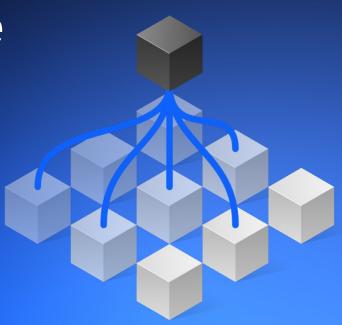
Db2 Click to Containerize

Tridex Db2 LUW

George Baklarz Principal Data & AI Technical Specialist – Americas



IBM

Agenda

- pureScale on AWS Update
- Introduction to Db2U and Containerization
- Db2 Click to Containerize Update
- Shift Scenarios
- Resources

IBM Db2 pureScale Primer

Db2 pureScale is an active-active database cluster solution for *transactional workloads*. It reduces the risk and cost of business growth.

IBM® Db2® pureScale® leverages the parallel sysplex architecture and data sharing technology to provide mainframe-class continuous availability (5 9s)¹

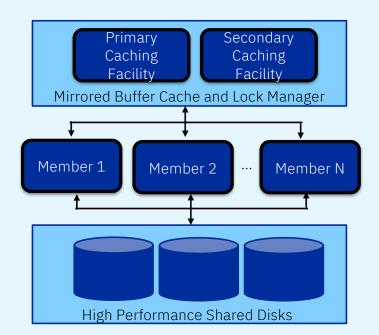
- 1. Always on No unplanned downtime
- 2. Extreme scalability Unlimited capacity
- 3. Application transparency No application changes, no data redistribution, and no performance tuning required

¹**An IBM c**ase study on PUMA with Db2 pureScale



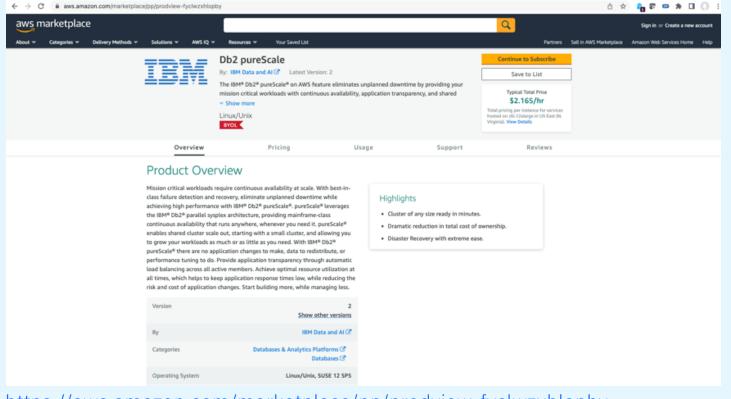
Business Challenge: Flexible scaling, microservices, and continuous availability of Db2 to interact with an increasingly digital world.

Solution: Add Db2 pureScale feature to support flexible scaling. Develop new functions as microservices in Red Hat[®] OpenShift[®] containers for faster digital innovation.



Db2 pureScale on AWS Marketplace





https://aws.amazon.com/marketplace/pp/prodview-fyclwzxhlopby

Db2 pureScale on AWS Documentation

- v Introduction		
 Visiteduction Visite Physical Science AM Use cases bipported Region Use cases bipported Region Prerequisities and Bar Architecture Diagram Security Costs Sizing Disployment Assets Hearth Check Backup and Recovery Rocified Maintenance Emergency Maintenan Support FAQ 	and	
	Security Root Privilege The deployment of the Db2 pureScale cluster on AWS using the Cloud Formation template does not require the usage of an AWS root account. It is recommended to use an administrative AWS 12	

https://ibm.biz/purescaleaws

The Challenges of Containerizing Databases

If a containerized application has a builtin database, that database disappears when the container shuts down and is recreated from scratch when it boots back up again. Many databases were not designed to be cloud native, compatible with containers or orchestrated by Kubernetes. Because of the benefits of moving to a containerized environment, there's a great deal of interest from enterprises in finding ways to run SQL and no-SQL databases on Kubernetes. The trick is to provide developers with a data fabric that just works, without forcing developers to struggle with security, auditability or scalability.

