

# Distributed Processing with DB2 for z/OS – Advanced Topics

Hugh Smith ([smithhj@us.ibm.com](mailto:smithhj@us.ibm.com))

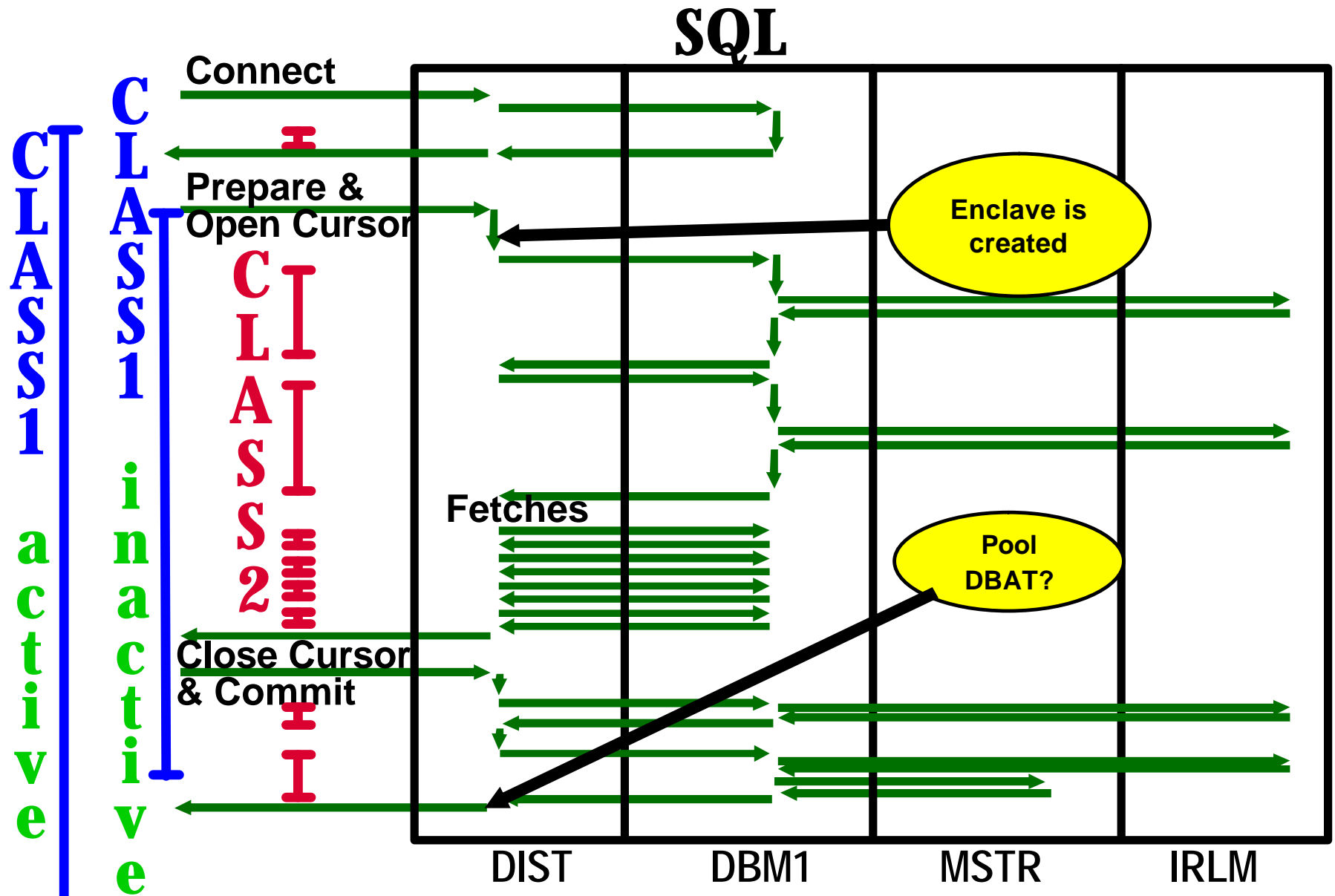
08/18/04 -- 1341

# Topics

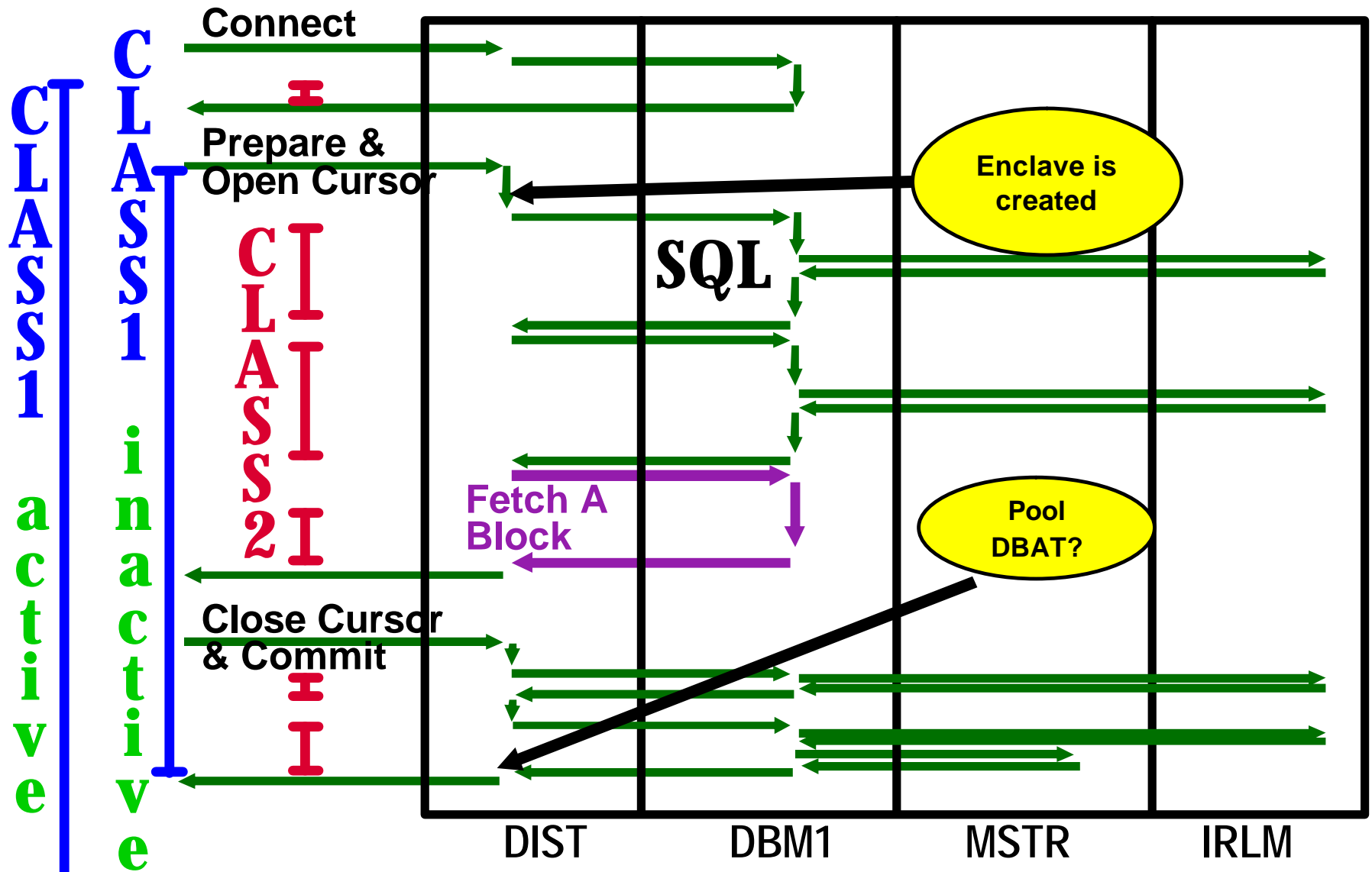
- Path of a Distributed Task
- What is inactive, pooled, etc.?
- Statistics and Accounting in a Distributed World
- Controlling DDF Work
- Data Sharing and Distributed
- DB2 and Dynamic Virtual IP Addressing
- DB2 Connect and Sysplex
- Subsetting a Data Sharing group



