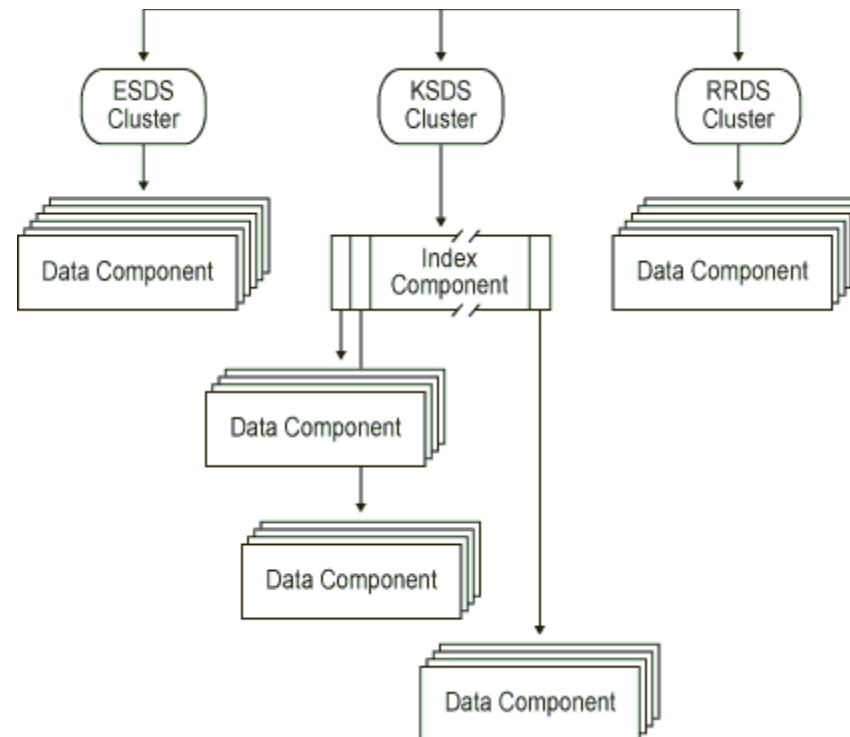


## VSAM – Virtual Storage Access Method

- ▶ 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 (Copybook)
- ▶ Single file might have multiple schemas
- ▶ Physical sequence of the records might have meaning
- ▶ “Dirty” data



## VSAM (contd.)





## Other File Types

- ▶ Physical Sequential Data Set (PS)
- ▶ Partitioned Data Set (PDS)
  - ▶ Directory
  - ▶ Members (typically related in some way)
- ▶ Schema imbedded in application program (Copybook)
- ▶ Single file might have multiple schemas
- ▶ Physical sequence of the data might have meaning
- ▶ “Dirty” data





