How to Setup Application Server to Access IBM DB2 z/OS with High Availability

Session Number IDZ 1702

Maryela Weihrauch,
Distinguished Engineer, IBM Silicon Valley Lab,
weihrau@us.ibm.com

IBM Software
Information On Demand 2011

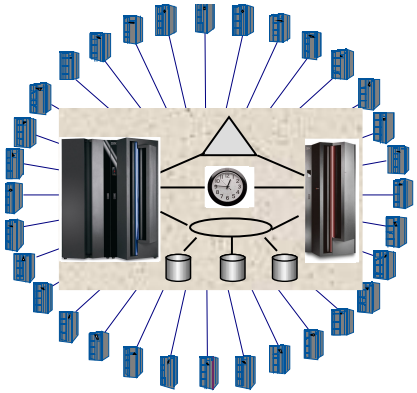
Agenda

- DB2 z/OS concepts for availability
- Client configurations
- DB2 Sysplex Workload Balancing
- Application server considerations
- DB2 10 enhancements

DB2 Data Sharing - Availability

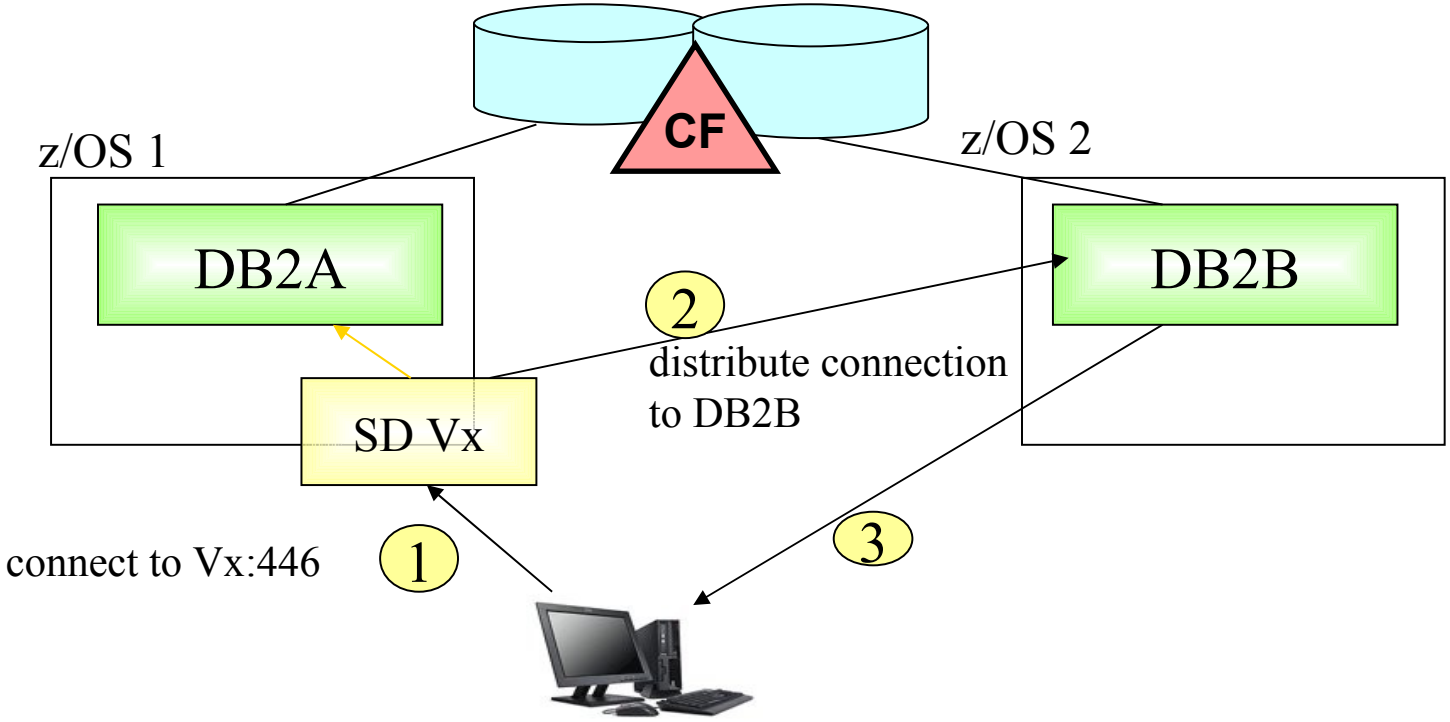
Goal: Continuous availability across any planned or unplanned outage across entire z HW/SW stack.

- Elimination of single point of failure (DB2 subsystem, z/OS system, CPC, I/O paths,..)
- Remove all causes for planned outages
  - rolling “on-line” upgrades and maintenance
  - online schema evolution
  - online utilities
- On a failure:
  - Isolate failure to lowest granularity possible
  - Automate recovery and recover fast



Challenge: the distributed application server needs to find the (best) available path to the data.

