DB2 Data Sharing Best Practices For Distributed Access

Shivram Ganduri
Senior Software Engineer, IBM

August 8, 2011
Session number : 9837

Important Disclaimer

Agenda

1. DB2 and network components of high availability
2. Sysplex Workload Balancing to a group
3. Sysplex Workload balancing to a subset.
4. Automatic Client reroute.
5. Benefits of z/OS WLM.
6. Sysplex awareness of clients.

High Availability Configuration for Remote Access

- TCP/IP Sysplex Distributor

  - Needs to be configured at the network layer and used for the initial connections to the DB2 group

- DB2 Sysplex Workload Balancing

  - Needs to be enabled on the drivers when accessing a DB2 group

- Load distribution managed by WLM

The network view of a Parallel Sysplex • a single large server with many network interfaces and many services like a DB2 data sharing group

My virtual z/OS IP host

VIPA#2 — DB2 Subsystem
VIPA#1 — TN3270e Server
VIPA#3 — FTP Services
VIPA#4 — DB2 subsystem
VIPA#6 — Web Services
VIPA#5 — CICS Appl-B

OSA IP#10   OSA IP#11   OSA IP#12

Connect to VIPA#1
Name server
Use IP address VIPA#2
Resolve DB2 Location
Connect to DB2 Location
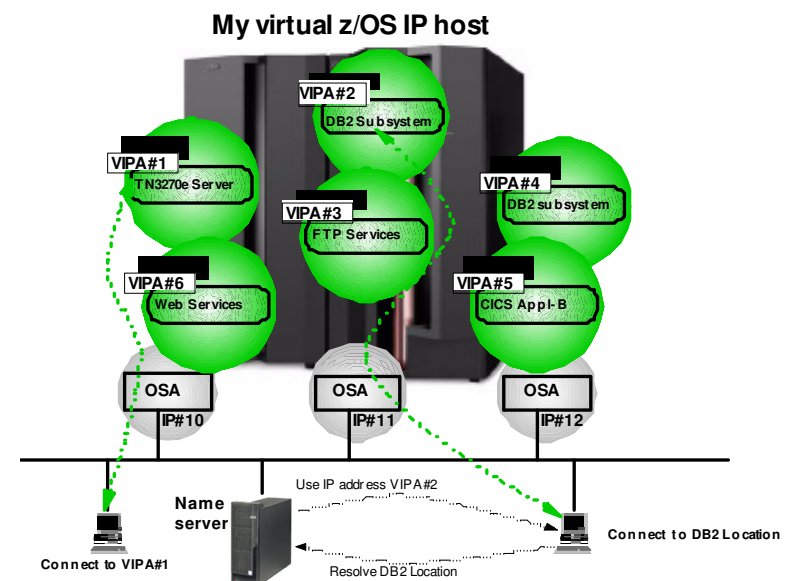
Balancing connections across group

- The TCP/IP Sysplex Distributor is used to establish the initial connection to the DB2 group

  - IP address used by clients to access DB2 is configured as a distributed dynamic VIPA

- Provides <u>connection</u> load distribution

- Ensures highest availability possible

- But…

  - Typical connections used for SQL have a long lifespan. For example, an application server that utilizes connection pooling…

Balancing transactions across group

- DB2 Sysplex Workload Balancing provides transaction level load distribution
  - Each DB2 member is configured with a unique dynamic VIPA and automatic VIPA takeover

- DB2 Sysplex Workload Balancing is enabled on the client.

- WLM ensures work is routed to the member with the most capacity and best health.

- Superior utilization of group resources

Configure network routing to the group

- Configure a group distributed DVIPA for the DB2 data sharing group.

