

IBM Data Virtualization Manager for z/OS

Session Number: 22621

John Casey Senior Solutions Advisor **Rocket Software** jcasey@rocketsoftware.com

Disclaimer



- © IBM Corporation 2018. All Rights Reserved.
- Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be
- IBM, the IBM logo, DB2, z/OS, CICS, IMS, MQ and QMF are trademarks of International Business Machines Corporation in the United States, other countries, or both.
- Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
- Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.
- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- Other company, product, or service names may be trademarks or service marks of others.

Critical Business Issues



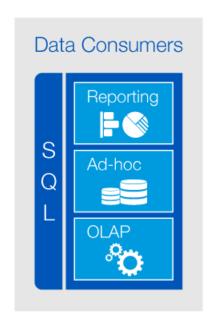




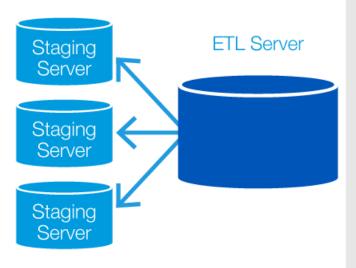
Traditional Data Integration

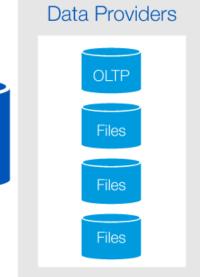
Challenges of moving data (Traditional ETL)

- Risk to data security
- Data inconsistency
- Rigid, limits business agility
- High cost and latency









IBM Data Virtualization Manager for z/OS



■ IBM Data Virtualization Manager for z/OS

➤ IBM® Data Virtualization Manager for z/OS® allows organizations to <u>virtualize z/OS data</u> with <u>other enterprise data</u> sources in <u>real-time</u> to provide comprehensive information that is readily consumable by analytics, digital applications and users.



Value Proposition:

➤ IBM Data Virtualization Manager for z/OS enhances IBM's premier z/OS data integration solution with new capabilities for <u>creating real-time</u>, <u>virtual views</u> of <u>enterprise data</u> - <u>mainframe and non-mainframe</u>. With Data Virtualization Manager, <u>data remains securely in place</u>, <u>ready when needed in the right format</u>, <u>without the cost and complexity associated with moving data</u>. Data Virtualization Manager unlocks the value of your mainframe data for real-time customer and business insights.

Key Features:

- Keeps data secure and in-place with real-time data virtualization on IBM Z
- ➤ Support for a <u>broad range of mainframe and non-mainframe data sources</u>
- ➤ Breadth of APIs and interface support for SQL, JSON, REST, SOAP and HTTP
- > z/OS resident optimization for improved performance and TCO
- > Abstraction layer to improve productivity and reduce reliance on mainframe skills

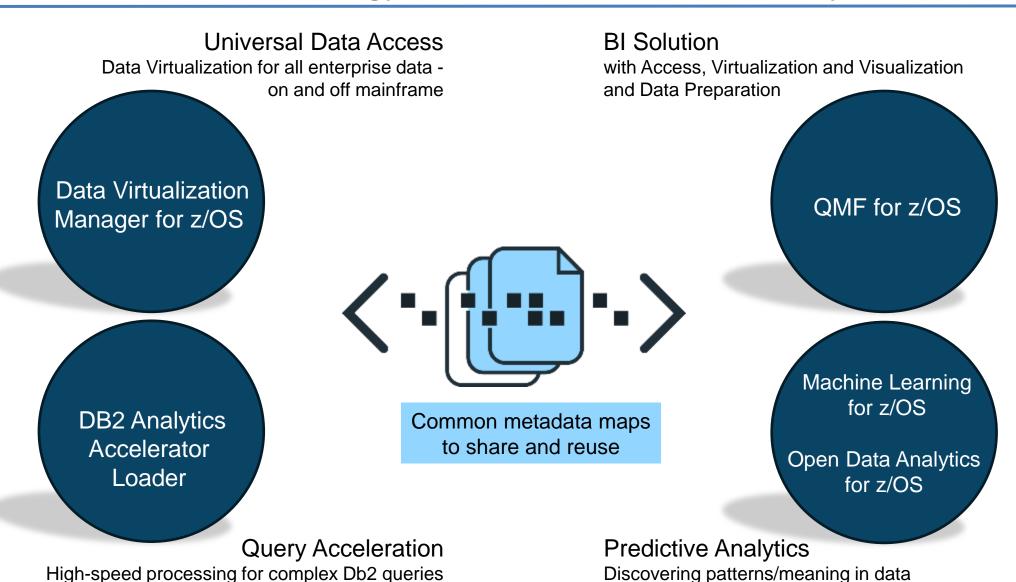
Data Virtualization Use Cases



Reduce the Faster, easier delivery of Need for immediate insight cost/complexity of modern systems of into your customer or accessing mainframe engagement business data Real-time Modernization Optimization Analytics



Data Virtualization Technology Accelerates Real Time Analytics

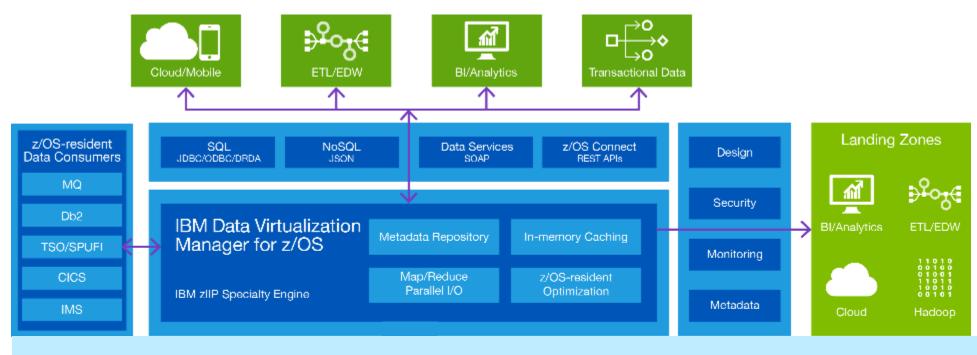


7



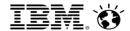


Virtualize z/OS data with other enterprise data sources in real-time to provide comprehensive information that is readily consumable by analytics, digital applications and users