DVIPA and Sysplex Distributor – the Concept



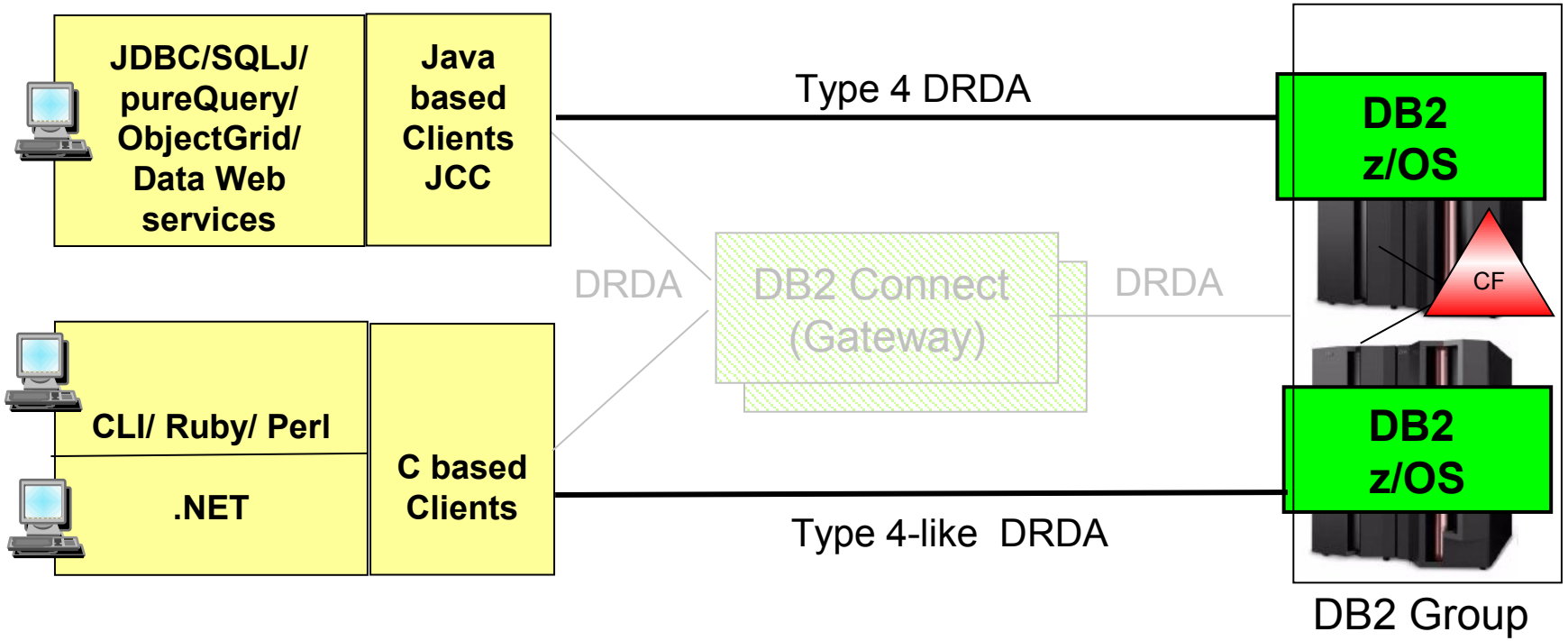
- Group DVIPA provides a virtual TCPIP address into the Data Sharing Group
- Sysplex Distributor routes the connection request to the most available member based on WLM recommendation

DVIPA and Sysplex Distributor – Consideration

- **Benefit:**
  - **Connections are successful as long as one member is up**
  - Connection level workload balancing between members
  - Setup is isolated to z/OS environment
- **Drawbacks**
  - SD on one lpar may route to a member on a different lpar, which results into slightly higher response time compared to direct member access.
  - Information about availability of data sharing members is only considered at creation of a "new" connection but application server typically maintain long-running connections.

✓ DVIPA and Sysplex Workload Balancing at the application server need to be enabled to ensure the highest availability characteristics in a distr. environment.