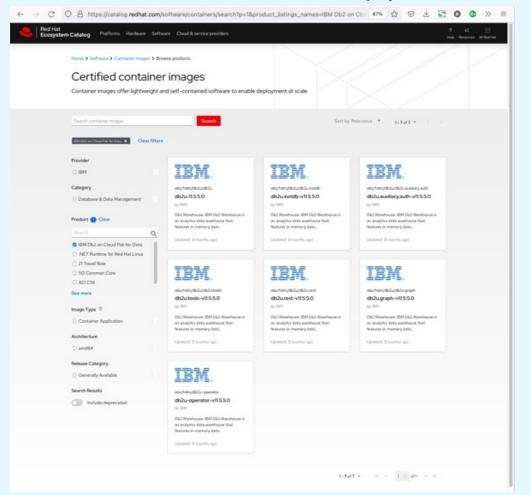
https://thenewstack.io/databases-finally-get-containerized/

Db2 Containerization Evolution



Phase I - Basic Install Automated application provisioning and configuration management Phase II - Seamless Upgrades Patch, minor version upgrades, application provisioning and configuration management Phase III - Full Lifecycle App lifecycle, storage lifecycle (backup, failure recovery

Db2 Containerization Support

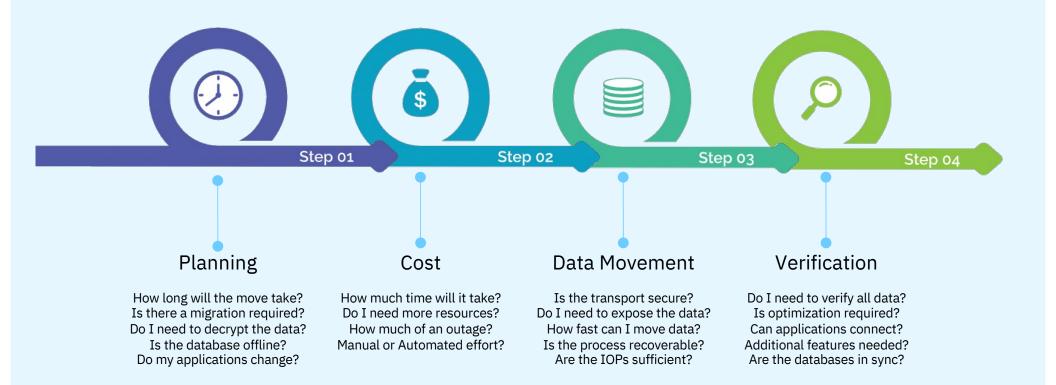


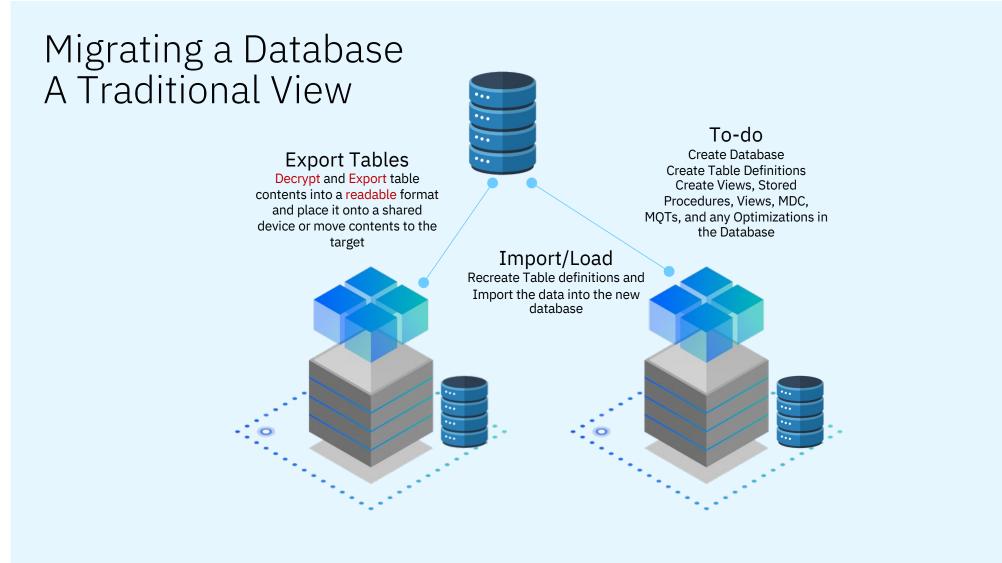


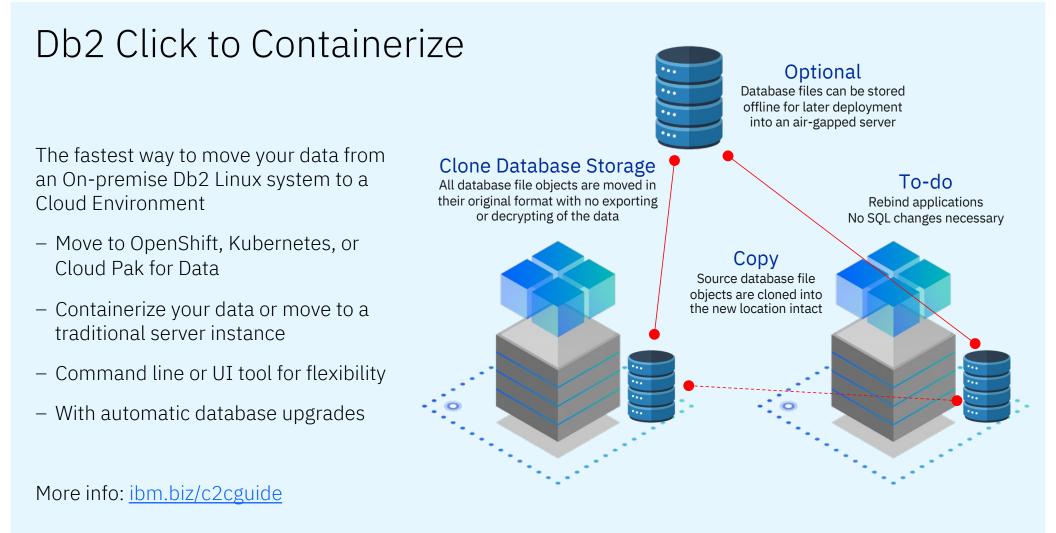


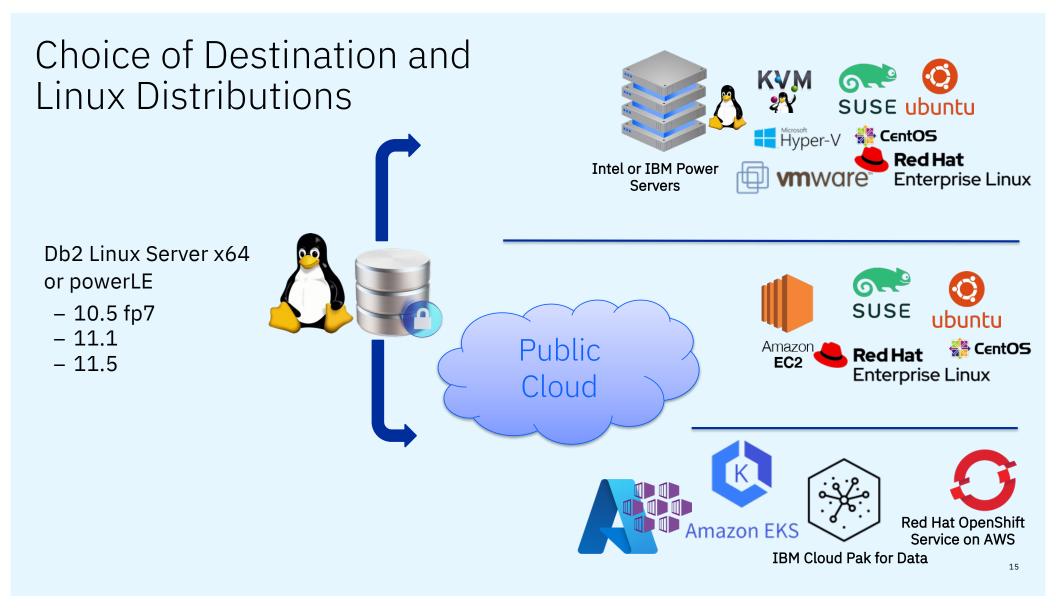


Migrating to Containers









Which Approach to Use?

Database Migration

Move between different O/S (AIX, zOS) Change configuration and codepages Unsupported database versions All or subset of data SQL Access only Topology change (MLNs for DPF) Object recreation required Data must be exported and decrypted