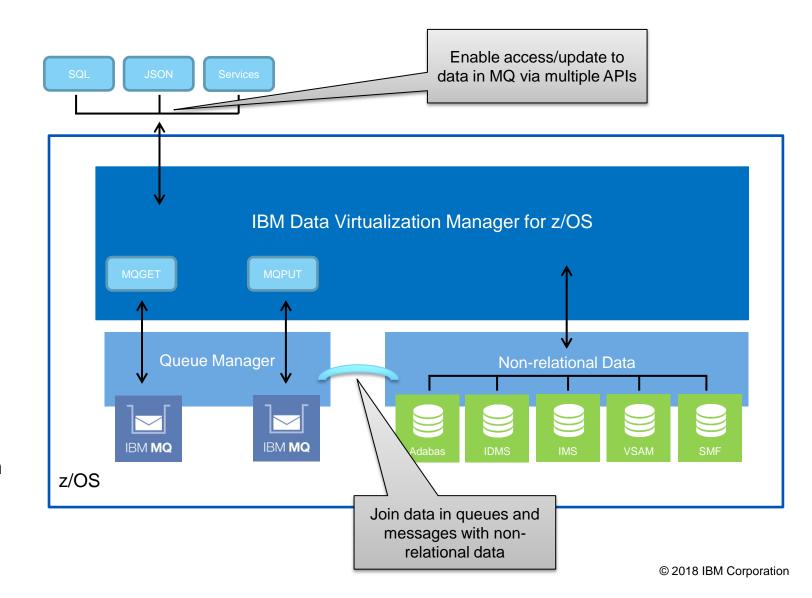
- zIIP eligible / avoid MLC costs
- Optimized for z14, z13, EC12
- Supports relational sources
- Rules engine

- Support any API:
 SQL, NoSQL, REST, SOAP, HTTP
- Python support
- Automated discovery though IBM ADDI
- Ships with industry mappings
- · Exploits 64-bit storage
- Built-in MapReduce
- Built-in parallel I/O



New Feature: MQ Queues and Messages

- Ability to perform SQL-based analytics directly against MQ queues/messages
 - treats MQ queues/messages as a new data source
 - ability to filter on individual messages to include in result set
- Join any DVM data sources with MQ messages to access as virtual table
- Support for MQGET (reads) and MQPUT (inserts)
 - Non-invasive, virtual access
 - Eliminates the need to extract data from queue



Virtual Parallel Data



- Allows multiple independent SQL requests to share a single result set
 - > Front ends
 - IDAA Loader
 - IzODA Spark Scala/Spark Java/Python DB API
 - QMF
 - JDBC/ODBC
 - Studio
 - Batch/DSSPUFI
 - > Back ends
 - Sequential/Tape
 - VSAM/IAM
 - Adabas
 - zFS
 - Logstreams
 - MQ Series
 - IMS
 - IDMS

Virtual Parallel Data

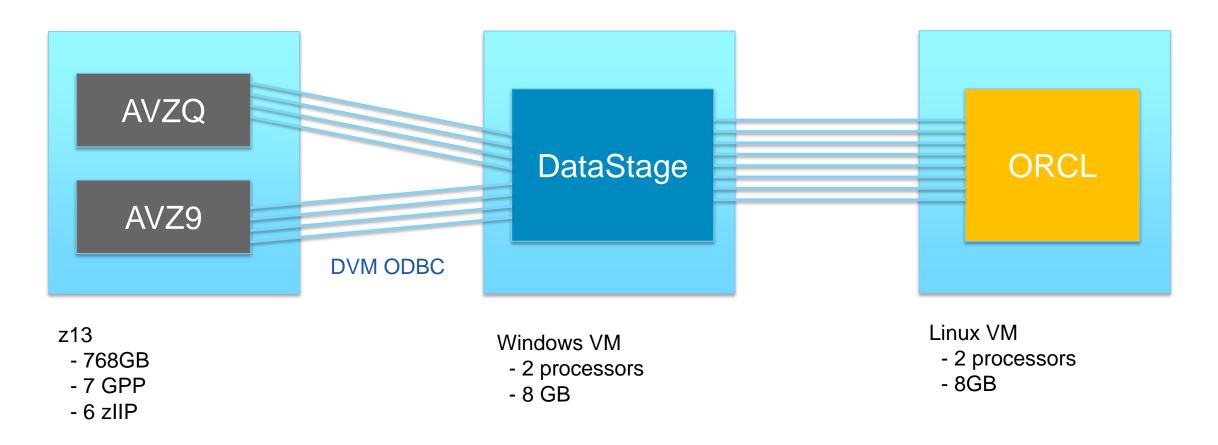


- VPD requests are joined in a named group
 - > Each request specifies the name of the group
 - ➤ The first request to arrive creates the group; I/O begins immediately.
 - > Requests arriving within a time interval are placed into the same group.
 - > The group is closed when the last request arrives, or the timeout is reached.
 - > Once the group is closed, the buffer can be wrapped/reused.
 - > The back-end I/O is done once, and buffered in one or more large, wrapping, 64-bit memory objects.
 - ➤ All group members share the buffered results.
 - Each VPD request can specify its own degree of parallelism using MRC.



Recent PoT Environment

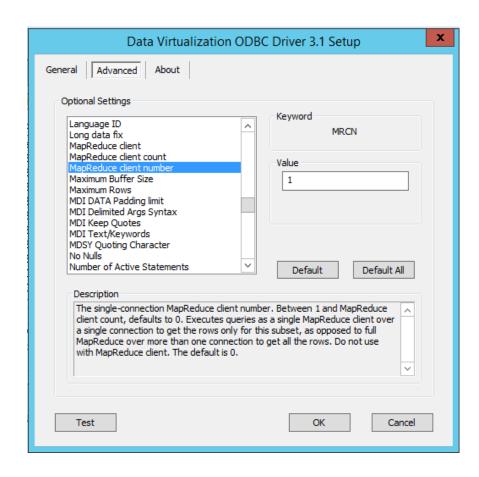
Objective: Replace multi-step ETL process to replicate IDMS tables to Oracle

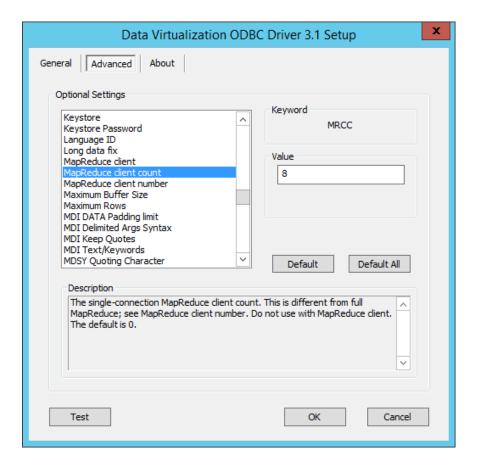


12



ODBC Map Reduce





13



How does IBM Data Virtualization Manager perform?