# Path of the Distributed SRB



# Path of the Distributed SRB (V8)



# DBAT Is Pooled When

- No WITH HOLD cursors are open

- f* temporary condition which can be resolved by client closing cursor and issuing another commit

- f* ODBC/CLI/JDBC/... client has a default of WITH HOLD

- f* Change default via db2cli.ini file set to CURSORHOLD=0

- No Declared Global Temporary Tables exist on the connection

- f* temporary condition which can be resolved by the application dropping the table and then issuing another commit

- No package (stored procedure, trigger, UDF, or nonnested task) with KEEP DYNAMIC YES bind option has been accessed

- f* permanent condition until thread is terminated or ROLLBACK has been issued



# Inactive/Pooling/What are we talking about?

- **INACTIVE** thread support originated in V3
  - f* Upon commit, a DBAT's memory footprint reduced and the DBAT goes inactive (-dis thd(\*) type(inactive))
  - f* New Unit-of-Work (UOW) request would cause DBAT to be returned to active set of DBATs and memory footprint expanded
  - f* Inactive DBAT tied to user's connection
- With V6, DRDA Connections now use Inactive Connection support
  - f* Upon commit, DBAT marked in DISCONN state (pooled) and connection becomes inactive
  - f* New UOW request from any user connection will cause DBAT in pool to be associated with connection
  - f* After 200 state switches, DBAT is purged
  - f* After POOLINAC of time in pool, DBAT is purged



## Inactive/Pooling/etc. ...

- Private protocol connections still use older inactive DBAT support
- MAXTYPE1 controls how many DBATs using private protocol can go inactive
  - $f0$  = any DBAT which uses private protocol will stay active (includes any DRDA DBAT which hopped out to another server via private protocol)
  - $fnnn$  = maximum number of DBATs using private protocol which can be inactive concurrently (DBAT/connection is aborted if number is exceeded)



# TCP/IP Performance Factors Summary

- Is the mainframe connected to network in full duplex mode? Are you sure?
- If delivering lots of data to client, maximize your TCP send buffer spaces/windows
  - f* on distributed clients utilize the DB2SOSNDBUF and DB2SORCVBUF registry settings, i.e. set to 65537 on Windows to get window scaling
  - f* TCPCONFIG on mainframe (DB2 sets them to a minimum of 64KB)
- Install PQ74041/UQ79222 in V7 to utilize tcp\_nodelay socket option
- How many router hops? i.e. message latency





# Other Performance Factors Summary

- CLI/ODBC/JDBC/VB/OLEDB/ADO/.NET standards have processing defaults which will cause extra message flows
  - f* AUTOCOMMIT - make sure applications turn it off and then issue commits where appropriate
  - f* CURSORHOLD - every cursor is requested WITH HOLD
    - prior to DB2 UDB for z/OS V7, set CURSORHOLD=0 in db2cli.ini
    - DB2 UDB for z/OS V8 will accept setting from DB2 V8 clients to early close cursors



# Statistics Report Observations

- Highlights section of DB2PM Statistics Report
  - f* Total Threads do not count any created distributed "server" threads
  - f* All per thread calculations do not factor in distributed "server" work
- DDF CPU time will have SRB processing when being a requester (SRB is TCP/IP or VTAM processing)
- DDF CPU time will not include any CPU for processing stored procedures
  - ▶ SMF 30 record for DDF address space will

# DDF Activity Statistics

GLOBAL DDF ACTIVITY	QUANTITY	/SECOND
-----	-----	-----
DBAT QUEUED-MAXIMUM ACTIVE	0.00	0.00
CONV. DEALLOC-MAX. CONNECTED	0.00	0.00
COLD START CONNECTIONS	0.00	0.00
WARM START CONNECTIONS	0.00	0.00
RESYNCHRONIZATION ATTEMPTED	0.00	0.00
RESYNCHRONIZATION SUCCEEDED	0.00	0.00
CUR TYPE 1 INACTIVE DBATS	0.00	N/A
TYPE 1 INACTIVE DBATS HWM	1.00	N/A
TYPE 1 CONNECTIONS TERMINAT	0.00	0.00
CUR TYPE 2 INACTIVE DBATS	5.00	N/A
TYPE 2 INACTIVE DBATS HWM	20.00	N/A
ACC QUEUED TYPE 2 INACT THR	3494.00	0.24
CUR QUEUED TYPE 2 INACT THR	0.00	N/A
QUEUED TYPE 2 INACT THR HWM	6.00	N/A
CURRENT ACTIVE DBATS	11.00	N/A
ACTIVE DBATS HWM	20.00	N/A
TOTAL DBATS HWM	24.00	N/A
CURRENT DBATS NOT IN USE	2.00	N/A
DBATS NOT IN USE HWM	17.00	N/A
DBATS CREATED	2.00	N/A
POOL DBATS REUSED	6986.00	N/A