# Data Provisioning



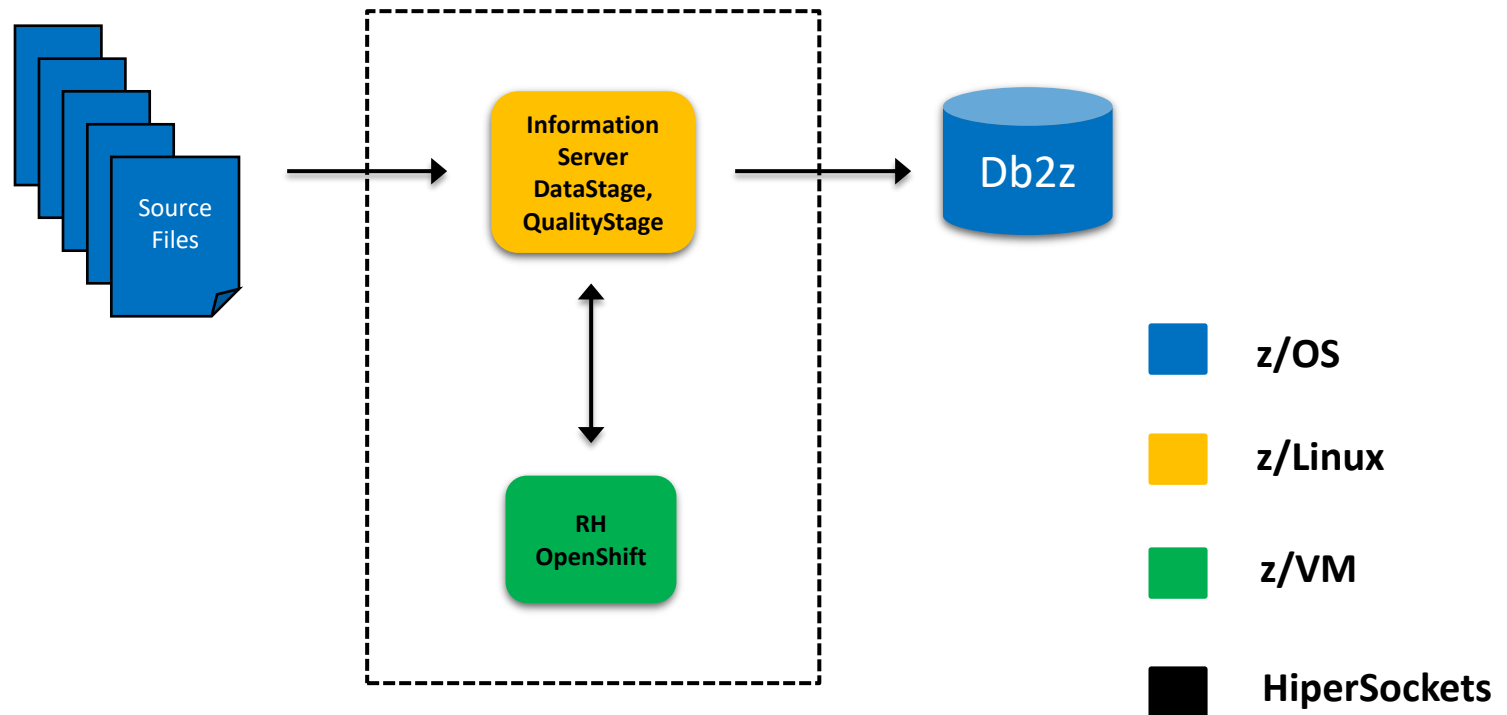


## Data Provisioning – Extract, Transform, Load

- ▶ IBM Information Server (DataStage, QualityStage)
  - ▶ Use case: periodic, batch-oriented
  - ▶ Complex File Format (CFF) Stage for input
    - ▶ Cobol Copybook can be used for formatting
  - ▶ Db2 Connector Stage for output
    - ▶ Bulk load – more efficient than INSERT
    - ▶ Delivers data to Unix System Services (USS) Pipes
    - ▶ Invokes Db2 for z/OS SYSPROC.DSNUTILU Stored Procedure from z/Linux