DB2 Client Configuration to Access DB2 z/OS – Strategy

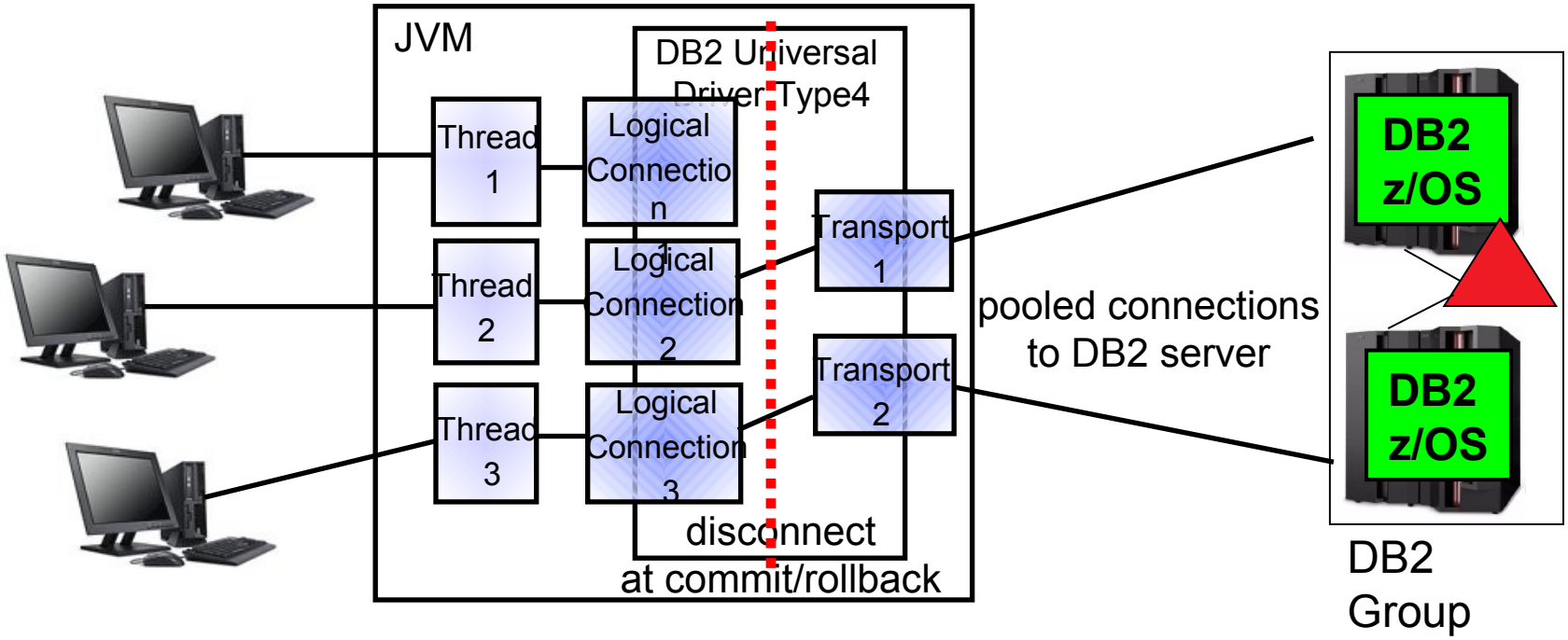


- Clients - especially application server - access DB2 z/OS directly (will not change licensing model)
- Sysplex Workload Balancing is supported by DB2 Connect gateway, JCC type 4, and Data Server Driver 9.5 FP3 for CLI and .NET application

JCC Type 4 Sysplex Workload Balancing

- **enableSysplexWLB**=YES
  - Enables Sysplex Workload Balancing
  - default is FALSE, disabled
- **maxTransportObjects** = -1
  - max # of concurrently used transports from this data source

  Can not be bigger than db2.jcc.maxTransportObjects
  - default value is -1, meaning no limit



JCC Type 4 Global properties defined in Global Properties File
It is Important to Set All!

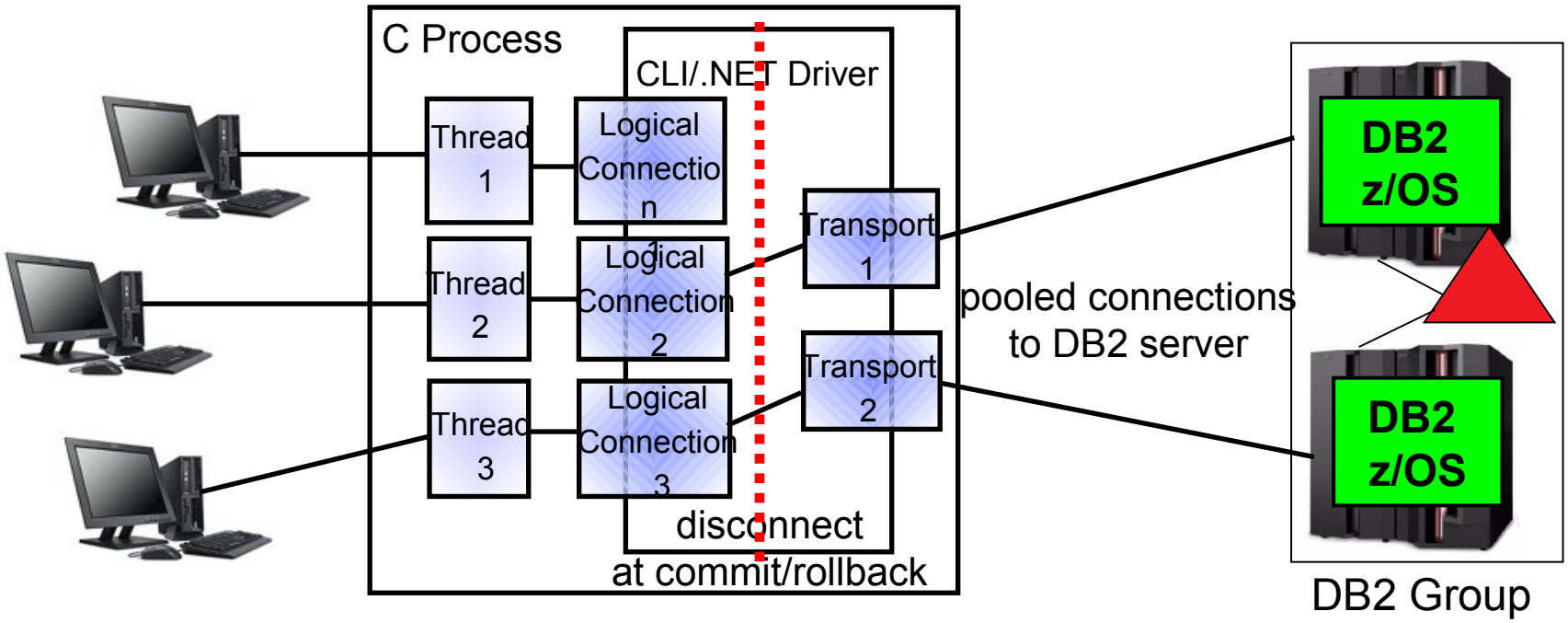
- **db2.jcc.maxTransportObjects**
  - max # of connections to DB2 server across all datasources
  - default value is -1, meaning no limit
  - Recommended: #steady state concurrent logical connections * # ds members
- **db2.jcc.maxTransportObjectIdleTime**
  - time in sec., a connection stays idle in the pool before it is closed,
  - default value is 60 sec, recommended
- **db2.jcc.maxTransportObjectWaitTime**
  - If maxTransportObjects is reached - time in sec., an application waits to get a connection before throwing a SQLException
  - default value is -1, unlimited wait
  - Recommended to set per response time agreement (e.g. 3 s)
  - Currently, do not set <3 s!
- **blockingReadConnectionTimeout**
  - Socket timeout value
  - Default value is -1, unlimited wait
  - Recommended: value greater by few seconds than the time it takes to execute the longest query in the application or response time agreement plus few sec

JCC Type 4 Global properties defined in Global Properties File
It is Important to Review and Set All!