Tests performed at the IBM Systems Benchmark Center, Poughkeepsie, NY in Nov. 2017 running on IBM z13 using 800GB of financial data - flat files, with a multitude of fields

zIIP exploitation

99% of data virtualization runs on zIIP

Parallelism

- Test Case 4 and 2 have same configuration
- With degree of parallelism at 8 elapsed time is reduced from <u>98.68</u> minutes to <u>17</u> minutes
- Furthermore, by adding 3 zIIPs, test case 8 shows even greater improvement bringing the elapsed time down to <u>13.83</u> minutes.
- With enough zIIPs it will not unusual for us to see 1000% improvement for elapsed times

zIIP engine exploitation

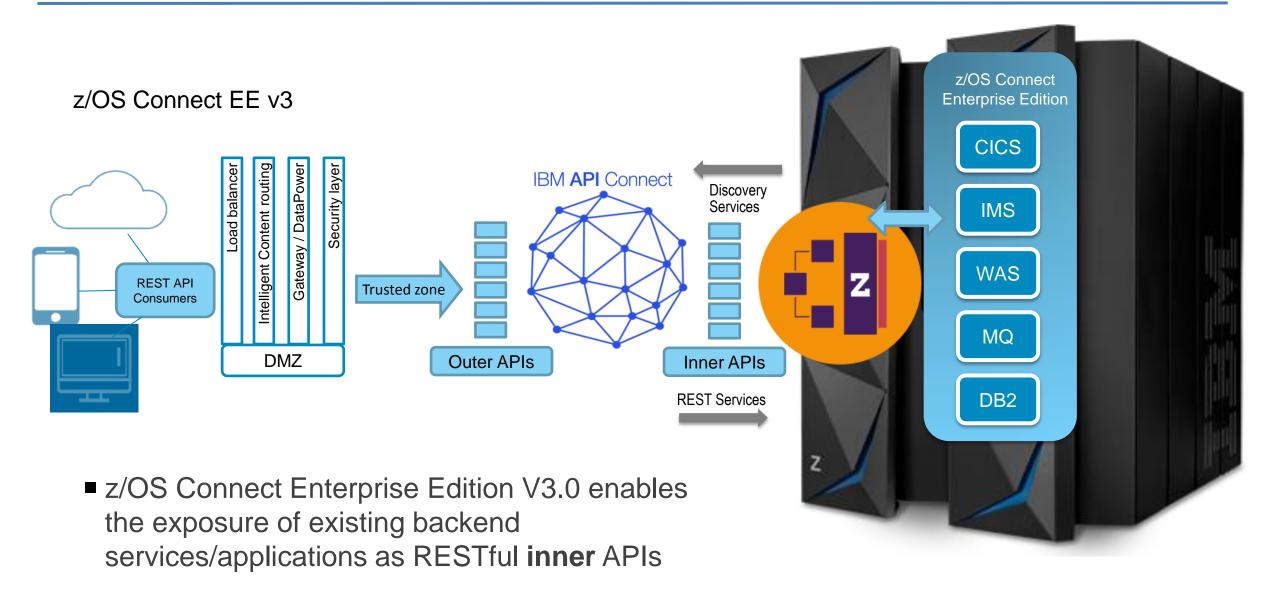
Row	Sum of CPU	Sum of zIIP	Sum of IIPCP	Sum of zIIP	%zIIP
Labels	Time	Time	Time	NTime	eligible
DVM 1	7099.03	5609.55	1389.58	5609.55	98.59%

Parallelism impact on elapsed time

Test Case	GPP's	Number of zIIP engines	Degree of parallelism	Elapse time in minutes	SMT
1	8	0	0	118.96	1
2	8	5	0	98.68	1
3	8	5	4	27.05	1
→ 4	8	5	8	17.14	1
5	8	5	8	20.84	2
6	8	5	10	17	2
	8	5	16	15.73	2
8	8	8	8	13.83	1
9	8	8	8	17.62	2
10	8	8	16	11.72	2