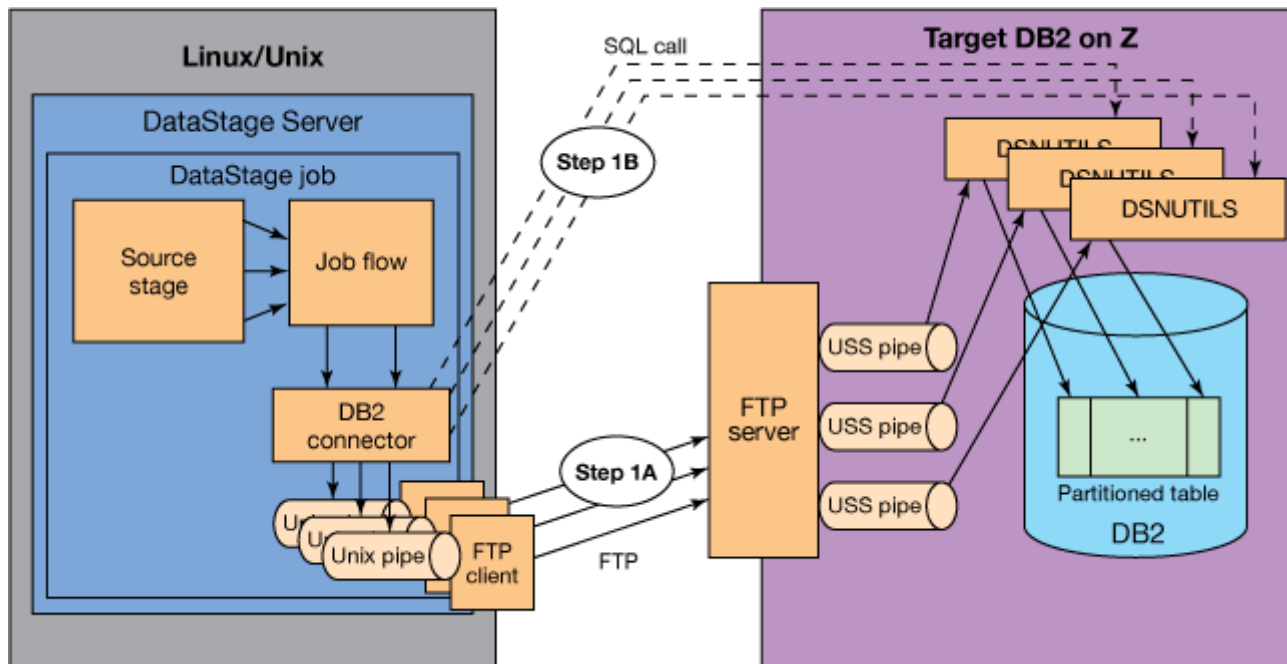


## Data Provisioning – ETL (contd.)





## Data Provisioning – Db2 for z/OS Bulk Load





## Data Provisioning – Read z/OS Files from RH OS

- ▶ **zdsfs Linux command**
  - ▶ Mount z/OS Direct Access Storage Device (DASD)
  - ▶ Read Physical Sequential and PDS files
  - ▶ Limitations
    - ▶ Read-only
    - ▶ z/OS authorization not enforced
    - ▶ No catalog access – specify volumes
    - ▶ File cannot span volumes

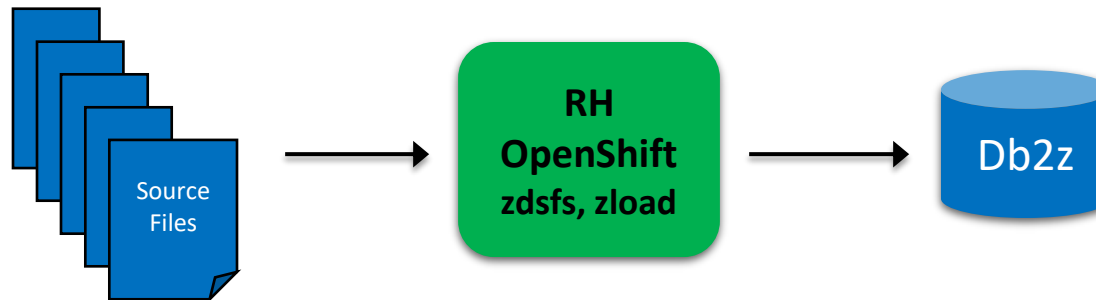





## Data Provisioning – Db2 12 DRDA Fast Load

- ▶ Db2 12 for z/OS Distributed Relational Database Architecture (DRDA) Fast Load
  - ▶ Db2 for LUW Call Level Interface (CLI) v11.1 FP 1
  - ▶ Db2 for LUW Command Line Processor (CLP) v11.1 FP 1
  - ▶ zload command
    - ▶ Bulk load – more efficient than INSERT
    - ▶ Loads local file
    - ▶ Invokes Db2 for z/OS SYSPROC.DSNUTILU Stored Procedure from RH OS
    - ▶ 100% of CPU offloaded to zIIP