- **maxRefreshInterval** in sec
  - Max time between refreshes of the WLM priorities
  - Default 30 sec – recommended except in very large environments
- **maxRetriesForClientReroute**
  - Maximum number of connection retries for automatic client reroute
  - Default -1 unlimited
  - Recommended to set to # of ds members
- **retryIntervalForClientReroute**
  - Number of seconds between consecutive connection
  - Default 0 s - recommended
- **loginTimeout**
  - only used for new connection
  - Recommended to set to 1 sec

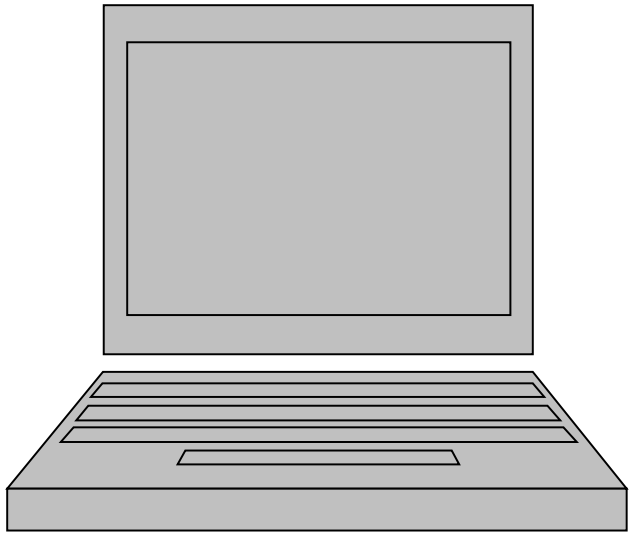
C-based Client Sysplex Workload Balancing

- Data Server Driver 9.5 FP3 for CLI and .NET supports Sysplex Workload Balancing
- Limited XA Support



XML Config. File – db2dsdriver.cfg

```
< configuration >
   < DSN_Collection >
      < dsn alias=, name=, host=, port= >
      < parameter name=, value= />
     < /dsn >
   < /DSN_Collection >
   < databases >
    < database name=, host=, port=,>
      < parameter name=, value= />
      < WLB >
        < parameter name=, value= />
      < /WLB >
      < ACR >
      < parameter name=, value= />
      < /ACR >
    < /database >
   < databases >
   < parameters >
     < parameter name=, value= />
   < /parameters >
< /configuration >
```



WLB Configuration

```
<databases>
   <database name="STLEC1" host="9.30.30.5" port="446">
      <WLB>
         <parameter name="enableWLB" value="true"/>
         <parameter name="maxTransports" value="100"/>
         <parameter name="maxTransportIdleTime" value="600"/>
         <parameter name="maxTransportWaitTime" value="30"/>
         <parameter name="maxRefreshInterval" value="30"/>
      </WLB>
       <ACR>
          <parameter name="enableACR" value="true"/>
          <parameter name="enableSeamlessACR" value="true"/>
       </ACR>
   </database>
</databases>
```

WLB Properties Comparison

Description	C-based client	JCC client
Enable Sysplex Feature	enableWLB	enableSysplexWLB
Maximum limit for the number of transports in transport pool	maxTransports	maxTransportObjects
Time that an unused transport stays in transport	maxTransportIdleTime	maxTransportObjectIdle Time
Time that an application waits for an available transport	maxTransportWait Time	maxTransportObject WaitTime
Max time in seconds between refreshes of the WLM priorities	maxRefreshInterval	maxRefreshInterval

ACR Properties and Reroute Timeout Comparison

Description	C-based client	JCC client
Enables automatic client reroute	enableAcr	
Enables seamless automatic client reroute	enableSeamlessAcr	
Maximum number of connection retries for automatic client reroute	maxAcrRetries	maxRetriesForClient Reroute
Number of seconds between consecutive connection	acrRetryInterval	maxTransportObject WaitTime
Max time in seconds between refreshes of the WLM priorities	maxRefreshInterval	retryIntervalForClient Reroute
Wait for a socket to a member when rerouting before attempting another member	MemberConnect Timeout	LoginTimeout

Fault Tolerant Sysplex Support

- Initial connection to DB2 data sharing group url
  - recommended to be a distributed DVIPA. This virtual IP address resolves to any available member based on WLM
  - or a DNS name that resolves into a list of IP address of each member that is used round robin
  - It always works if at least 1 member is up.
- The initial connection (or any ACCRDB responds) returns a member IP list with WLM weights
- Any following connection: uses client sysplex workload balancing algorithm to determine which member to use if reuse is OK:
  - No open WITH HOLD cursor
  - No declared global temporary tables must exist
    - (used declared global temp tables must be explicitly or implicitly dropped)
  - No reference to packages bound with KEEPDYNAMIC YES
  - Rollback always leaves a connection in a reusable state

Behavior at Connection Errors

- DB2 returns indicator as part of COMMIT responds if transport can be reused and returns SET statements to replay connection state at reuse
- **If first SQL stmt in transaction fails and reuse OK**
  - No errors reported back to application
  - SET statements associated with the logical connection are replayed with first SQL on another transport
- **If subsequent SQL fails and reuse OK,**
  - -30108 reuse error returned to application (transaction is rolled back and reconnected).
  - SET statements are replayed to recover connection state
  - Up to application to retry transaction
- **If subsequent SQL and reuse not OK**
  - -30081 connection failed error returned to application.
  - Connection returned to initial (default) state
  - application needs to reestablish connection state and retry transaction
- If all members in the member list are tried and none seems to be available, the initial data source url (DVIPA/DNS name) is retried to make sure that really no member is available.