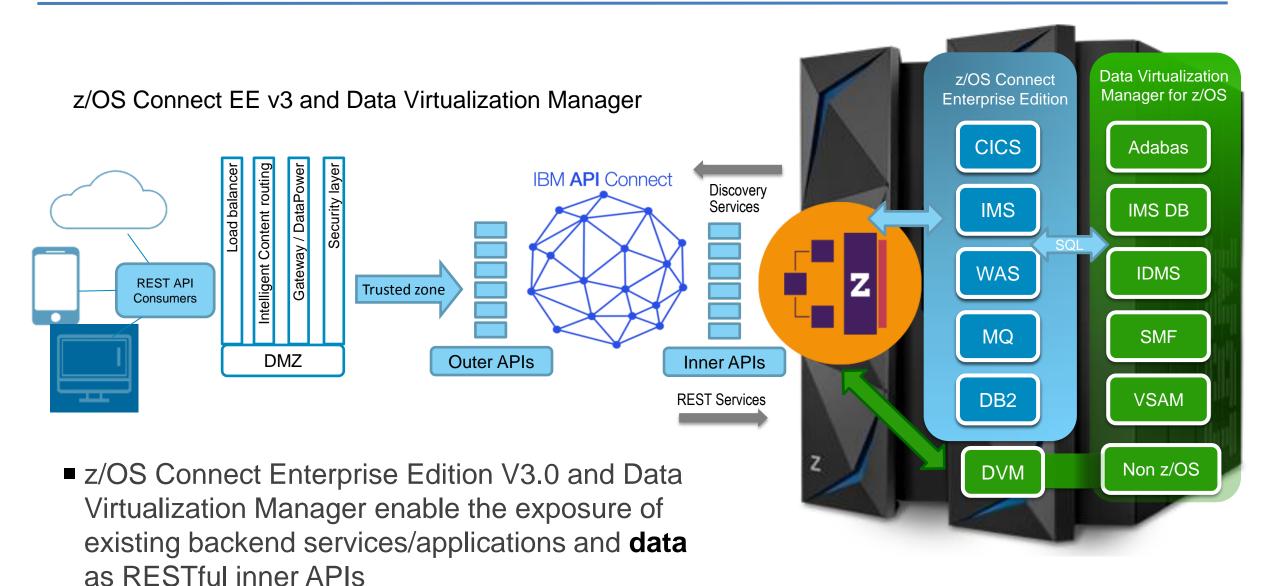




15 15





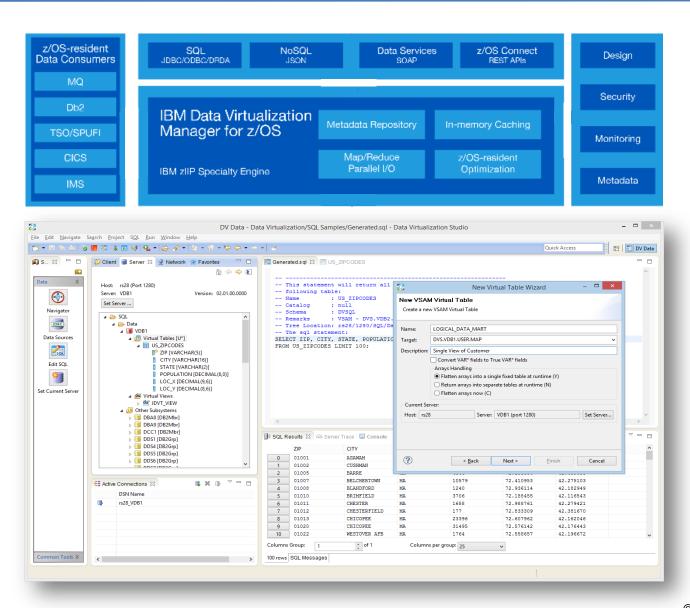


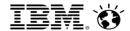


IBM Data Virtualization Manager for z/OS

Installable components

- DVM Server (SMP/E)
- Studio (Eclipse-based)
- Drivers (ODBC/JDBC)





Installing the IBM Data Virtualization Manager for z/OS (DVM)

	O R	R E C F	L R E C	BLK
Name	G	M	L	SIZE
SMPMCS	SEQ	FB	80	6400
IBM.HAVZ110.F1	PDS	FB	80	8800
IBM.HAVZ110.F2	PDS	FB	80	8800
IBM.HAVZ110.F3	PDS	VB	256	27998
IBM.HAVZ110.F4	PDS	FB	80	8800
IBM.HAVZ110.F5	PDS	FB	80	8800

IBM.HAVZ110.F6	PDS	FB	80	8800
IBM.HAVZ110.F7	PDSE	U	0	6144
IBM.HAVZ110.F8	PDS	FB	2048	18432
IBM.HAVZ110.F9	PDS	FB	80	8800
IBM.HAVZ110.F10	PDS	FB	80	8800
IBM.HAVZ110.F11	PDS	FB	80	8800
IBM.HAVZ110.F12	PDS	FB	80	8800
IBM.HAVZ110.F13	PDS	FB	80	8800
IBM.HAVZ110.F14	PDS	FB	2048	18432
IBM.HAVZ110.F15	PDS	FB	80	8800
IBM.HAVZ110.F16	PDS	FB	80	8800
IBM.HAVZ110.F17	PDS	FB	80	8800
IBM.HAVZ110.F18	PDS	FB	80	8800
IBM.HAVZ110.F19	PDS	FB	80	8800
IBM.HAVZ110.F20	PDS	FB	80	8800
IBM.HAVZ110.F21	PDS	FB	80	8800
IBM.HAVZ110.F22	PDS	FB	80	8800

BLK SIZE Standard SMP/E Installation

Program Directory lists the 22 Rel files shipped with the product

SAVZLOAD, SAVZCLOD and SAVZPRC must be PDSEs



Configuring the DVM Server

Step	Task description	For more information		
1	Create the server data sets using the hlq.SAVZCNTL members AVZDFDIV, AVZGNMP1 and AVZEXSWI.	See "Creating server data sets" on page 7.		
2	Set up the security application to use with the server using one of the following <i>hlq</i> .SAVZCNTL members: AVZRAVDB, AVZA2VDB, AVZTSVDB.	See "Defining security authorizations" on page 8.		
3	Configure Workload Manager (WLM) for optimum performance of the server.	See "Configuring Workload Manager (WLM)" on page 8.		
4	APF-authorize the product LOAD library.	See "APF-authorizing the LOAD library" on page 9.		
5	Create a copy of the product libraries (optional).	See "Copying target libraries" on page 9.		
6	Configure the server to support DBCS (optional).	See "Configuring support for code pages and DBCS" on page 10.		
7	Customize the server to access your data sources in hlq.SAVZEXEC(AVZSIN00).	See "Customizing the server initialization member" on page 10. Then, see Chapter 4, "Configuring access to data sources," on page 15 for the specific types of data sources the server should access.		

Installation and Customization Guide contains high level checklist

7 Required tasks

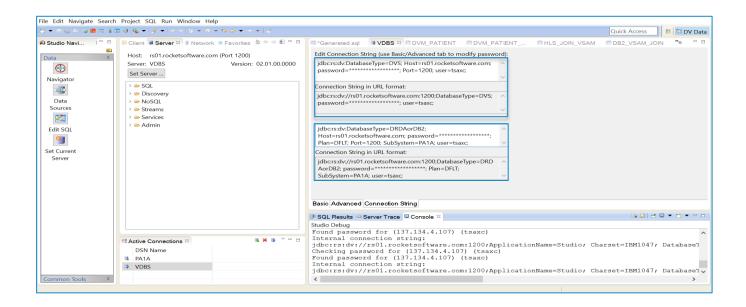
3 Optional tasks

	Table 1.	Cusi	tomi	zation	checklist	(continued)	
١		T					_

Step	Task description	For more information
8	Configure the started task JCL located in <i>hlq</i> .SAVZCNTL(AVZ1PROC) before you can start the server.	See "Configuring the started task JCL" on page 11.
9	Configure the CLIST that invokes the ISPF panels by using hlq.SAVZEXEC(AVZ).	See "Configuring the ISPF application" on page 12.
10	Verify the installation by creating a virtual table and accessing its underlying VSAM file (optional).	See "Verifying the Data Virtualization Manager server installation" on page 13.