## Data Provisioning – DIY ETL



-  z/OS
-  z/VM
-  HiperSockets



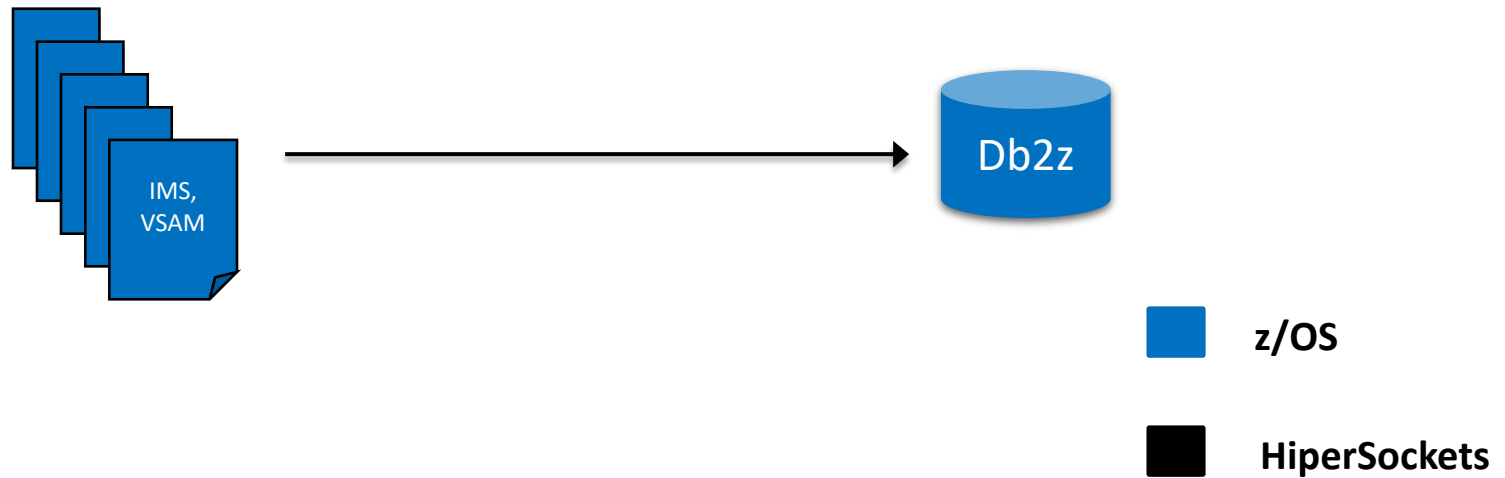
## Data Provisioning – Replication

- ▶ IBM Data Replication (IDR)
  - ▶ Use case: near real-time consistency between source, target
  - ▶ IDR Classic Change Data Capture (CDC) for z/OS
    - ▶ Non-relational mainframe sources: IMS, VSAM
    - ▶ Runs natively on z/OS
    - ▶ Log-based capture
  - ▶ IDR VSAM for z/OS Remote Source
    - ▶ Minimal footprint on z/OS
    - ▶ Log-based capture: VSAM logging must be enabled for source files