WebSphere Connection Pool Properties

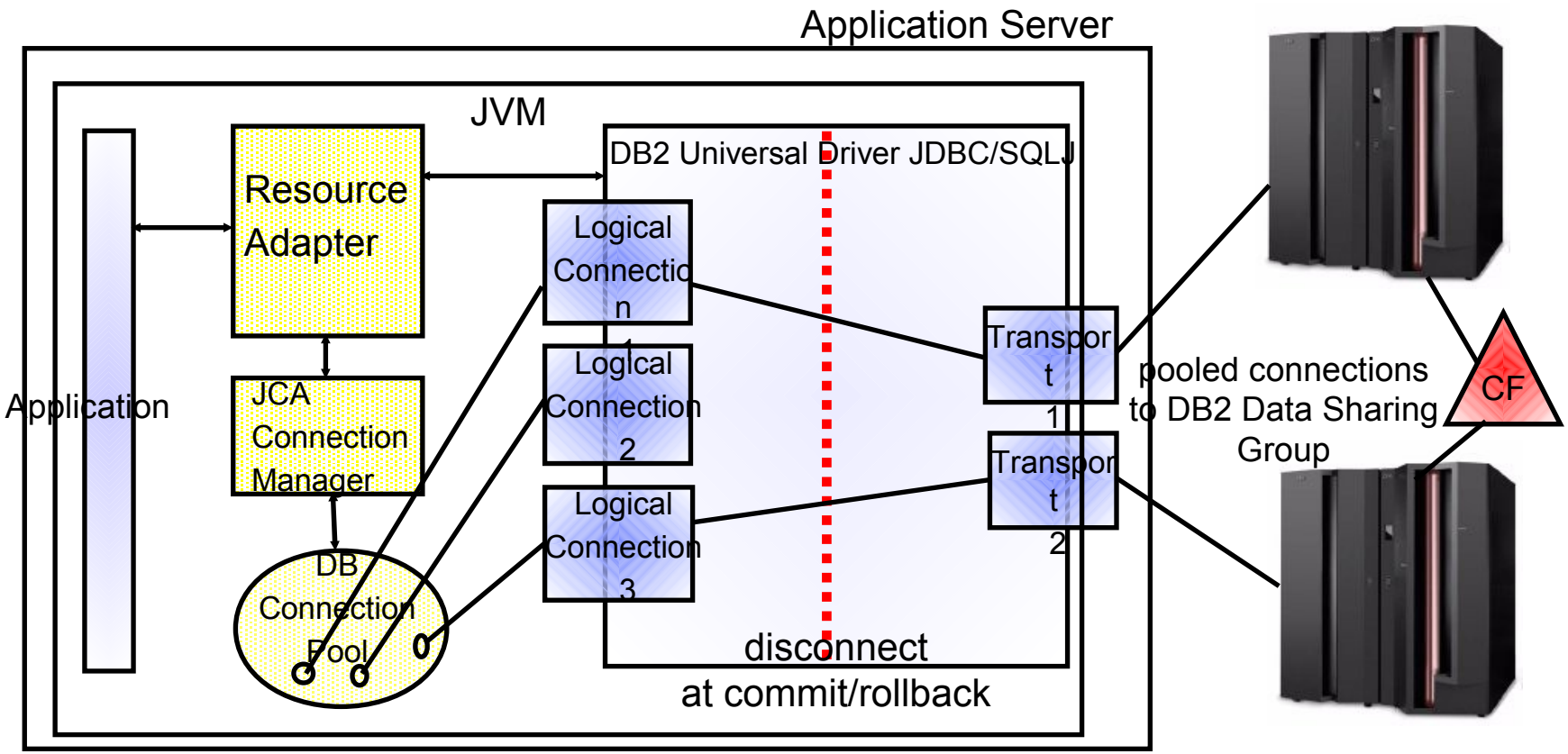
- Max Connections
  - max connections from JVM instance
- Min Connections
  - lazy minimum number of connections in pool
- Reap Time
  - How often cleanup of pool is scheduled in seconds
- Unused Timeout
  - How long to let a connection sit in the pool unused
- Aged Timeout
  - How long to let a connection live before recycling
- Purge Policy
  - After StaleConnection, does the entire pool get purged or only individual connection

WebSphere Connetion Pooling

- Connection object pool currently maintained by WebSphere
  - saves creating/destroying Connection objects which is relatively expensive
  - part of the Connection object creation is creating a physical connection to DB2
- Connection object holds references to other Java objects like preparedStatement objects that would be destroyed with the Connection object

WebSphere Application Server - End-to-end Example

- Always use connection pool closest to application
- Enable Sysplex Workload Balancing for best availability