# DDF Activity Statistics

## (Terminology the way I want it!)

GLOBAL DDF ACTIVITY	QUANTITY	/SECOND
-----	-----	-----
DBAT QUEUED-MAXIMUM ACTIVE	0.00	0.00
CONV. DEALLOC-MAX. CONNECTED	0.00	0.00
COLD START CONNECTIONS	0.00	0.00
WARM START CONNECTIONS	0.00	0.00
RESYNCHRONIZATION ATTEMPTED	0.00	0.00
RESYNCHRONIZATION SUCCEEDED	0.00	0.00
CURRENT INACTIVE DBATS	0.00	N/A
INACTIVE DBATS HWM	1.00	N/A
PRIV PROT CONNS TERMINATED	0.00	0.00
CUR INACTIVE CONNECTIONS	5.00	N/A
INACTIVE CONNECTIONS HWM	20.00	N/A
ACC QUEUED INACT CONNS	3494.00	0.24
CUR QUEUED INACT CONNS	0.00	N/A
QUEUED INACTIVE CONNS HWM	6.00	N/A
CURRENT ACTIVE DBATS	11.00	N/A
ACTIVE DBATS HWM	20.00	N/A
TOTAL DBATS HWM	24.00	N/A
CURRENT DBATS NOT IN USE	2.00	N/A
DBATS NOT IN USE HWM	17.00	N/A
DBATS CREATED	2.00	N/A
POOL DBATS REUSED	6986.00	N/A



# DRDA Remote Locs Statistics

DRDA REMOTE LOCS	SENT	RECEIVED
-----	-----	-----
TRANSACTIONS	0.00	0.00
CONVERSATIONS	0.00	0.00
CONVERSATIONS QUEUED	0.00	
SQL STATEMENTS	0.00	1459.2K
SINGLE PHASE COMMITS	0.00	442.7K
SINGLE PHASE ROLLBACKS	0.00	0.00
<b>ROWS</b>	<b>2047.8K</b>	0.00
<b>MESSAGES</b>	<b>1009.6K</b>	1009.5K
BYTES	703.6M	685.2M
<b>BLOCKS</b>	<b>350.4K</b>	0.00
MESSAGES IN BUFFER	2047.9K	

- Blocking: **rows** are put into **blocks** which are then sent out in **messages**



# Accounting Report/Trace Observations

- Processing "in DB2" (Class 2) should be the same regardless of connection type (almost, there is extra fetch column processing for DDF)
- Time in DB2 server is:
  - f* Class 2 nonnested elapsed time +
  - f* Class 1 stored procedure, UDF, and trigger elapsed time +
  - f* Nonnested (Class 1 CPU - Class 2 CPU)
- Time outside of DB2 server is total Class 1 elapsed less previous calculation
- Active thread accounting records are created at thread deallocation
- Inactive DBAT/Connection accounting records are created at DBAT inactive (look for DBAT inactive)



# Using Accounting Trace Information

			SQL DML	TOTAL
-----			-----	-----
AVERAGE	APPL(CL. 1)	DB2 (CL. 2)	SELECT	50974
-----	-----	-----	INSERT	30949
ELAPSED TIME	24: 59. 7823	3: 50. 24315	UPDATE	13029
NONNESTED	24: 59. 7823	3: 50. 24315	DELETE	1292
STORED PROC	0. 000000	0. 000000	DESCRIBE	0
UDF	0. 000000	0. 000000	DESC. TBL	0
TRIGGER	0. 000000	0. 000000	PREPARE	0
CPU TIME	1: 05. 93530	40. 218570	OPEN	5670
AGENT	1: 05. 93530	40. 218570	FETCH	15100
NONNESTED	1: 05. 93530	40. 218570	CLOSE	5670
STORED PRC	0. 000000	0. 000000		
UDF	0. 000000	0. 000000		
TRIGGER	0. 000000	0. 000000		
PAR. TASKS	0. 000000	0. 000000		
SUSPEND TIME	N/A	1: 18. 03509		
AGENT	N/A	1: 18. 03509		
PAR. TASKS	N/A	0. 000000		
NOT ACCOUNT.	N/A	1: 51. 98949		
DB2 ENT/EXIT	N/A	213306		
EN/EX- STPROC	N/A	0. 00		
EN/EX- UDF	N/A	0. 00		

	DML - ALL	122684
---	DISTRIBUTED ACTIVITY	-----
REQUESTER	:	10. 10. 10. 10
COMMITTS(1) RECEIVED:	:	4738
SQL RECEIVED	:	101914
MESSAGES SENT	:	106653
MESSAGES RECEIVED	:	106653
BYTES SENT	:	12943505
BYTES RECEIVED	:	21614937
MESSAGES IN BUFFER	:	9430
ROWS SENT	:	13326
BLOCKS SENT	:	5670