DVM Studio connections

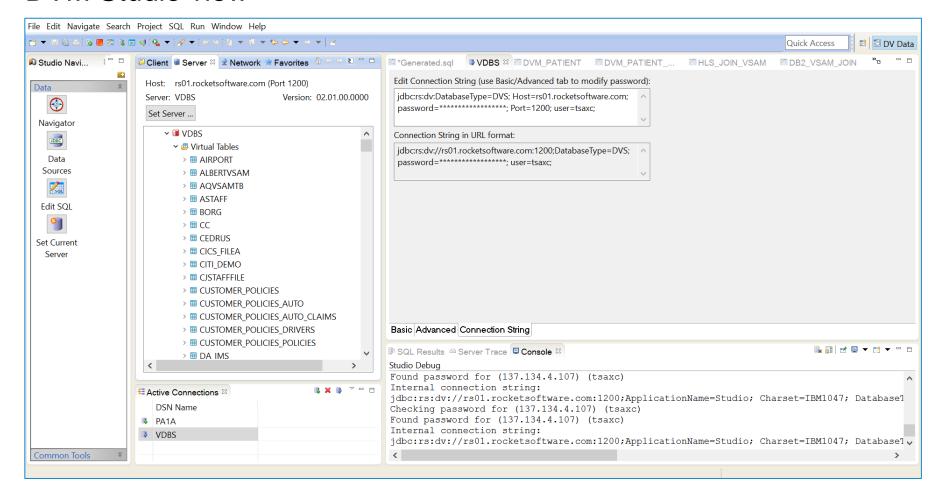


SDSF OUTPUT DISPLAY VDBS STC07224 DSID 2 LINE 110 COMMAND INPUT ===> 13.33.32 STC07224 VDS1290H VDBS OEPORTNUMBER '1200'

- Connection String
 - JDBC Connection string to DVM Started Task on z/OS
 - > JDBC Connection string to DB2



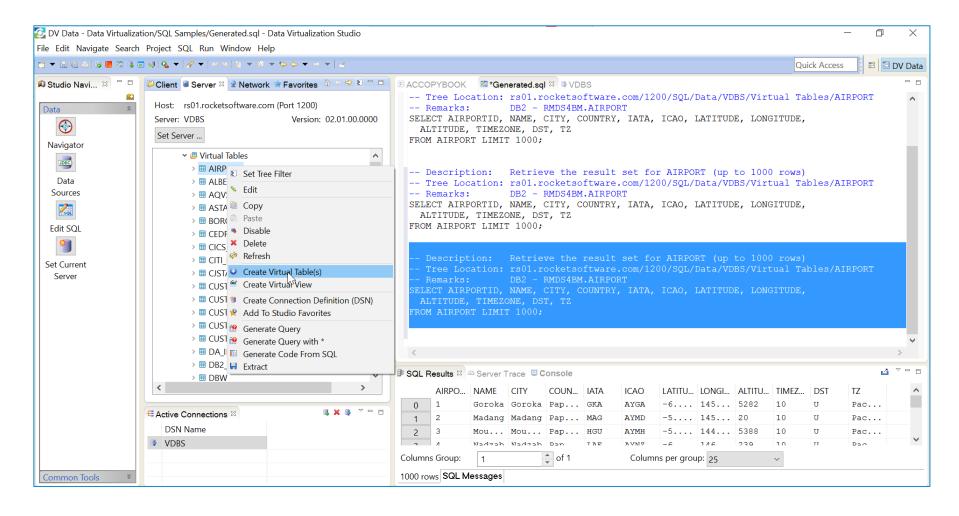
DVM Studio view



- Virtual table view within the studio
 - View existing virtual tables
 - Create new virtual tables
 - Execute sample test queries



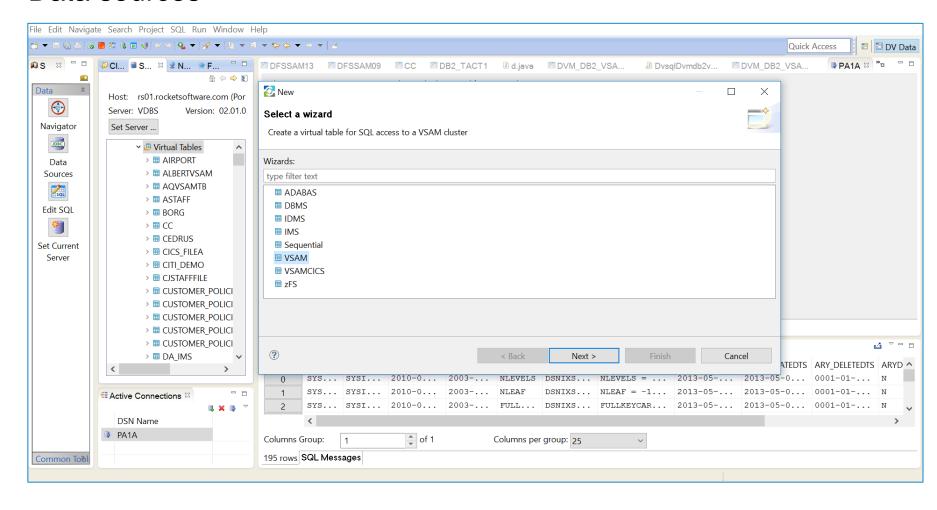
Creating virtual tables



- Virtual Tables
 - A right-click supplies various virtual table options
 - Select the "create virtual tables" option



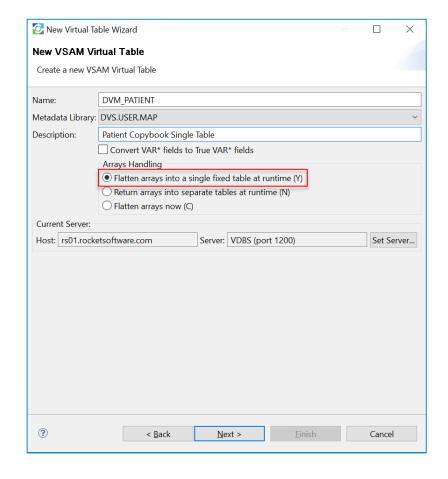
Data sources

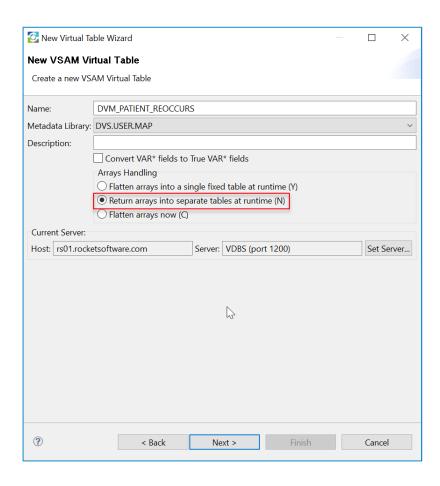


- Data sources
 - List of available z/OS data sources
 - VSAM will be the data source in this example