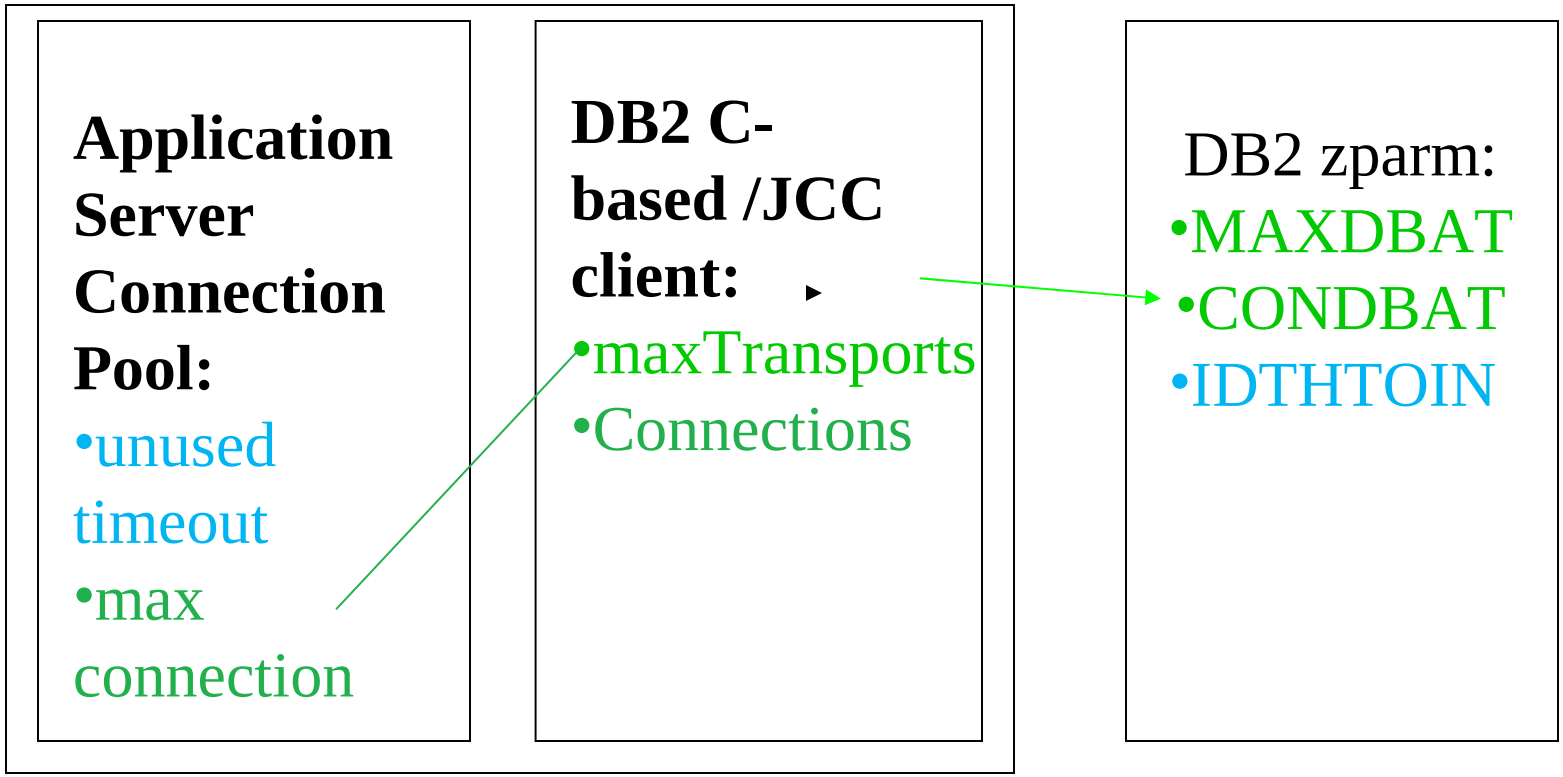
WAS/DB2 Active Thread – Tuning Considerations

- WAS connections in connection pool that keep an active thread in DB2 are target of the "idle thread timeout“
  - Type 2 on z/OS (RRSAF)
  - Type 4 connection that holds on to resources, e.g. WITH HOLD cursor, KEEPDYNAMIC YES.
- Set WAS "connection unused time" to a smaller value than DB2 "idle thread timeout" to avoid stale connection conditions.
- Consider setting min connections to 0 (zero) and connection unused time to not higher than 10 - 15 min to free up unused resources in DB2 in a controlled way and to reduce the exposure of long living threads
- In DB2 10 - CPU reduction with client package bind option RELEASE(DEALLOCATE)
  - Consider setting WAS “aged timeout” to less than 5 min, recommended 120 sec to reduce exposure of long living threads

Important DB2 zparm Values

- CONDBAT – Max. # of distributed connection into DB2 system
  - includes inactive and active connections, may be large because of many inactive connections
  - **Queues DBAT requests to become active up to CONDBAT**
    - **PM43293: Drain MAXDBAT queue zPARM for earlier reroute**
- MAXDBAT - Max # database access threads (DBATs) that can be active concurrently.
  - In many installations, max. value determined by available storage in DBM1 (check IFCID 225)
  - Set this value conservatively
- CMTSTAT INACTIVE – make a thread inactive after it successfully commits or rolls back and thread does not hold resources
  - prerequisite for sysplex workload balancing
  - inactive connections use less storage and free up DBM1 resources
  - Prereq for DB2 10 High Performance DBAT
- IDTHTOIN – time in sec an active server thread remain idle before it is canceled
  - inactive connections are not subject to idle thread timeout
  - Strongly recommended to not set to 0 – disable, default works well