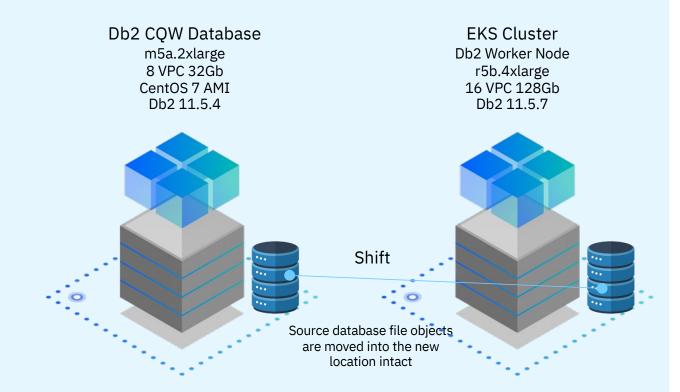


Db2 Click to Containerize

Move only Linux to Linux environment All database features moved intact Current releases only 10.5, 11.1, 11.5 Entire database movement Direct file access, no SQL used No change in configuration allowed No rebuild of objects required No exposure of the data

Demonstration Scenario

Table	Size
Database Size (RAW)	100Gb
Customers	200000
Addresses	941566
Customer Demographics	1920800
Date Dimensions	73049
Household Demographics	7200
Income Bands	20
Items	202960
Promotions	1000
Stores	394
Store Sales	262185014
Time Dimensions	36000

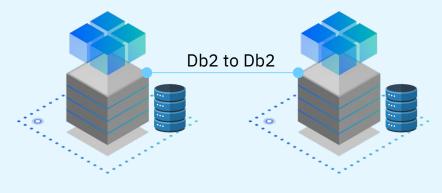


Shift Targets

- The Shift targets have been expanded in the newest release
 - Continues to support OpenShift 3.11, and 4+ clusters
 - CP4D V3.5 and V4+ are supported
 - New! Kubernetes Distributions (1.19+) are supported
 - New! MicroK8S, MiniKube for small systems
- Operating System Support
 - CentOS 6, 7, 8, RHEL 6, 7, 8, CentOS Stream
 - Ubuntu 18.04, 20.04. 22.04
 - Power LE
 - SUSE 15

Shift Modes

- The Shift mode has been expanded
 - Shift between Db2 instance and another Db2 instance (traditional Db2)
 - Create Database on target if it doesn't exist (Db2 instance only)
 - Set up HADR at target and initialize source and target (Pod or Instance)
 - Create cloned databases without requiring target information (Pod or Instance)
 - Useful in air-gapped environments
 - Clone copy to and from a POD
 - Reverse shift (back-out) from POD back to an instance

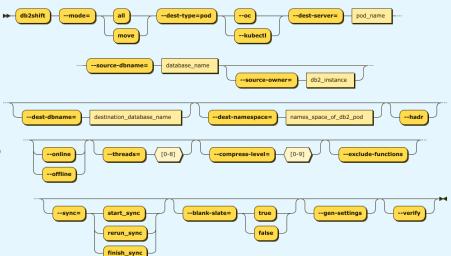


Db2 Shift Scenarios

- The Db2 Shift utility provides support for 9 scenarios
 - Shift Operations
 - Shift a Db2 database to OpenShift, Kubernetes or CP4D
 - Shift a Db2 database to another Db2 instance
 - Clone and Deploy a Database
 - Create a Cloned copy of the Db2 database for later deployment
 - Deploy a clone into an OpenShift, Kubernetes or CP4D container
 - Deploy a clone into another Db2 instance
 - HADR Operations
 - Initialize HADR between Source and Target POD
 - Initialize HADR between Source and Target Instance
 - Initialize DMC and LDAP Authentication for CP4D
 - Copy Cloned Databases to a POD

Containerize Db2 to OpenShift, Kubernetes, or CP4D

- Containerizing an existing Db2 database on an on-premise system to OpenShift, Kubernetes, or CP4D is the most common scenario for using Db2 Shift
- In order to move a Db2 database to a POD you will require the following information:
 - Source Database details
 - Destination location and POD details
 - Type of Containerization environment
 - Shift Options
- The target of the Db2 Shift operation can be OpenShift, a Kubernetes cluster, or Cloud Pak for Data



Containerize Db2 to OpenShift, K8s or CP4D Dialog

• The Source and Destination Details are the only required fields

	Shift to Db2U on OpenShift or Kubernetes
Source Details	
Database Name Instance Owner	: db2inst1
instance owner	
Destination Detai	ls
	: 📀 OpenShift 🜔 Kubernetes
Database Name	: db2oltp_
POD Project	: db2u+ : c-demo-db2u-0+
POD Name	: c-aemo-ab2u-0
Metadata	: • Refresh 📀 None 🔿 Settings Only 💿 Verify
	: • None • Initialize • Refresh • Finalize
Move Options	: 💿 Database 📀 Database/LDAP/DMC 📀 Move Database for HADR
	: O Do not move external routines
	: • Online Move • Offline Move
Thread count	
Compression Overrides	
overrides	