Time in DB2 = 3:50.24315+(1:05.93530-40.218570) = 4:15.95988

Time outside of DB2 = 24:59.7823-4:15.95988 = 20:43.82242



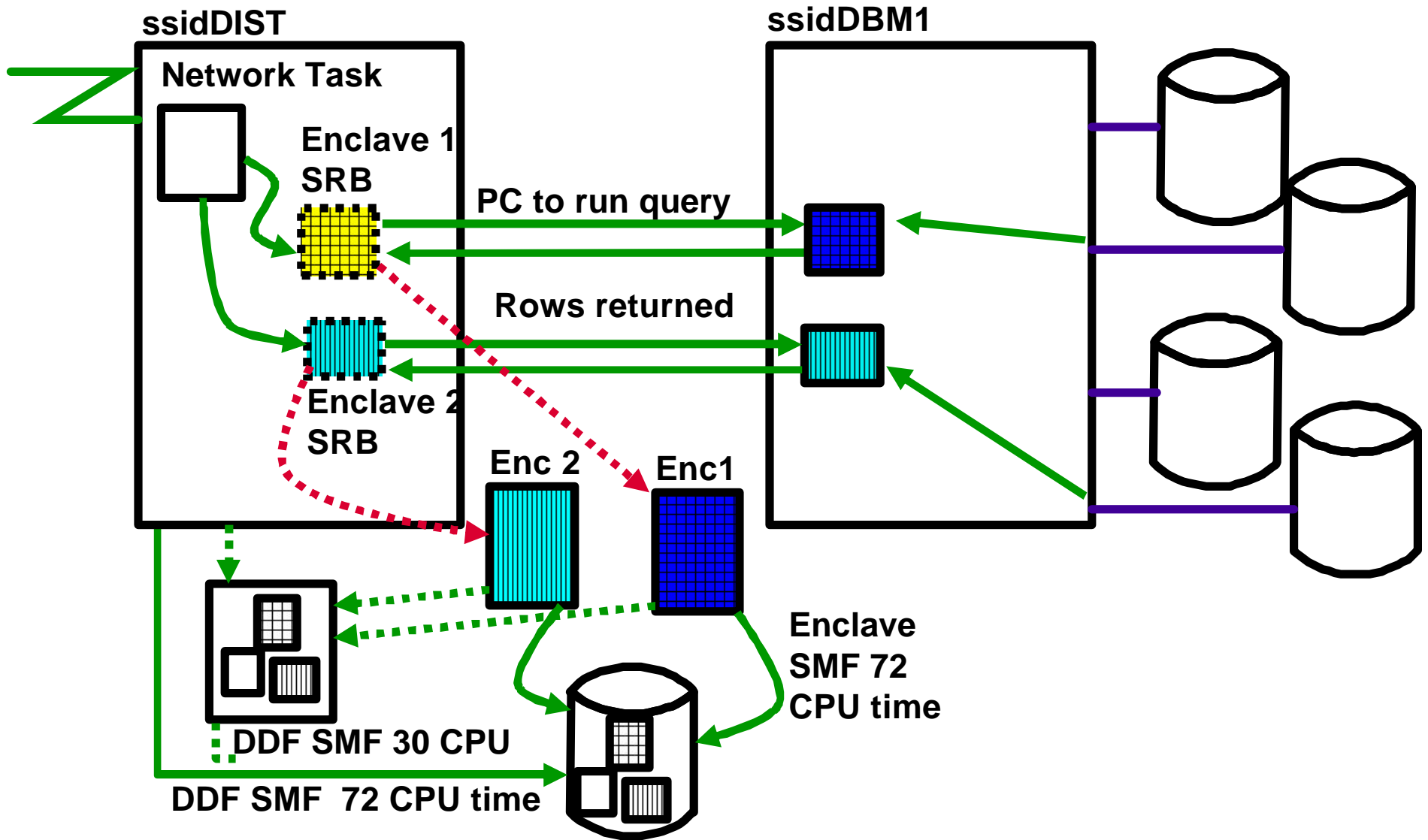
# Requester Accounting Report/Trace

- If TCP/IP connection, then time waiting for server is found in class 3 wait, Service Task Switch – Other Service
- If SNA or private protocol connection, then time waiting for server is included in NOT ACCOUNT.'d time while i DB2  
*f* NOT ACCOUNT. time can also mean time waiting for processor
- Time spent in communication stacks and/or network is the difference between the requesting DB2 time waiting for server value and the calculated time in DB2 server





# Managing DDF Work in DB2



# Classify DDF Transaction

- All enclaves, including DDF transactions, are classified using the active WLM policy in both compatibility and goal modes
- In compat mode the ICS can be used to map a service class to a performance group, and likewise for reporting



# DDF Work Classification Panel

Subsystem-Type Xref Notes Options Help

-----  
 Modify Rules for the Subsystem Type Row 1 to 1 of 1  
 Command ===> \_\_\_\_\_ SCROLL ===> PAGE

Subsystem Type . : DDF Fold qualifier names? Y (Y or N)  
 Description . . . DB2 Distributed Transactions

Action codes: A=After C=Copy M=Move I=Insert rule  
 B=Before D=Delete row R=Repeat IS=Insert Sub-rule  
 More ===>

Action	Type	Qualifier Name	Start	Class	Service	Report
_____	1	_____	_____	_____	DEFAULTS: MI SC_1	_____

\*\*\*\*\* BOTTOM OF DATA \*\*\*\*\*



# DDF Work Classification Attributes

Attribute	Type	Description
Accounting Information	<u>AI</u>	Can be passed from a DB2 Client via Client Information APIs
Correlation Information	<u>CI</u>	DB2 Connect assigns application program name by default but application can set via Client Information APIs
Collection Name	<u>CN</u>	Collection name of the first SQL package accessed by the DRDA requester in the unit of work
Connection Type	<u>CT</u>	Always 'DIST ' for DDF server threads
Package Name	<u>PK</u>	Name of the first DB2 package accessed by the DRDA requester in the unit of work
Plan Name	<u>PN</u>	Always 'DISTSERV' for DDF server threads accessed via DRDA requesters
Procedure Name	<u>PR</u>	Name of the procedure called as the first request in the unit of work
Process Name	<u>PC</u>	Client application name by default but can be set via Client Information APIs
Subsystem Collection Name	<u>SSC</u>	Usually the DB2 data sharing group name
Subsystem Instance	<u>SI</u>	DB2 server's MVS subsystem name
Sysplex Name	<u>PX</u>	Name assigned to sysplex at IPL
Userid	<u>UI</u>	DDF server thread's primary AUTHID
Subsystem Parameter	<u>SPM</u>	V8 - assigned the concatenation of client userid/workstation name



# Providing Workload Classification Attributes from Client

- ODBC/CLI/VB/ADO ... applications
  - ? Use SQLSetConnectionAttr on:
    - SQL\_ATTR\_INFO\_ACCTSTR - accounting string (AI)
    - SQL\_ATTR\_INFO\_APPLNAME - application name (PC)
    - SQL\_ATTR\_INFO\_USERID - client userid
    - SQL\_ATTR\_INFO\_WRKSTNNAME - client workstation name (V8 - last 2 will be concatenated and used for SPM)
- Non-ODBC... use sqleseti Administrative API function
- Universal Driver for Java applications (JCC T2 or T4)
  - ? Use methods against connection class instance
    - setClientUser, setClientApplicationInformation, setClientWorkStation, setClientAccountingInformation



# DDF Classification Defaults

- **Goal mode**

- f* Enclaves default to the SYSOTHER service class which has a discretionary goal!