Virtual table options

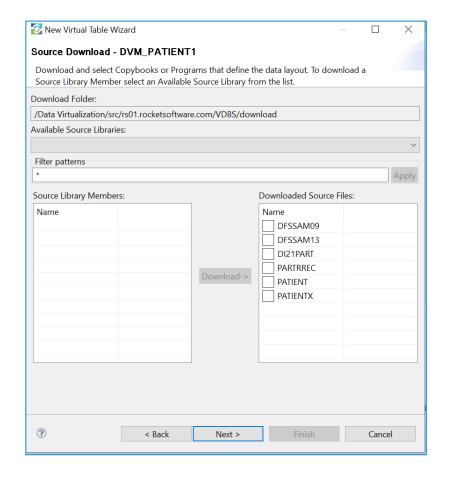


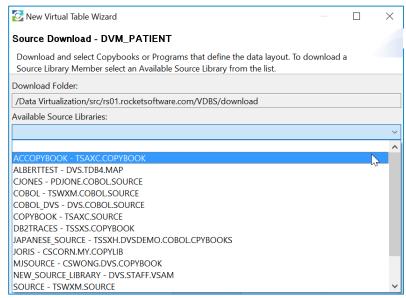


- Virtual table options
 - Supply a name for the virtual tables, names are nonqualified names
 - Option to create one flat table or parent/child tables for reoccurs
 - PDSE for metadata library



Copybooks

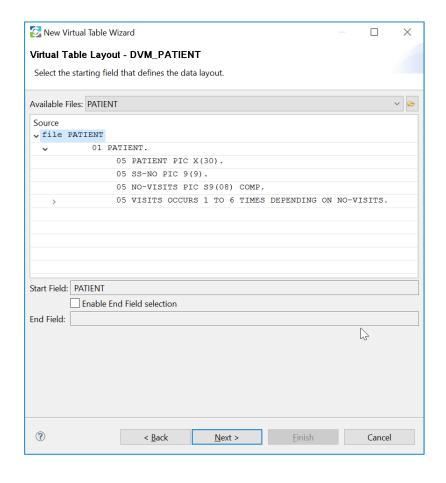




- Copybook options
 - Pull down menu to choose what copybook dataset to access
 - Members will be displayed
 - Select/download the appropriate VSAM copybook



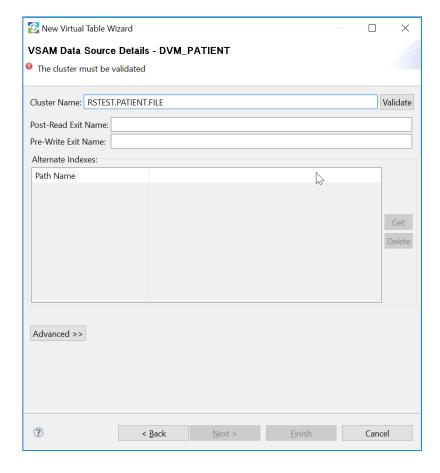
Copybook - continued

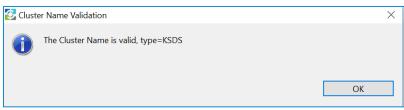


- Copybook options
 - Visual representation of selected copybook
 - ➤ Ability to select start/end fields
 - Example not selecting a "Filler" field



VSAM cluster name

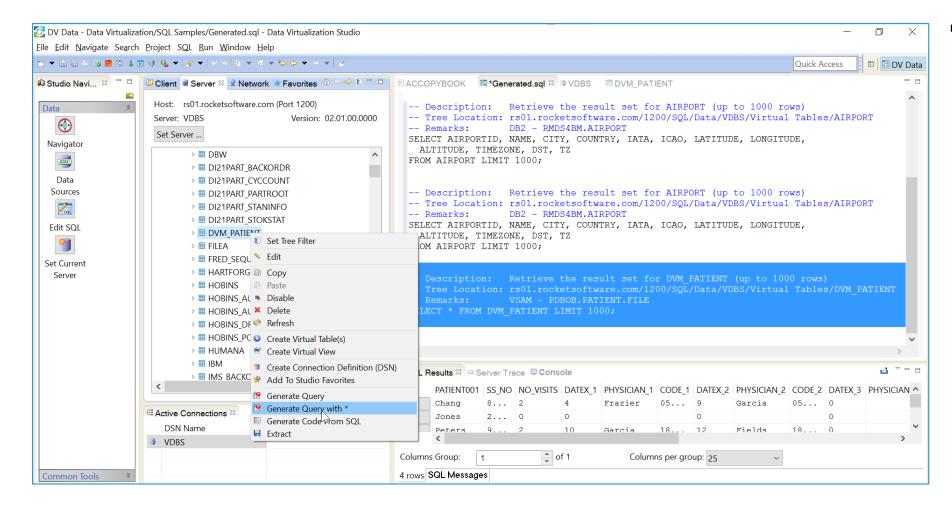




- Final options
 - Input VSAM cluster name and click to validate
 - ➤ Post/pre exit options
 - Ability to use an alternate index
 - > Finish to create



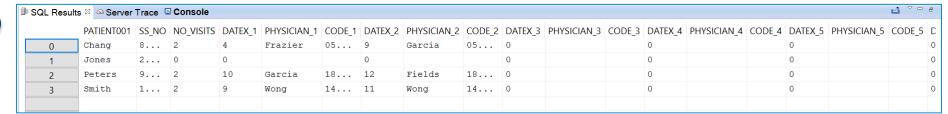
Data validation



- Data validation
 - Right-click the virtual table name
 - Select the option to query the data
 - A SQL virtual table query will be generated to validate the data at rest



Data validation - continued



	☐ SQL Results ☐ Server Trace Console ☐ SQL Results ☐ A Server Trace Console						
2		PATIENT001	SS_NO	NO_VISITS	CHILD_KEY	BASE_KEY	
	0	Chang	876543210	2	C388819587	C388819587	
	1	Jones	234567890	0	D1969585A2	D1969585A2	
	2	Peters	987564321	2	D785A38599	D785A38599	
	3	Smith	123456789	2	E29489A388	E29489A388	

8	SQL Resul	ts ඎ ⊶ Server Trac	ce 🖳 Console					⊿ ▽ □ <i>s</i>
		DATEX	PHYSICIAN	CODE	ROW_INDEX	PARENT_KEY	BASE_KEY	
Ш	0	4	Frazier	05-234	1	C388819587	C388819587	
	1	9	Garcia	05-234	2	C388819587	C388819587	
	2	10	Garcia	18-278	1	D785A38599	D785A38599	
	3	12	Fields	18-278	2	D785A38599	D785A38599	
	4	9	Wong	14-265	1	E29489A388	E29489A388	
	5	11	Wong	14-265	2	E29489A388	E29489A388	