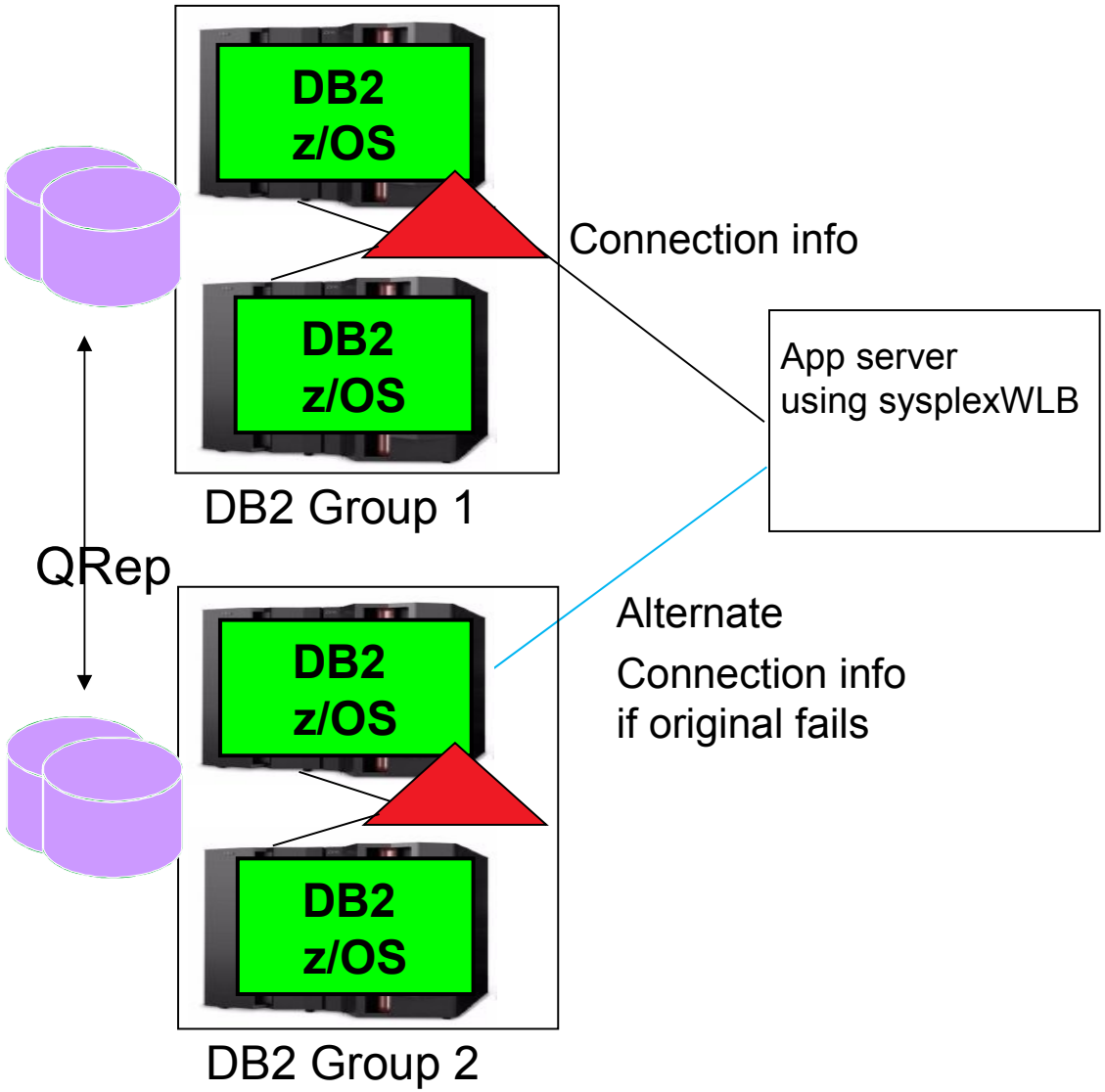
WAS/DB2 Inactive Thread - Tuning Considerations



Cancel Long Running SQL from Distributed Requestor

- Application Server cancel DB2 connections if responds does not come back in given time.
- Client properties that affect server availability
  - Recommended dropping connection, not interrupt SQL
    - Default client behavior not appropriate for z/OS
    - z/OS Lock manager is not listening to interrupt
    - Does not break out of long waits
  - C-based client
    - Set Interrupt=2 to always drop connection
    - SQLCancel to cancel SQL
  - Java client
    - Datasource property interruptProcessingMode = INTERRUPT_PROCESSING_MODE=CLOSE_SOCKET
    - Statement.cancel() to cancel SQL
- Investigate with z/OS dev. to support “cancel SRB” to interrupt distributed requests that execute in DB2 in a tight loop.
  - Currently, DB2 does not always immediately cancel work in DB2 so resources are burned unnecessarily

Client Failover to Alternate Data Sharing Group planned for 9.7 FP5



Primary DVIPA Group PG1

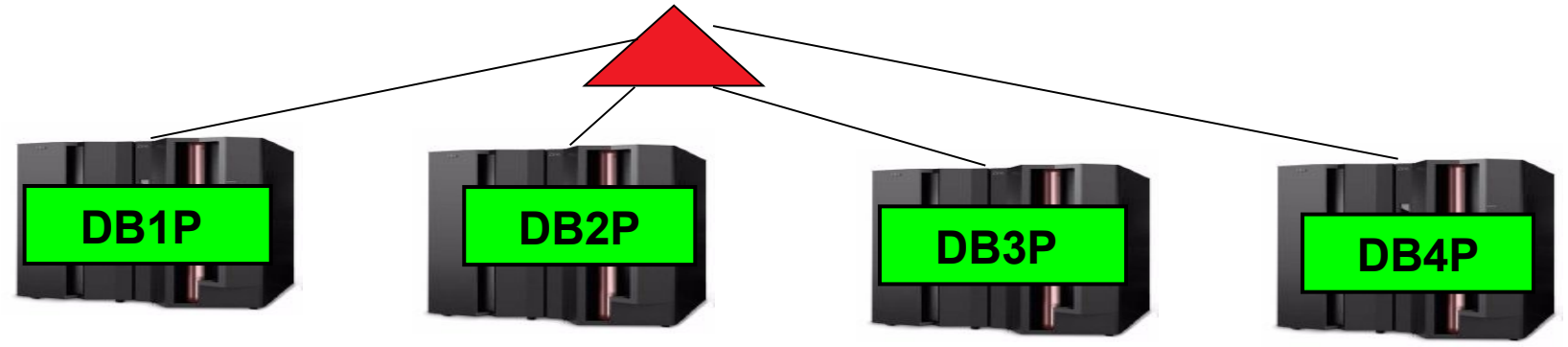
serverName = host1
portNumber = port1,
databaseName = dbname1

Alternate DVIPA group AG1

serverName = host2
portNumber = port2,
databaseName = dbname2

DS Group 2

DB2 10 Location Alias Enhancements



LocationAlias1
LocationAlias2

LocationAlias1

LocationAlias3

LocationAlias1
LocationAlias2

LocationAlias1

LocationAlias3

Application1	Application2	Application3
LocationAlias1	LocationAlias2	LocationAlias3
To all members	To DB1P, DB3P	To DB2P, DB3P

New –MODIFY DDF command with the ALIAS keyword to configure and manage aliases dynamically without taking a DB2 or DDF outage.