- **Compat mode**

- f* Any enclaves run in the pgn/rpgn of the owning address space (DDF) as they did before

- Defaults apply if you do not provide any classification rules for DDF work

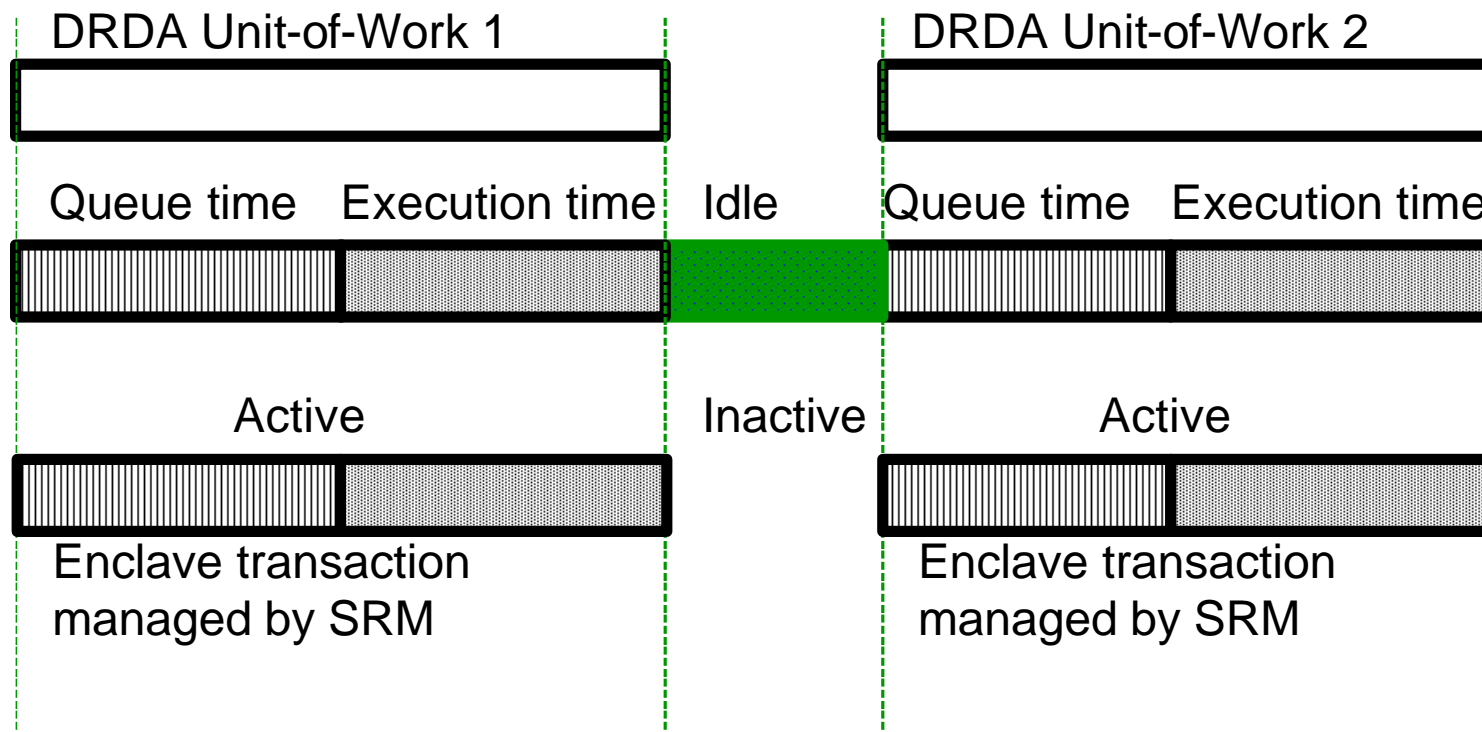


# Managing DDF Work (Enclaves)

- Transactions are subject to period switch
- Goal mode:
  - f* All goal types allowed
  - f* WLM manages enclaves with its own dispatch priority, etc.
- Compat mode:
  - f* Performance groups and report performance groups can be assigned using SUBSYS=DDF
  - f* Limit of one rpgn
  - f* Time slicing is ignored
  - f* Domain is ignored

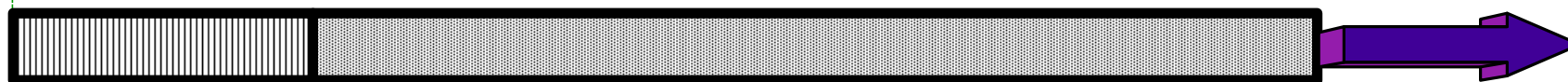
# What is a Transaction?

**Threads = Inactive**



**Threads = Active**

Database thread is active from creation until termination



Enclave transaction managed by SRM exhibits conversational behavior



# What Goals Should I Use?

- WLM compatibility mode

- f* Goal does not matter. Service class used only to associate an enclave with a performance group.

- WLM goal mode

- f* THREADS=INACTIVE and DBATs are pooled

- DDF creates one enclave per active interval
    - Response times do not include user think time
    - Response time goals and multiple periods can be used

- f* THREADS=ACTIVE

- DDF creates one enclave for the life of the thread
    - Enclave response time includes think time
    - Response time goals should not be used
    - Multiple periods should not be used