• The additional fields provide more options for the Shift process

Optional Fields

- Metadata
 - Used for gathering information on the source and target destination
 - No need to use this unless you want to test the connections between the source and the target (verify) or you are shifting a database that needs to be offline (upgraded)
- Sync Options
 - By default, all shifts are performed from beginning to end
 - You can shift incrementally by using the Initialize/Refresh/Finalize Options

• Move Options

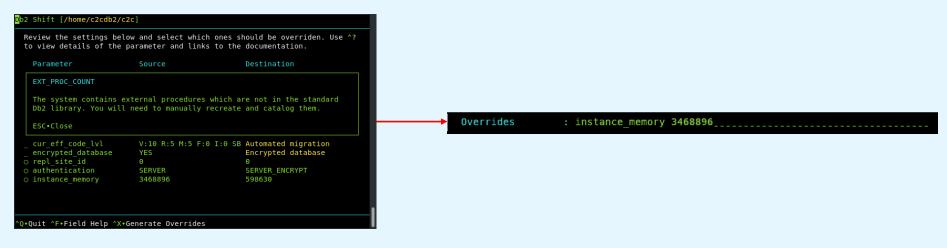
- Use Database for OpenShift and K8s, Database/LDAP/DMC for CP4D, and HADR if you
 want the target POD created as a standby for the primary database
- Exclude Routines
 - Do not move any external routines found on the source database library path

Optional Fields (Continued)

- Database Mode
 - The database can be online during a shift operation (11.5 only)
 - For databases that require an upgrade (10.5 and 11.1), the database needs to be offline for the shift to work properly (Db2 requirement)
- Thread Count
 - The Db2 shift utility can use up to 8 threads when communicating with the target location
 - The more threads, the more throughput, but at the expense of CPU resources
 - The limit is usually the Ethernet performance, not the threads
- Compression
 - The data stream can be compressed when communicating with the target
 - If the network is >1Gb/sec then there is no benefit from using this setting
 - Anything less than 1Gb/sec may find compression helps at the expense of CPU cycles

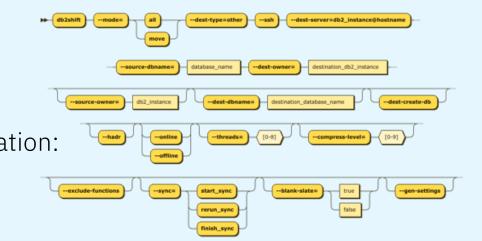
Optional Fields (Continued)

- Overrides
 - All source database settings are retained at the target location
 - Instance settings do NOT move to the target
 - You may need to adjust the target's instance settings
 - This field provides a way of setting instance values that will be updated during the shift process
 - The Analyze option (^A) will provide a summary of the differences between the source and target and allow you to select which settings to override



Shift Db2 to another Instance

- This format of the Db2 Shift will take an existing Db2 database on an on-premise system, and shift it to another traditional Db2 system to:
 - Another on-premise server
 - A Cloud virtual machine (i.e. EC2)
- This does not containerize Db2!
- The command requires the following information:
 - Source Database details
 - Destination location
 - Shift Options



• The db2shift program assumes that you are currently connected to the instance that has the Db2 database and have ssh connectivity to the target server

Shift Db2 to another Instance

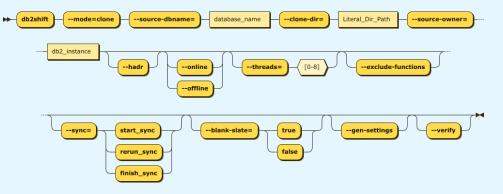
• The Source and Destination Details are the only required fields

Source Details	Shift Db2 to another Db2 instance
Database Name	: banking_
	: db2inst1
Destination Detai	
	: banking_ 🛛 💿 Force Database Creation
Instance Owner	: db2inst1
Server Address	: 10.0.0.2
	: • Refresh 🔿 None 🔿 Settings Only 🔿 Verify
Sync Options	: 💿 None 💿 Initialize 💿 Refresh 💿 Finalize
HADR Option	: 🔿 Move Database for HADR
	: 🔿 Do not move external routines
Database Mode	: 💿 Online Move 💿 Offline Move
Thread count	: 4
Compression	
0verrides	:
inter INSTANCE over	rrides as setting value; setting value; etc