  - All members listen to this IP address for the SQL port

  - Connections are distributed across all members

  - Connections are successful as long one member is up

  - Used by clients to access the group providing a single image

- Configure a unique member IP address for each subsystem

  - This member DVIPA is not distributed since used to route connections to a specific member

  - Allows routing even if a member fails over to another LPAR using VIPA takeover

  - WLM weight and member DVIPA are provided to the client for all subsystems registered

Network routing to the group

DB2 LOCATION

z/OS-1

DB2A

Vx, 446
V1, 446
V1, 5001

SD:Vx

② Dispatch connection to DB2B

① Initial connection to Vx, 446

④ Workload balancing connection to V1, 446

z/OS-2

DB2B

Vx, 446
V2, 446
V2, 5002

③ Resync V2, 5002
SRVLST (V1:W1,V2:W2, V3:W3)

z/OS-3

DB2C

Vx, 446
V3, 446
V3, 5003

④ Workload balancing connection to V3, 446
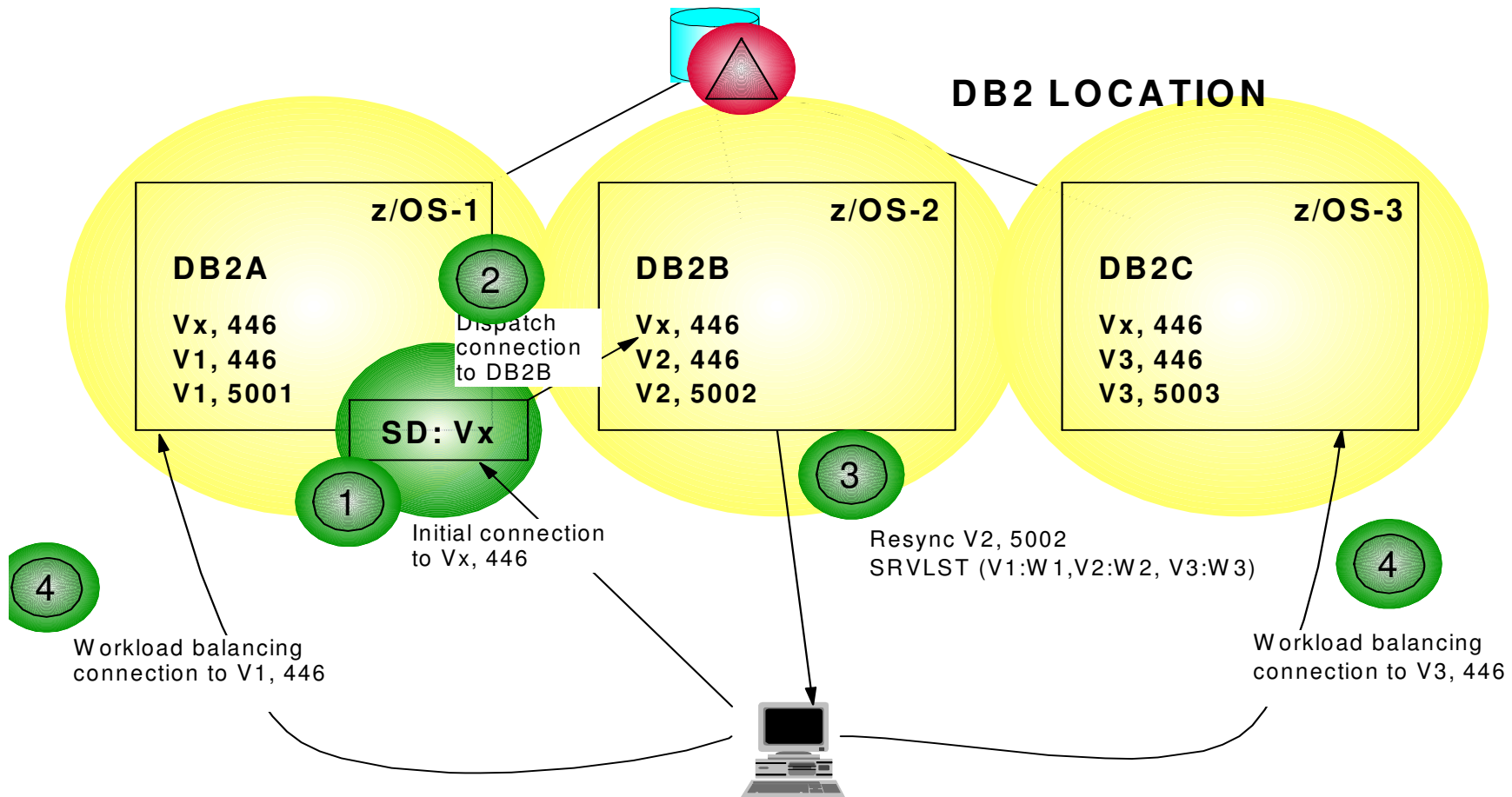
Configure using TCP/IP BIND keyword : Method 1