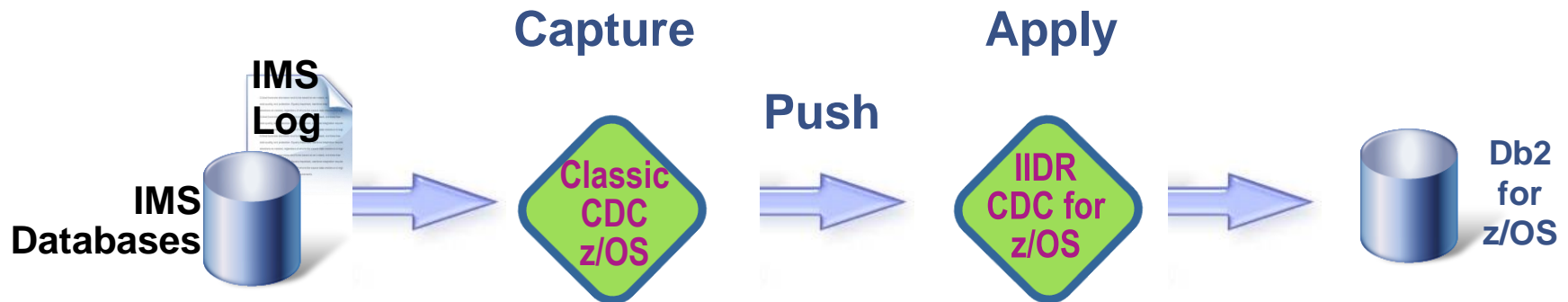


## Data Provisioning – Classic CDC for z/OS



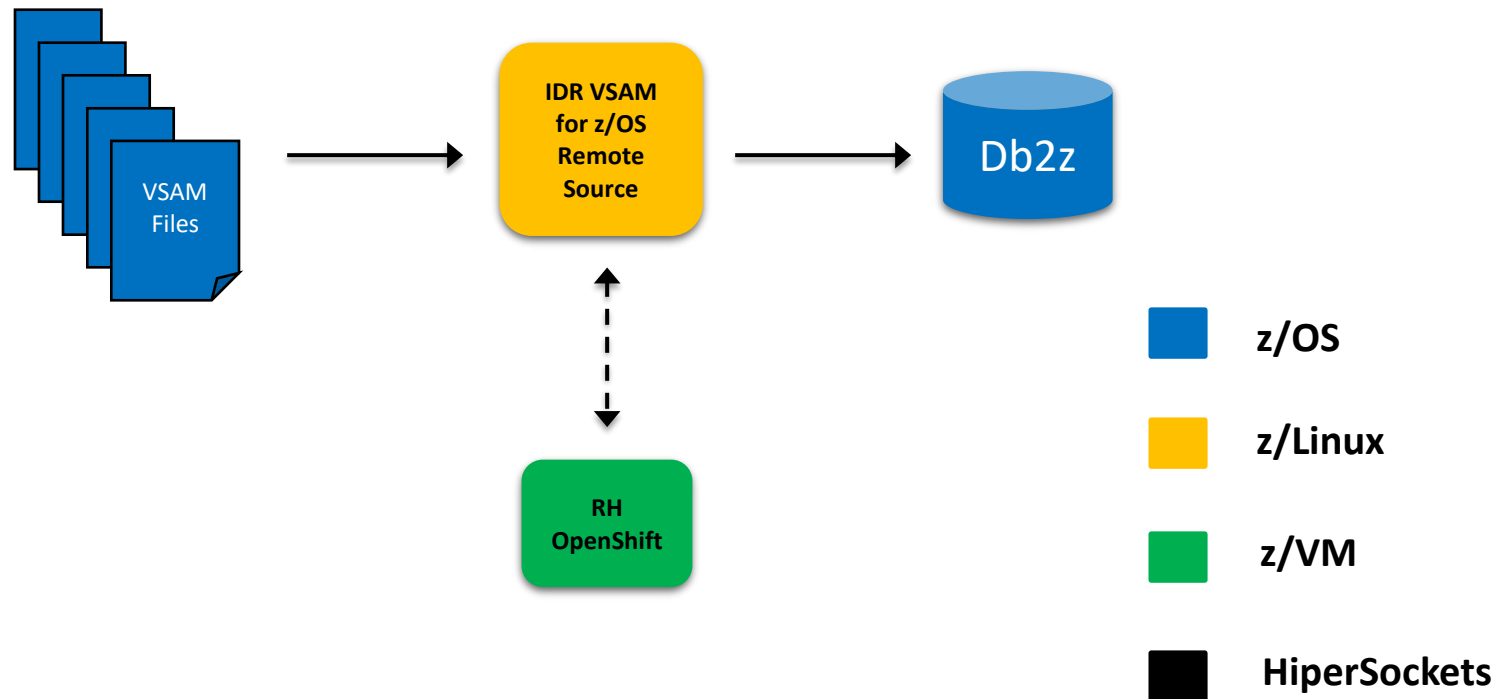


## Data Provisioning – Classic CDC (contd.)





## Data Provisioning – VSAM for z/OS Remote Source



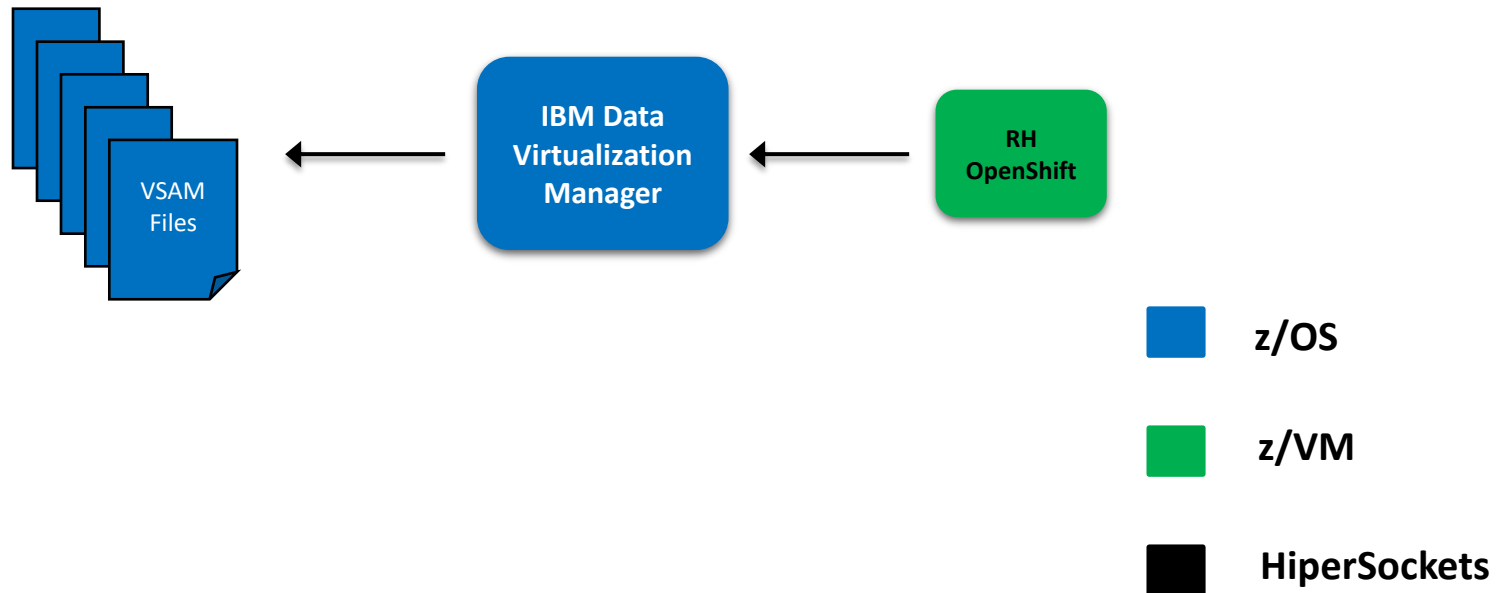


## Data Provisioning – Federation

- ▶ IBM Data Virtualization Manager (DVM) for z/OS
  - ▶ Accesses a variety of data sources in place
    - ▶ Mainframe: Db2z, IMS, VSAM, Adabas, IDMS...
    - ▶ Non-mainframe: Db2 for LUW, SQL Server, Oracle...
  - ▶ APIs: HTTP, SOAP, SQL, no SQL (REST with z/OS Connect)
  - ▶ Up to 99% zIIP offloadable
  - ▶ Prerequisite for DataStage to directly access VSAM files



## Data Provisioning – Federation (contd.)





# Gartner Magic Quadrant for Data Integration Tools

August 2019

ID G00369547

<https://www.gartner.com/doc/reprints?id=I-IQJ15E39&ct=190923&st=sb>





## IBM Strengths - Gartner

- ▶ **“Depth of integration offering.** Reference customers highlighted the completeness of IBM’s holistic data integration suite, including its rich functionality, variety of prebuilt functions and connectors, and its overall performance.”
- ▶ **“Diverse data integration delivery styles.** Reference customers use IBM’s products for traditional data delivery styles (data replication, batch processing), as well as more complex data delivery styles (including data synchronization and stream data integration). They