- The FORCE DATABASE CREATION option will create the database at the target
- The additional fields provide more options for the Shift process and are identical to the Db2 Shift to POD command

Clone a Database

- The Db2 Shift clone option is used to take an existing Db2 database that is currently on-premise, and clone it into a directory
 - This cloned database can be transported to another server and deployed at that location
 - The advantage of cloning is that the destination does not need to be connected to the source location and the deployment of the clone can be done at a more convenient time
- The Db2 Shift program requires the following information:
 - Source Database details
 - Clone Options
- The destination details are not required to clone a database



Create a Clone of a Database

• The Source is required as well as the directory to place the cloned copy

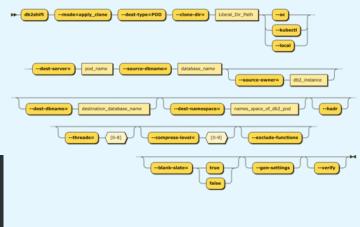
Db2 Shift [/home/o	c2cdb2/c2c]
	Clone a Database for Future Deployment
Source Details Database Name Instance Owner Directory	: <mark>f</mark> lights_ : db2instl : /tmp/cache+
Sync Options Exclude Routines	 Refresh O None O Settings Only O Verify None O Initialize O Refresh O Finalize Do not move external routines Online Move O Offline Move 4
	Help ^D•Directory ^X•Review and Execute

- The additional fields provide more options for the Shift process and are identical to the Db2 Shift to POD command
- HADR and Compression options are not used

Deploy a Clone to a POD

- This Db2 Shift option will take a database clone and deploy it into a Db2u pod running on OpenShift, Kubernetes or CP4D
- The requirements for deploying to a POD are identical to a direct shift
- The command requires the following information:
 - Database name and clone location
 - Destination POD details
 - Type of Containerization environment
 - Shift Options

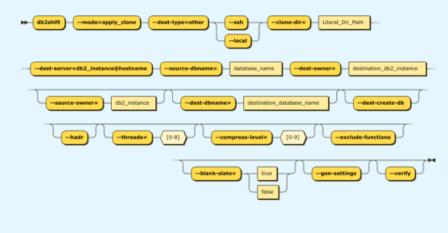
Source Details	Deploy a Clone Copy into a Container
Database Name	Diabte
	db2inst1
	: /tmp/cache
Destination Deta	its
	: 💿 OpenShift 💿 Kubernetes 💿 Local
Database Name	
POD Project	: db2u
POD Name	: c-demo-db2u-0
Metadata	: • Refresh • None • Settings Only • Verify
HADR Option	: O Move Database for HADR
Exclude Routines	: 🖕 Do not move external routines
Thread count	
Compression	
Overrides	



Deploy a Clone an Instance

- This Db2 Shift option will take a database clone and deploy it into another Db2 instance running natively (in any environment including Cloud VMs)
- The requirements for deploying to an Instance are identical to a direct shift
- The command requires the following information:
 - Database name and clone location
 - Destination instance details
 - Shift Options

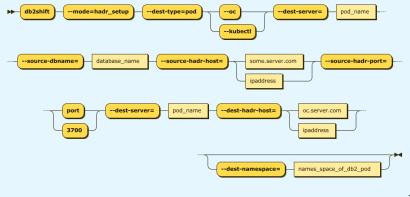
Db2 Shift [/home/c2	2cdb2/c2c]
	Deploy a Clone Copy into a Db2 Instance flights_ db2inst1 /tmp/cache
Database Name Instance Owner	lls : ● Remote ○ Local : flights_ ○ Force Database Creation : db2inst1 : some.server.com
HADR Option Exclude Routines Thread count Compression	
ESC•Quit ^?•Field H	Help ^D+Select Directory ^A+Analyze ^X+Review and Execute



Initialize HADR to a POD

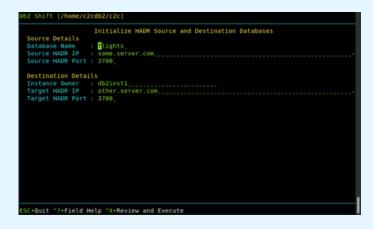
- This Db2 Shift option will take a source and destination (POD) database and start the HADR service between them
 - The Db2u pod must have been created with the HADR option during the shift step
 - Note that a destination HADR port is not required since it is automatically generated for you by the Db2 Shift program
- The command requires the following information:
 - The source database name and server
 - The destination POD and server details