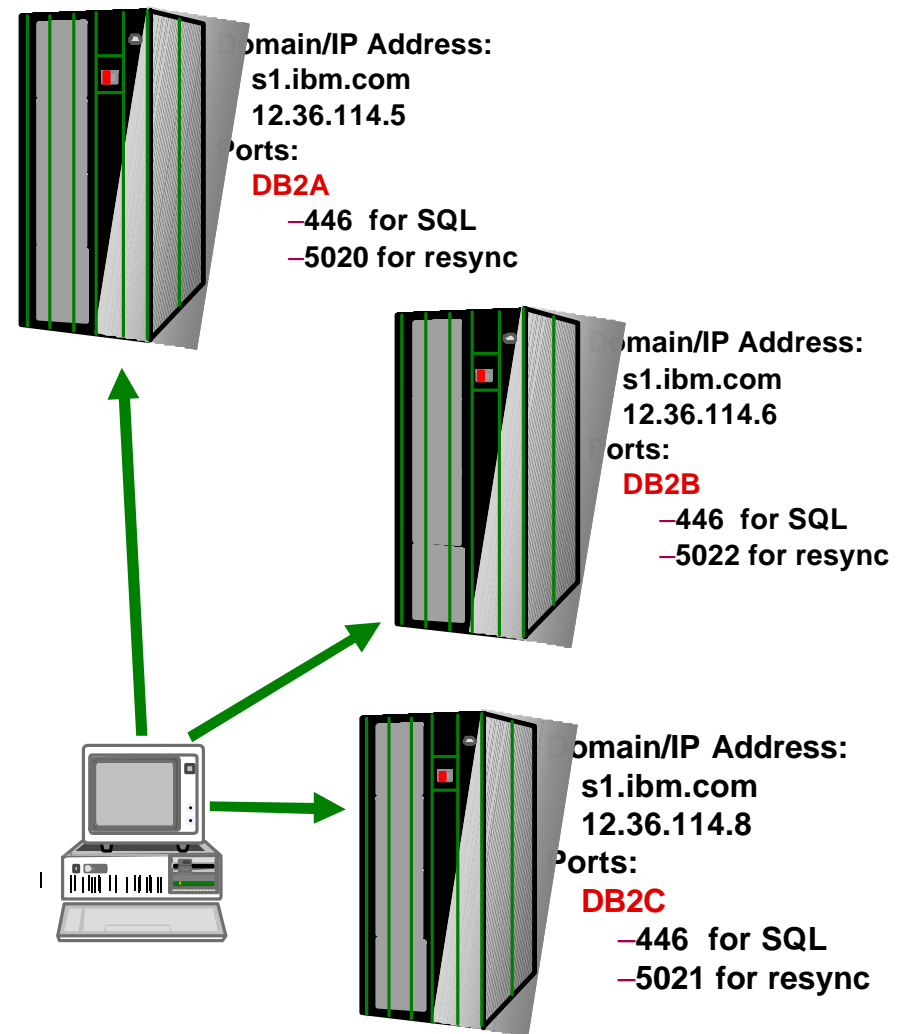
# Enclave System Effects (DDF)

- Each DDF transaction now seen by SRM, i.e. lots more transactions
- Increased active time since more transactions exist
- Large decrease in DDF SRB time/service in SMF 30 record
- Corresponding increase in CPU time/service in the service classes/performance groups where enclaves are running
- MSO and I/O service is unchanged
- DDF logons are not held up by existing DDF work



# Data Sharing and Distributed

- All members share same LOCATION, PORT, GENERICLU
- Each member has:
  - f* Unique RESPOR for resync work
  - f* Unique LUNAME
  - f* IP address can change after DB2 restart
- TCP/IP SHAREPORT PORT statement option allows multiple members on LPAR
- V6|V7 APAR PQ46659



# CDB Definition -- Data Sharing

LOCATION/Domain Name:

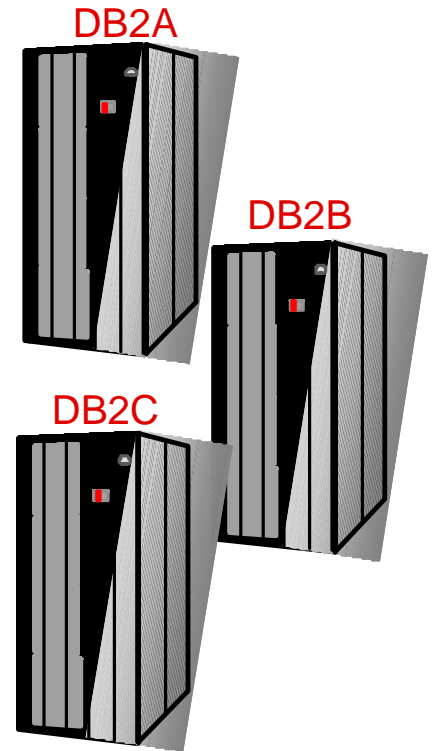
**DB2DSHR**  
s1.ibm.com

Ports:

**DB2A**  
-446 for SQL  
-5020 for resync

**DB2B**  
-446 for SQL  
-5021 for resync

**DB2C**  
-446 for SQL  
-5022 for resync



## SYSIBM.LOCATIONS

LOCATION	LINKNAME	PORT
'DB2DSHR'	'LINK1'	'446'

## SYSIBM.IPNAMES

LINKNAME	SECURITY_OUT	IPADDR	USERNAMES
'LINK1'	'P'	's1.ibm.com'	'O'

## SYSIBM.USERNAMES

TYPE	LINKNAME	NEWAUTHID	PASSWORD
'O'	'LINK1'	'DB2XUSR'	'DB2XPW'



# DB2 Connect - Data Sharing Example

- Create node profile:

```
f db2 catalog tcpip node db2dshr remote s1.ibm.com  
server 446
```

- Create DCS DB profile:

```
f db2 catalog dcs db db2dshrx as db2dshr parms  
',,,,,sysplex'
```

- Create DB alias profile:

```
f db2 catalog db db2dshrx as db2dshr at node db2dshr  
authentication server
```

- Can connect to any available member
- If using DNS hostname in node profile, each IP address from gethostbyname will be tried until success or all addresses exhausted