Result Sets

- 1. All data flattened in one table
- 2. Parent table for VSAM cluster
- 3. Child table for reoccurs

Parent/child to connect tables

29



Application Access



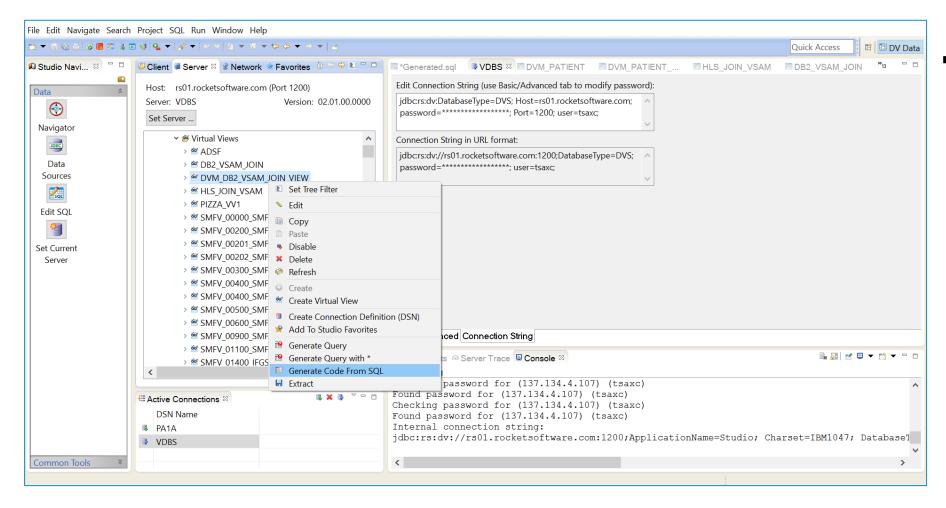
Access virtual tables from COBOL

300001
310001
320001
330001
340001
350001
360001
370001
380001
390001
400001
410001
420001
430001
440001
450001
460001
470001

- Traditional COBOL coding
 - Access data in place, VSAM, Sequential, IMS,etc.
 - DSCLIENT embedded in COBOL code to access virtual tables



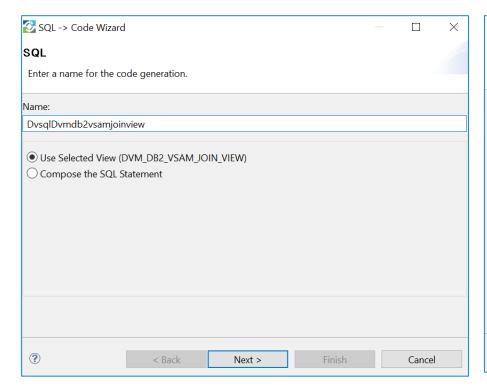
DVM code generator

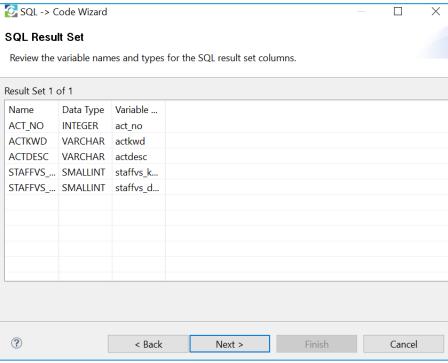


- Generating Code
 - Virtual table/view definition complete
 - Right-click on table/view and choose "Generate Code From SQL"



DVM code generator

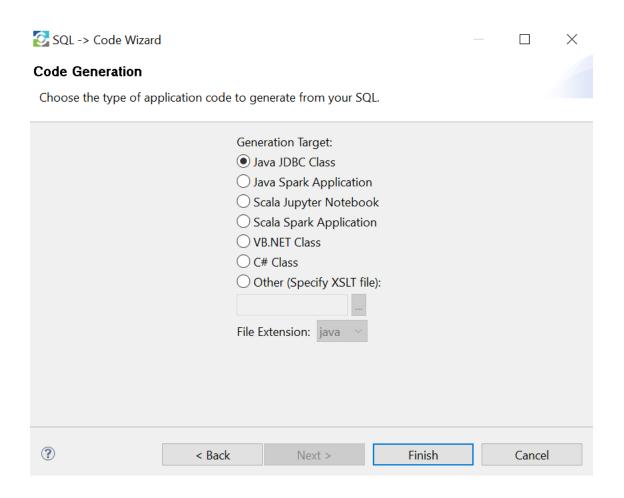




- Generating Code
 - > Input and target of code
 - Variables



Choose application type



Application code options



Generated Java code

```
public static void main(String[] args) {
    // Set up some local variables to hold things that we need
    Connection con = null;
    DvsqlDvmdb2vsamjoinview sql_obj= null;
    String connection_string = "jdbc:rs:dv://rs01.rocketsoftware.com:1200;DatabaseType=DVS; password=73e54bd68d6c3aaf

try {
    con = DriverManager.getConnection (connection_string);

    // Create the sql statement call object
    sql_obj = new DvsqlDvmdb2vsamjoinview(con);

    // Make the call to the sql statement
    sql_obj.prepareAndExecute();
```

- Generated code is a complete working program
 - Samples show the connection to DVM and the SQL prepare and execute

```
/** This method handles preparing our sql statement, and binding the
    * input parameters to our member variables.
    */
public void prepareSQLStatement() throws SQLException {
        // The double quotes in the sql statement have been escaped
        // and the newline characters have been removed.
        String sql = "SELECT ACT_NO, ACTKWD, ACTDESC, STAFFVS_KEY_ID, STAFFVS_DATA_NAME_L FROM DVSQL.DVM_DB2_VSAM_
        m_statement = m_connection.prepareStatement(sql );

        // bind all of the inputs
}

/** This method executes our sql statement.

*/
public void executeSQLStatement() throws SQLException {
        boolean has_resultset = m_statement.execute();
        if (has_resultset) {
```