Member BSDSs:
>   DDF LOCATION=GROUP1,PORT=446,RESPORT=5001
>   DDF LOCATION=GROUP1,PORT=446,RESPORT=5002
>   DDF LOCATION=GROUP1,PORT=446,RESPORT=5003

TCP/IP PORT statement entries:
>   446 TCP DB2ADIST SHAREPORT BIND Vx
>   446 TCP DB2BDIST SHAREPORT BIND Vx
>   446 TCP DB2CDIST SHAREPORT BIND Vx
>   5001 TCP DB2ADIST BIND V1
>   5002 TCP DB2BDIST BIND V2
>   5003 TCP DB2CDIST BIND V3

TCP/IP VIPADYNAMIC entries:
>   **Member-specific DVIPA:**
>   VIPARANGE DEFINE 255.255.255.255 V1
>   VIPARANGE DEFINE 255.255.255.255 V2
>   VIPARANGE DEFINE 255.255.255.255 V3

>   **Group distributing DVIPA:**
>   VIPADEFINE 255.255.255.255 Vx ; defined on primary
>   system
>   VIPABACKUP n Vx ; defined on backup systems
>   VIPADISTRIBUTE DEFINE Vx PORT 446 DESTIP ALL

Configure using DB2 BSDS : Method 2

- **Member BSDSs:**

  DDFLOCATION=GROUP1,PORT=446,RESPORT=5001,IPV4=V1,GRPIPV4=Vx
  DDFLOCATION=GROUP1,PORT=446,RESPORT=5002,IPV4=V2,GRPIPV4=Vx
  DDFLOCATION=GROUP1,PORT=446,RESPORT=5003,IPV4=V3,GRPIPV4=Vx

- **TCP/IP Profile PORT statement entries:**

  446 TCP DB2ADIST SHAREPORT
  446 TCP DB2BDIST SHAREPORT
  446 TCP DB2CDIST SHAREPORT
  5001 TCP DB2ADIST
  5002 TCP DB2BDIST
  5003 TCP DB2CDIST

- **TCP/IP VIPADYNAMIC entries remain the same as in method 1.**

Balance to a subset of subsystems

- Managed load distribution across a subset of members

- Again, recommend the use of Sysplex Workload Balancing in conjunction with Sysplex Distributor
  - Configured with dynamic VIPA and automatic VIPA takeover
  - Connection successful if one member of subset started
  - A subset port is defined to distinguish between group and subset

- Connections and load distribution occurs only to the members listening on the subset port

- Connection failures can occur even when other members outside of the subset are up