# Data Sharing, Dynamic VIPA, and Sysplex Distributor (PQ46659)

```

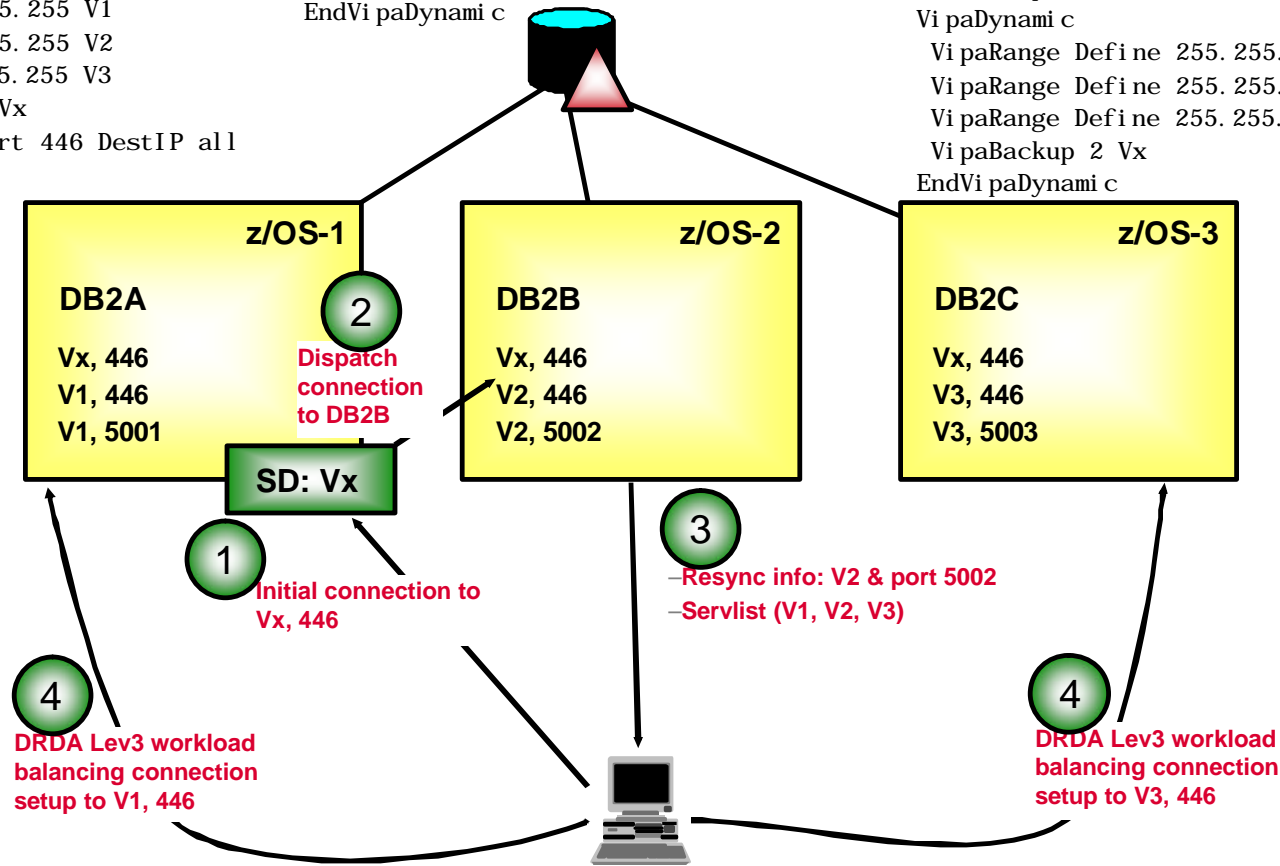
Port
  446 tcp db2adist shareport bind Vx
  5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
  5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
  5449 tcp db2cdist bind V3
Vi paDynami c
Vi paRange Define 255.255.255.255 V1
Vi paRange Define 255.255.255.255 V2
Vi paRange Define 255.255.255.255 V3
Vi paDefine 255.255.255.255 Vx
Vi paDi stribute Define Vx Port 446 DestIP all
EndVi paDynami c
  
```

```

Port
  446 tcp db2adist shareport bind Vx
  5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
  5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
  5449 tcp db2cdist bind V3
Vi paDynami c
Vi paRange Define 255.255.255.255 V1
Vi paRange Define 255.255.255.255 V2
Vi paRange Define 255.255.255.255 V3
Vi paBackup 1 Vx
EndVi paDynami c
  
```

```

Port
  446 tcp db2adist shareport bind Vx
  5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
  5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
  5449 tcp db2cdist bind V3
Vi paDynami c
Vi paRange Define 255.255.255.255 V1
Vi paRange Define 255.255.255.255 V2
Vi paRange Define 255.255.255.255 V3
Vi paBackup 2 Vx
EndVi paDynami c
  
```



# PQ46659 Requirements/Uses

- TCPIP stack must have SYSPLEXROUTING, DYNAMICXCF links, and datagram forwarding enabled
- Each DB2 member has its own IP address wherever it starts or is started in the sysplex
- Any available member will be reached if the dynamic VIPA address of the data sharing group used:
  - f* DB2 Connect PE and Univ. Java Driver T4 for any connection
  - f* DB2 Connect EE for first connection only



# DRDA Server List Support

- Architected request/response
  - f* Can be issued if implemented from any DRDA level 3 requester
  - f* DB2 only DRDA server to support server list request
- DB2 to DB2 already supported
  - f* Asks for server list during TCP/IP connections





# DB2 Connect EE Sysplex Support

- Catalog DCS entries with "sysplex" parameter  
*f* db2 catalog dcs db xyz as location parms ',,,,,sysplex'
- Client Configuration Assistant has support
- Local applications to DB2 Connect EE can use via registry setting:  
*f* db2set DB2CONNECT\_IN\_APP\_PROCESS=NO
- DB2SYSPLEX\_SERVER=0 disables



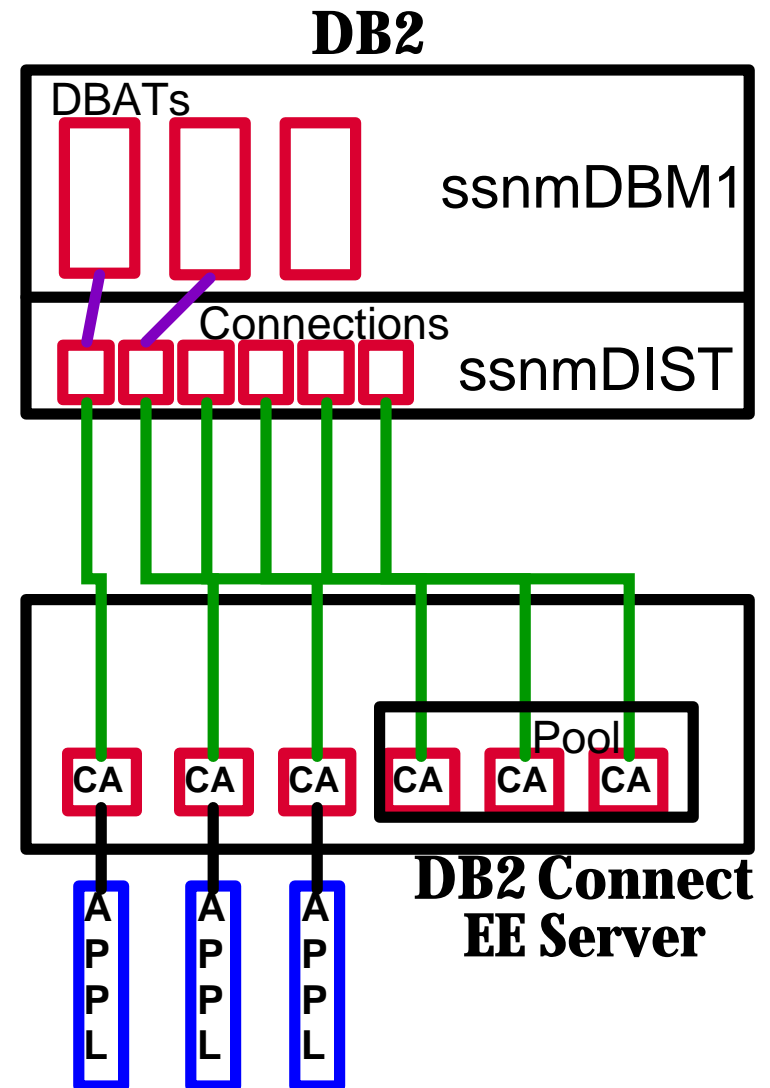
# DB2 Connect EE Sysplex Support (*cont'd*)