praised IBM’s data integration tool portfolio for its ability to deliver complex data integration requirements that demand combinations of traditional and modern data integration styles, such as data replication, data virtualization and stream data integration for real-time analytics.”



## IBM Strengths – Gartner (contd.)

- ▶ **“Brand awareness and market presence.** IBM’s size and the global coverage of its business systems, infrastructure platforms and analytics solutions enable it to draw on a huge customer base and a wide product distribution model for positioning its data integration tools. Broad usage of IBM technologies within its customer base has driven the wide availability of implementation service providers and approaches to solving complex integration challenges.”





# Data Source Considerations

Premier  
Business  
Partner





## Single File – Multiple Formats

- ▶ Example: Purchase Order Header, PO Line Items
  - ▶ Multiple Cobol Copybooks
    - ▶ FILE Section
    - ▶ WORKING-STORAGE Section
  - ▶ REDEFINES clause in Copybook
  - ▶ OCCURS DEPENDING ON clause in Copybook



## Data Encoding - EBCDIC

- ▶ Extended Binary Coded Decimal Interchange Code (EBCDIC)
  - ▶ Typical for z/OS files of all types
  - ▶ Eight-bit, 256 character set
  - ▶ Sample hexadecimal to character conversion

x' C1 ' - "A"      x' C2 ' - "B"      x' C3 ' - "C" ...  
x' F1 ' - "1"      x' F2 ' - "2"      x' F3 ' - "3" ...