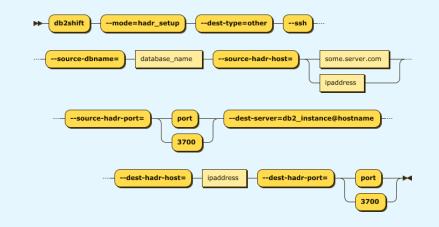
Source Details Database Name Source HADR IP Source HADR Port	
Destination Detai	ls
Target Location	: • OpenShift • Kubernetes
POD Name	
Target HADR IP	
	ielp ^X-Review and Execute



Initialize HADR to an Instance

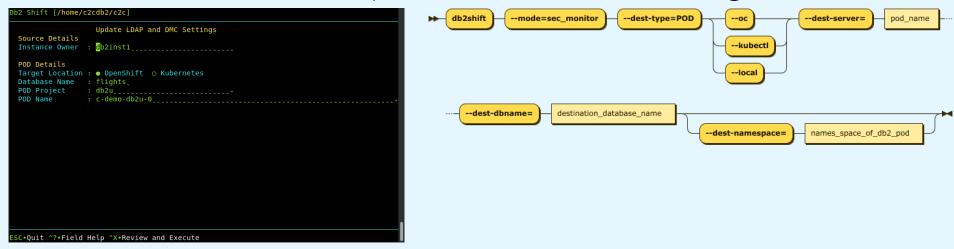
- This menu is similar to the previous one where the HADR service is setup between the source Db2 database and another Db2 instance
 - The Db2 database on the target system must have been created with the HADR option during the Shift operation
- The command requires the following information:
 - The source database name and server
 - The destination server details





Initialize LDAP and DMC (CP4D Only)

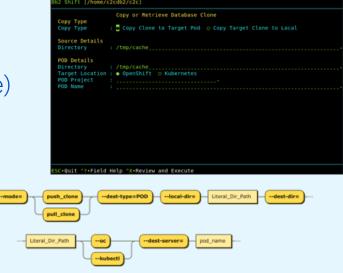
- Db2 Shift will automatically add the appropriate userids to the LDAP service and reset the Data Management Console (DMC) when shifting to CP4D
 - If you choose to shift a database and set it up as an HADR secondary, the LDAP and DMC setup cannot be performed until after you have converted the secondary POD into the primary server
- You must supply the name of the database, POD name and project as well as the name of the instance owner to update the LDAP and DMC settings



34

Copy a Clone to a POD

- This Db2 Shift command provides a feature that allows a user to copy an existing database clone copy to a POD, or to retrieve a database clone from a POD
- Once a database clone has been generated, the copy can be moved to any location and then deployed locally
 - This option provides a convenient way of copying the database using Db2 Shift without having to use OpenShift or Kubernetes commands
- The panel requires the following information:
 - Type of copy (From Source to Target or Target to Source)
 - Source cloned database directory
 - Target cloned database directory
 - The destination POD and server details

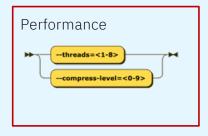


names_space_of_db2_po

35

Db2 Shift Performance

- There are two settings which will impact the performance of a Shift operation:
 - Threads the amount of parallelism that Db2 Shift will use
 - Compression How much compression will be applied to the data stream during shifts
- The impact of these two parameters can be substantial
- Deciding what setting to use is based on network speed, CPU capacity and overall load on the source system



Threading • Syntax: ► --threads=[1...8] ► Thread count : 4

- The copy phase of the Db2 Shift program can use multiple threads to transmit data to a destination
- This setting allows you to increase the parallelism up to 8 threads
- As you increase the number of threads, the amount of data being transmitted increases, at the expense of greater CPU usage and network congestion
- The default value is 4 which strikes a balance between overhead and network performance

Compression

- Syntax: ► --compression=[1...8] ►
 Compression : 4
- RSYNC compresses the data during the transfer process to allow for faster movement of data
- The amount of compression can be adjusted from 0 to 9 with 0 turning off compression and values between 1 and 9 increasing the amount of compression applied to the data
 - A value of 4 has been found to be a good compromise between compression overhead and data size on slow networks (< 1Gb/s)
 - For high-speed networks, a value of 0 is recommended unless there is a requirement to reduce network traffic
- Higher compression values will result in more CPU usage and may not significantly reduce the size of the data stream

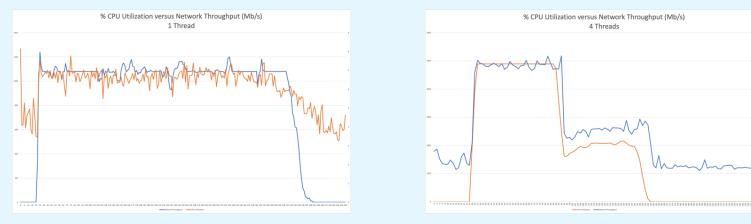
Elapsed Time

- 289M rows of transactional database with database size of 50Gb
- Best elapsed time 237 secs, with 131 sec for the data transfer
- Why 4 threads?