DB2 10 Location Alias Enhancements …

Examples :

-MODIFY DDF ALIAS(alias1) ADD
Alias1 is created and is stopped by default.
-MODIFY DDF ALIAS(alias1) PORT(9000)
Alias1 is associated with port 9000.
-MODIFY DDF ALIAS(alias1) IPv4(2.2.2.2)
Alias1 is associated with IP address 2.2.2.2
-MODIFY DDF ALIAS(alias1) START
DDF will accept requests for alias1 on port 9000. When a client connects to alias1, IP address 2.2.2.2 is returned in the server list.
-MODIFY DDF ALIAS(alias1) STOP
Alias1 is stopped and will not accept new requests. Existing requests will be allowed to complete.

DB2 10 High Performance DBAT

- **High Performance DBATs reduce CPU consumption by**
  - Supporting RELEASE(DEALLOCATE) to avoid repeated package allocation/deallocation
  - Avoids processing to go inactive and then back to active
  - Bigger CPU reduction for short transactions
- Targeted for users that do not use keepdynamic yes
- **High Performance DBAT behavior**
  - DBAT will stay active with connection when DBAT is about to be pooled and there is at least one RELEASE(DEALLOCATE) package still existing
  - Connections will turn inactive after 200 times (not changeable) to free up DBAT
  - Normal idle thread time-out detection will be applied to these DBATs. If DBAT is in completed unit-of-work status, Connection will turn inactive instead of being canceled
- **Default bind option for client packages changes from RELEASE(COMMIT) to RELEASE(DEALLOCATE) in 9.7 FP3**

DB2 10 High Performance DBAT…

- **New -MODIFY DDF PKGREL(BNDOPT/COMMIT) command**
  - to alter DDF's inactive connection processing which is activated via the ZPARM, CMTSTAT=INACTIVE.
  - 2 options
    - PKGREL(BNDOPT) honors package bind option
    - PKGREL(COMMIT) forces package bind option RELEASE(COMMIT)
      - Same as V9 inactive connection behavior
      - Will allow BIND and DDL to run concurrently with distributed work

Monitoring

- ***JCC trace*** ***http://publib.boulder.ibm.com/infocenter/db2luw/v9/index.jsp?topic=/com.ibm.db2.udb.apdv.java.doc/doc/c0020930.htm***
- ***DB2 command -display ddf detail***
  http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/index.jsp?topic=/com.ibm.db29.doc.comref/db2z_cmd_displayddf.htm

Reference

- DB2 9 for z/OS Data Sharing: Distributed Load Balancing and Fault Tolerant Configuration
  http://www.redbooks.ibm.com/abstracts/redp4449.html?Open
- DB2 Version 9.1 for z/OS Data Sharing: Planning and Administration (SC18-9845)
  http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/index.jsp?topic=/com.ibm.db29.doc.dshare/db2z_dshare.htm
- DB2 for z/OS best practice education
  https://www.ibm.com/developerworks/data/bestpractices/db2zos/

Information Management Communities

- **On-line communities, User Groups, Technical Forums, Blogs, Social networks, and more**
  - Find the community that interests you…
    - **World of DB2 for z/OS http://db2forzos.ning.com/**
    - **Information Management ibm.com**/software/data/community
    - **Business Analytics ibm.com**/software/analytics/community
    - **International DB2 User Group** www.idug.org

- **IBM Champions**
  - Recognizing individuals who have made the most outstanding contributions to Information Management, Business Analytics, and Enterprise Content Management communities
    - **ibm.com**/champion

Useful DB2 for z/OS URLs

- **DB2 10 Launch Website** → http://bit.ly/DB210Launch
- **DB2 for z/OS Website** → http://www-01.ibm.com/software/data/db2/zos/
- **DB2 Product Library** → http://www.ibm.com/software/data/db2/zos/library.html
- **DB2 Newsletter** → http://www.ibm.com/vrm/newsletter/11065

- **Latest Whitepapers**
  - Business Value of DB2 10 – Julian Stuhler
  - A Matter of Time: Temporal Data Management
  - Why DB2 for z/OS is BETTER than Oracle RAC ?
- **DB2 for z/OS e-Kit** → http://bit.ly/DB210e-Kit
- **Upcoming Conferences/Events**
  - **IDUG DB2 Tech Conference EMEA Prague - 13th -18th Nov 2011**
  - **DB2 10 Migration Planning Workshop Prague - 13th November 2011**
  - **IDUG DB2 Tech Conference AG Denver - 14th -16th May 2012**
  - **IDUG 10 Migration Planning Workshop Denver - 13th May 2012**

Top DB2 for z/OS e-Communities

- World of DB2 for z/OS - 1700+ members → http://db2forzos.ning.com/
- DB2 10 LinkedIn - 1000+ members → http://linkd.in/IBMDB210
- DB2 for z/OS What's On LinkedIn – 2000+ members → http://linkd.in/kd05LH
- DB2 for z/OS YouTube → http://www.youtube.com/user/IBMDB2forzOS
- WW IDUG LinkedIn Group - 2000 +members → http://linkd.in/IDUGLinkedIn
- IBM DeveloperWorks → http://www.ibm.com/developerworks/data/community/

Acknowledgements and Disclaimers:

Availability. References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.

The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.

Thank You!
Your Feedback is Important to Us

- Access your personal session survey list and complete via SmartSite
  - Your smart phone or web browser at: iodsmartsite.com
  - Any SmartSite kiosk onsite
  - Each completed session survey increases your chance to win an Apple iPod Touch with daily drawing sponsored by Alliance Tech
- Visit Us at

  www.ibm.com/software/data/db2/db210