- Utilizes only a subset of the group resources

Network routing to a subset
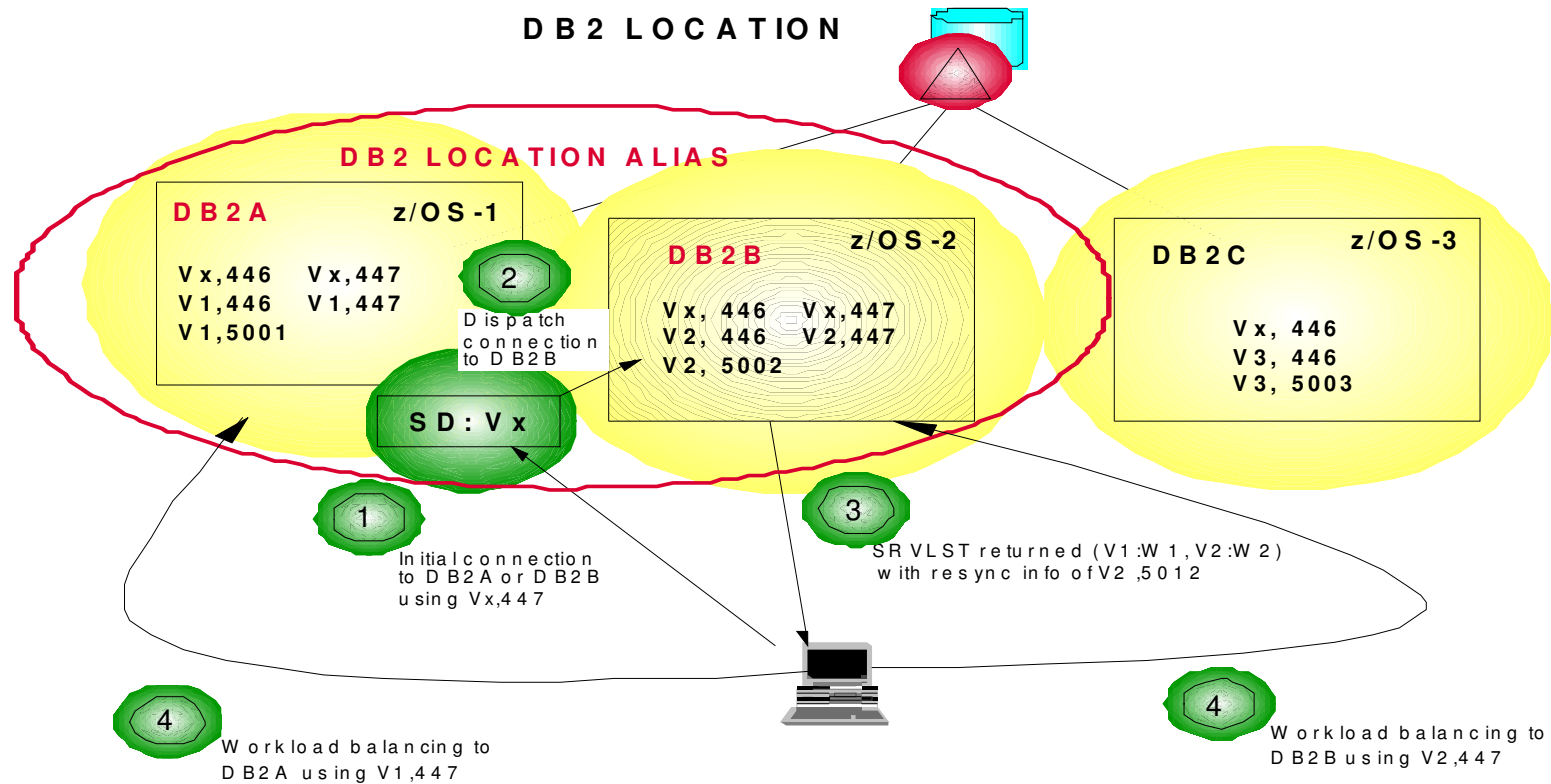
Configure using TCP/IP BIND keyword : Method 1

- Member BSDSs for subsetting using alias port 447

    DDF LOCATION=GROUP1,PORT=446,RESPORT=5001,ALIAS=ALIAS1:447
    DDF LOCATION=GROUP1,PORT=446,RESPORT=5002,ALIAS=ALIAS1:447
    DDF LOCATION=GROUP1,PORT=446,RESPORT=5003

- TCP/IP PORT statement for subsetting : Same as that for the Group. In addition, reserve the alias ports as follows

    447 TCP DB2ADIST SHAREPORT
    447 TCP DB2BDIST SHAREPORT

- TCP/IP VIPADYNAMIC statements for subsetting:

    - All the statements, except VIPADISTRIBUTE are the same as that for the Group.