Combined CPU and Network

• CPU Usage is directly correlated to Network Throughput



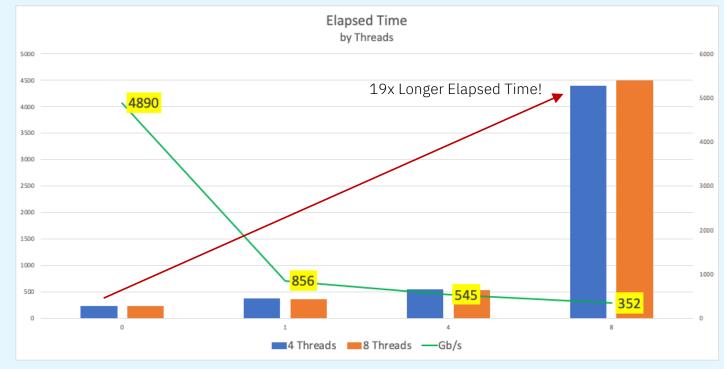
1 Thread



- Maximum network throughput
 - 1 thread was 1240 Mb/sec
 - 4 threads was 4890 Mb/sec
- Throughput was capped by processor speed and network limit of 5Gb/s

Compression

- The Compression setting has a negative impact on elapsed time
 - It will reduce the amount of network traffic but at the expense of elapsed time and CPU usage



Summary

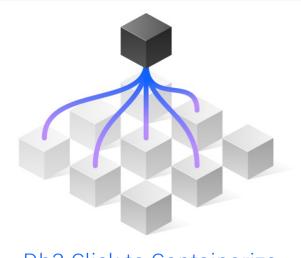
- Moves Databases to and from:
 - Normal Db2 Instance
 - Containerized Db2 Instance
- Support Multi-tenant Db2
- Supports offline Clones
- Inflight Version Upgrade
- Db2 Warehouse / OLTP
- Volumes = GBs to Terabytes
- Cloud or Local Deployments
 - IBM Cloud Pak For Data Deployments
 - Kubernetes
 - OpenShift
- Tunable Transportation

Offline Move

Seamless Failover-Zero Downtime



Resources



Db2 Click to Containerize

Visit <u>ibm.biz/c2cdownload</u> for more details on the Technical Preview of Db2 Click to Containerize

Documentation is available on <u>ibm.biz/c2cebook</u> (PDF) or <u>ibm.biz/c2cguide</u> (HTML)

Self-paced lab is available for business partners and IBMers at <u>ibm.biz/c2cv2lab</u>. Lab details are described in the following PDF <u>ibm.biz/c2ck8slab</u>

Single page fact sheet can be found at <u>ibm.biz/c2cfactsheet</u>

Thank you

© Copyright IBM Corporation 2022

IBM Corporation New Orchard Road, Armonk, NY 10504

Produced in the United States of America March 2021

IBM, the IBM logo, ibm.com, IBM Cloud, IBM Cloud Pak, Db2, InfoSphere, DataStage, Cognos and IBM Watson are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft 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.

Red Hat[®], and OpenShift[®] are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs. THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

Statement of Good Security Practices: IT system security involves protecting systems and information through prevention, detection and response to improper access from within and outside your enterprise. Improper access can result in information being altered, destroyed, misappropriated or misused or can result in damage to or misuse of your systems, including for use in attacks on others. No IT system or product should be considered completely secure, and no single product, service or security measure can be completely effective in preventing improper use or access. IBM systems, products and services are designed to be part of a lawful, comprehensive security approach, which will necessarily involve additional operational procedures, and may require other systems, products or services to be most effective. IBM DOES NOT WARRANT THAT ANY SYSTEMS, PRODUCTS OR SERVICES ARE IMMUNE FROM, OR WILL MAKE YOUR ENTERPRISE IMMUNE FROM, THE MALICIOUS OR ILLEGAL CONDUCT OF ANY PARTY