ISPF interface



```
IBM Data Virtualization Manager for z/OS
Option 0
 Interface Facilities:
                                                     SSID
                                                            : AVZQ
                                                     Version :
    ACI
                        IDMS
                                                              01.01.00
     Adabas
                        IMS
                                                     Date
                                                          : 18/03/11
  3
     CICS
                        VSAM/Sequential
                                                     Time
                                                            : 17:22
                        DSSPUFI
  4
     DB2
 Data Virtualization Server Administration:
     Remote User - Manage Remote Users
    Server Trace
                      - Server Trace Facility - SIS SSID: AVZQ
                        Manage Data Virtualization Server
     AVZ Admin.
                        Data Mapping Facility
     Data Mapping
                        Event Facility Management
     Rules Mgmt.
                      - Monitor Server Activity
     Monitor
                        Streams Administration
     Streams
     Services - Services Administration
     Instrumentation - Instrumentation Server Administration
```

Trace browse



```
Cols 001 044
                           Server Trace --- 17:21:38 11 MAR 18
Command ===>
                                                               Scroll ===>
                        MESSAGENUM ----+---1----2----+---3----+---4----
DDMMM HH:MM:SS TCBADDR
                                   SWITCH CONTEXT EXECUTED - CTX SVCS NEONRRS.F
                                   BIND - ODBC - RC 0
11MAR 17:21:38 C08A8378
                                   AV04101T Password validated for USERID 'TSQA
11MAR 17:21:38 C08A8378
                                   AVQ4103T RACF MESSAGE - ALLOW
11MAR 17:21:38 C08A8378
                                   LOGON - TSQA - RACF MESSAGE - ALLOW
                                   DSNRLI BYPASSED OPEN - RC 0 REASON 00000000
11MAR 17:21:38 C08A8378
                                   WRITE EXECUTED
11MAR 17:21:38 C08A8378
                                                            - SOCK 0001 - WRITE
                                   READ EXECUTED
                                                            - SOCK 0001 - READ
11MAR 17:21:38 C08A8378
                                   SQL ENGINE OPEN DATABASE - RC 0
11MAR 17:21:38 C08A8378
                                   SET AUTO-OFF - SQLCODE 0
11MAR 17:21:38 C08A8378
                                   WRITE EXECUTED
                                                            - SOCK 0001 - WRITE
                                   READ EXECUTED
                                                            - SOCK 0001 - READ
11MAR 17:21:38 C08A8378
                                   SELECT DBKEY, T ID, T DTS, T ST ID, T TT ID,
11MAR 17:21:38 C08A8378
                                   WRITE EXECUTED
                                                            - SOCK 0001 - WRITE
11MAR 17:21:38 C08A8378
                                   READ EXECUTED
11MAR 17:21:38 C08A8378
                                                             - SOCK 0001 - READ
                                   SOL ENGINE ROLLBACK - SOLCODE 0
11MAR 17:21:38 C08A8378
                                   WRITE EXECUTED
                                                            - SOCK 0001 - WRITE
                                   READ EXECUTED
                                                            - SOCK 0001 - READ
11MAR 17:21:38 C08A8378
                                   DSNRLI BYPASSED CLOSE THREAD - RC 0 REASON 0
11MAR 17:21:38 C08A8378
11MAR 17:21:38 C08A8378
                           1376419 SQL ENGINE CLOSE DATABASE - RC 0
11MAR 17:21:38 C08A8378 0001376420
                                   WRITE EXECUTED
                                                            - SOCK 0001 - WRITE
```

Pink – executed on zIIP Green – executed on GPP

TCBADDR:

C* - zIIP and GPP same speed

D* - zIIP faster than GPP

Line commands:

P/PP – print

S/SS – print with zoom

Primary commands:

DISPLAY

PROFILE



A Modern Option to Data Federation

 InfoSphere Classic Federation Server for z/OS - provides integration of non-relational z/OS data with distributed tools and applications through SQL-driven access to z/OS data sources IBM® Data Virtualization Manager for z/OS® - allows organizations to virtualize z/OS data with other enterprise data sources in real-time to provide comprehensive information.

API support limited to SQL	Support for any API – SQL, NoSQL, REST, SOAP
No support for zIIP specialty engine	Up to 99% zIIP eligible for low TCO
No machine specific software versions	Support for 4 optimized machine specific versions (z14, z13, EC12, and z196)
Performance - does not support parallelism	High Performance - built in MapReduce parallelism
No support for DRDA	Full DRDA Support
No industry data model support	Pre-built metadata maps for ISO 8583 and 20022

38



ANY Data for ANY Application

Simple

Get transactional access, no data movement

Open to all Apps

Modern APIs enable access

Secure

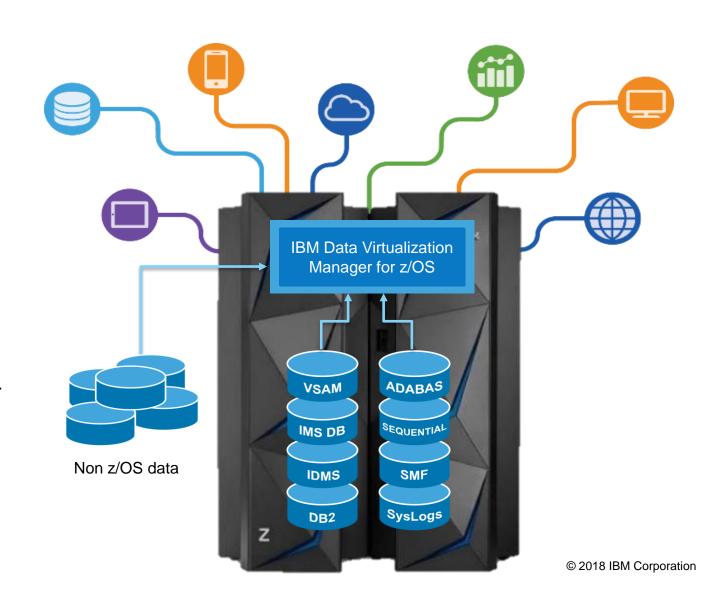
Avoid risk by reducing moving data off Z Systems

Fast

Exploits Z architecture, including parallelism and inmemory processing

Cost Effective

Keeps Z costs down with up to 99% zIIP offload





Where to Go for More Information



IBM Announcement

http://www-

01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/4/897/ENUS217-404/index.html&lang=en&request_locale=en

Product documentation

http://www-

304.ibm.com/support/docview.wss?uid=swg27020910 #ibmdvmz-lib

YouTube "IBM Data Virtualization Manager for z/OS" channel

https://www.youtube.com/channel/UCtbd_4oHoH-uKDYgSSRL7SA

Be sure to subscribe for new videos and to hit "Like" button

40