[https://www.ibm.com/support/knowledgecenter/zosbasics/com.ibm.zos.zappldev/zappldev\\_14.htm](https://www.ibm.com/support/knowledgecenter/zosbasics/com.ibm.zos.zappldev/zappldev_14.htm)



## Data Encoding - ASCII

- ▶ American Standard Code for Information Interchange (ASCII)
  - ▶ Typical for Windows, Linux, etc.
  - ▶ Seven-bit, (one parity bit), 128 character set
  - ▶ Note difference in “collating sequence” when sorting
    - ▶ In EBCDIC letters come before numbers
    - ▶ In ASCII numbers come before letters
  - ▶ Sample hexadecimal to character conversion

x' 31' - "1"      x' 32' - "2"      x' 33' - "3" ...  
x' 41' - "A"      x' 42' - "B"      x' 43' - "C" ...



## Data Encoding - Unicode

- ▶ Unicode – international standard
  - ▶ Typical for Linux
  - ▶ Up to forty eight-bit (up to six bytes)
  - ▶ 143,859 character set, many national languages
  - ▶ Special characters might convert to more than one byte
    - ▶ Example: cent sign “¢” converts to two bytes
    - ▶ Expansion might cause data not to fit in fields previously defined in EBCDIC or ASCII

<https://www.unicode.org/faq/>



## Other Data Considerations

- ▶ Physical sequence in a file might hold business meaning
  - ▶ For example: must process transactions in the order in which they reside in the file
  - ▶ Collating sequence differences in data encoding might change the record sequence
  - ▶ Db2 will not load or retrieve rows in a particular sequence unless specified (e.g. SQL ORDER BY clause)
  - ▶ Might require adding a record sequence number attribute



## Other Data Considerations (contd.)

- ▶ "Dirty" Data
  - ▶ Data typing enforcement frequently missing from VSAM, Physical Sequential files
  - ▶ Data cleansing is a must when moving to a repository with strong data type enforcement like Db2 for z/OS
    - ▶ Dates require particular focus
  - ▶ Embedded special characters might wreak havoc with delimited file input to a Db2 load
    - ▶ Characters in a free-form field (e.g. Description or Comment) might be record or field delimiter (e.g. comma , or vertical bar |)
    - ▶ Consider using fixed column formats



## Putting it all together...



**Business Analyst**



**DBA**



**DataStage Designer**



**Java coder**



**z/OS  
Communications Specialist**



**Cobol Programmer**



**Systems Programmer**



**DataStage Admin**



**RH OpenShift Admin**



**TCP/IP  
Communications Specialist**



**z/OS**



**z/Linux**



**z/VM**



**HiperSockets**





# Contacts and Resources





## Professional Services and Training

- ▶ Architecture and Implementation
  - ▶ Resilient, robust, scalable deployments
- ▶ Healthchecks
  - ▶ Monitoring and automation
  - ▶ Scripting
  - ▶ Schema evolution
- ▶ Version Upgrades
- ▶ IBM Authorized Training

**Kim May**

- [kim.may@thefillmoregroup.com](mailto:kim.may@thefillmoregroup.com)
- 410-465-6335



## Resources

- ▶ Mainframe hardware: Processing units

[https://www.ibm.com/support/knowledgecenter/zosbasics/com.ibm.zos.zmainframe/zconc\\_mfhwPUs.htm](https://www.ibm.com/support/knowledgecenter/zosbasics/com.ibm.zos.zmainframe/zconc_mfhwPUs.htm)

- ▶ IBM HiperSockets Implementation Guide

<https://www.redbooks.ibm.com/abstracts/sg246816.html?Open>

- ▶ Perform scalable data exchange using InfoSphere DataStage DB2 Connector (DataStage bulk load)

<https://www.ibm.com/developerworks/data/library/techarticle/dm-1508datastage-connector-db2-zos/index.html>



---

## Resources (contd.)

- ▶ IBM Db2 12 for z/OS Performance Topics (DB2 12 Fast Load)  
<https://www.redbooks.ibm.com/abstracts/sg248404.html?Open>



# Thank you!

Kim May, Vice President Business Development

[kim.may@thefillmoregroup.com](mailto:kim.may@thefillmoregroup.com)

[www.thefillmoregroup.com](http://www.thefillmoregroup.com)