    - Just add the alias port to the VIPADISTRIBUTE statement as follows:

        VIPADISTRIBUTE DEFINE Vx PORT 446 447 DESTIP ALL

Configure using DB2 BSDS : Method 2

- Member BSDSs for subsetting:

  DDF LOCATION=GROUP1,PORT=446,RESPORT=5001,IPV4=V1,GRPIPV4=Vx,
         ALIAS=ALIAS1:447
  DDF LOCATION=GROUP1,PORT=446,RESPORT=5002,IPV4=V2,GRPIPV4=Vx,
         ALIAS=ALIAS1:447
  DDF LOCATION=GROUP1,PORT=446,RESPORT=5003,IPV4=V3,GRPIPV4=Vx

- TCP/IP PORT and VIPADYNAMIC settings are the same as method 1.

Seamless automatic client reroute

- *Connection rerouting is possible because transaction pooling disassociates the logical connection from the transport on commit boundary. This allows the next transaction to be routed using a different transport that is connected to an available member of the data sharing group.*
- *The DB2 Connect clients (licensed IBM Data Server packages) give further capability of providing seamless client reroute, which is enabled by default when sysplex WLB is enabled, where the driver retries the transaction on another available data sharing member without notifying the application.*

Seamless automatic client reroute - continued

- *Seamless failover is enabled, by default when sysplexWLB is enabled, which is the recommended setting.*
- *If the first SQL stmt in a transaction fails either because the DB2 member is quiesced or due to any other error, then the transaction will be replayed on a different member without reporting errors back to the application.*
- *If the second or subsequent SQL fails, then SQLCODE of -30108 is returned to the application. The application will need to roll back the transaction and retry.*
- *If all the members in the server list are tried and none seem to be available, the initial data source url (Distributed DVIPA or a domain name that equates to it) is retried.*

Importance of z/OS WLM

- *WLM plays important roles in terms of using sysplex workload balancing (sysplexWLB). The sysplex WLB is based on a DRDA "server list", which consists of lists of IP address for DB2 data sharing member and "WLM server weight". The server list is created and maintained by WLM, and will be passed to DRDA client through DRDA flow during connect processing. When transaction pooling is used, DRDA clients and drivers will use the updated server list on each transaction to distribute workload to DB2 data sharing members.*

- *You need to know how your WLM policies are defined for your DRDA workload and how your DDF workloads meet the goal you defined, using RMF reports and server list information.*

- *In V10, you can monitor the weights of each DB2 data sharing member using the -DISPLAY DDF DETAIL command.*

Importance of z/OS WLM - Continued.

There are no configuration parameters related to enabling sysplex WLB on DB2 for z/OS server.
When DDF is started, it registers itself to WLM, unless MAXDBAT is set to 0 or DDF is stopped.
In either of these exception cases, that member's DDF does not appear in the server list.

There are several factors WLM will take into account when creating and updating weights.

1. Displaceable capacity of systems.
2. Enclave service class achievement (Performance Index, or PI)
   ➢ WLM goal should be attainable when system is not under stress, which would result in a PI < 1.
3. Enclave service class request queuing.
4. DB2 for z/OS Health
   ➢ DB2 will report its health factor of 0 to 100 to WLM based on the current storage consumption within the DBM1 and DIST address spaces.

Sysplex awareness of drivers

Drivers can exploit the sysplex as follows:

1. They can connect applications to a DB2 data sharing group as though it were a single database server, and spread the workload among the different members, based on server lists dynamically provided by WLM.

2. They can recognize when a member of a DB2 data sharing group fails and can automatically route new connections to other members.

Note : The drivers can establish direct connections to DB2 for z/OS bypassing DB2 Connect, although DB2 Connect license keys are still required.

Both the Java and non-Java based IBM Data Server drivers provide the sysplex awareness capability. Enabling this automatically enables transaction level balancing and automatic client reroute.

1. Java drivers - The property is called enableSysplexWLB.
2. Non-Java based drivers – The property is called enableWLB.