- New server list obtained during each "new" connection by a client
- DB2 asks WLM for server information during server list request processing
- Information lasts as long as server is up (db2start/db2stop)
- Connection time-outs may occur
- Connection errors can occur for in-flight transactions (pre-V8)



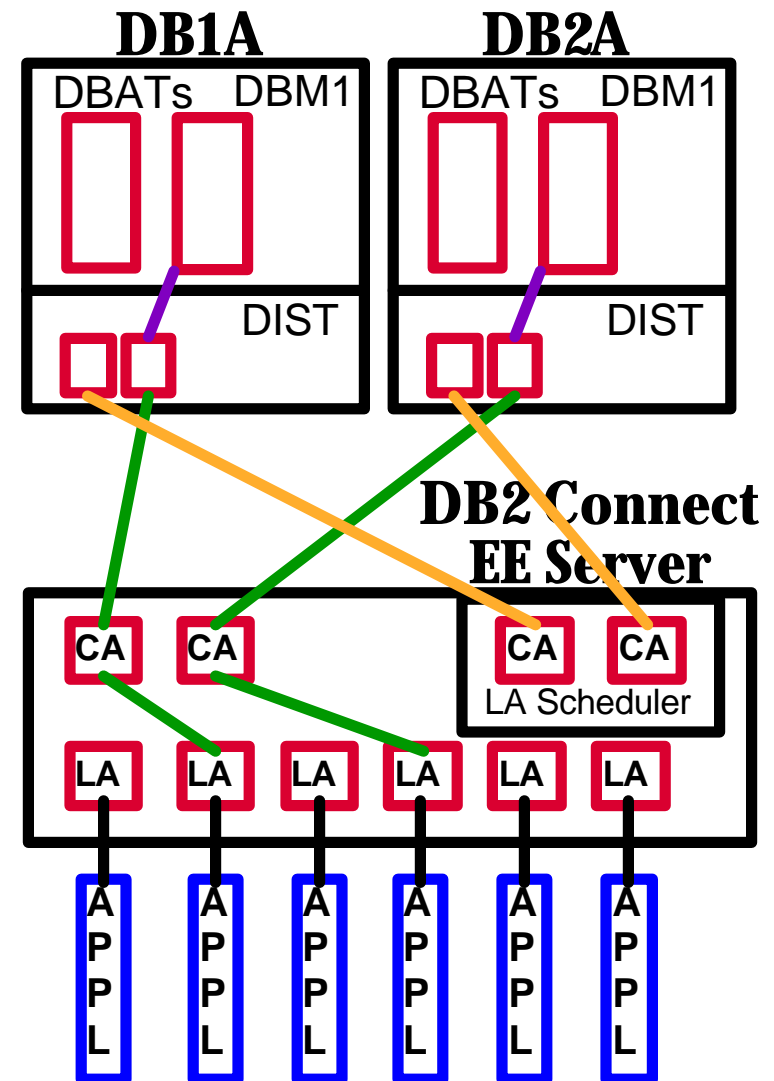
# DB2 Connect EE Connection Pooling

- Pooling is on by default
  - f* 2% of MAXAGENTS (different calc in V8)
  - f* Controlled by NUM\_POOLAGENTS
  - f* Can prestart pool agents via NUM\_INITAGENTS
- 1-1-1 relationship between applications, coordinating agents, and connections into DB2
- CA and corresponding host connection returned to the pool on disconnect
- Connections can come from any machine or process
  - f* DB2CONNECT\_IN\_APP\_PROCESS=NO for local processes/threads to utilize
- Connection pooled from any user with different attributes
- Resource-based management



# DB2 Connect EE Connection Concentrator

- Concentrator is not on by default
  - f* MAX\_CONNECTIONS > MAX\_COORDAGENTS
  - f* Make NUM\_POOLAGENTS = MAX\_COORDAGENTS
- N-1 relationship between applications and connections into DB2
- CA and corresponding host connection returned to the pool on commit/rollback
- Connection can be used by another application as soon as transaction completes
- Many different applications can reuse the same DB2 thread (DBAT)



# Connection Pooling/Concentrator & Sysplex Support

- First DB2 member in server list usually used to create next new connection
  - f* Connections initially spread across "available" members of DB2 data sharing group
- Reuse of existing connections to members goes through "best" member algorithm
  - f* As connections are reused, server list information is refined and connections rebalanced (even dropped when member should no longer receive any new work)
- Server list requested on connection reuse



# Connection Pooling/Concentrator & Sysplex Support (*cont'd*)

If just connection pooling:

- f* client is considered connected to DB2 member until disconnect by client
- f* for duration of connection, connection failures can be returned to client at any time

If connection concentrator:

- f* client is considered connected to DB2 Connect EE server
- f* connection failures only returned to client during an in-flight transaction with DB2 member (pre-V8)
  - for V8, EE returns -904 to client (changing to -30108)**
- f* reuse existing or create new connection determination occurs on new transaction from client
- f* controlled by max\_connections database manager configuration parameter



# Subsetting a Data Sharing Group (V8)

## SYSIBM.LOCATIONS

LOCATION	LINKNAME	PORT	DBALIAS
'DB2DSHR'	'LINKGRP1'	'446'	' '
'DB2AB'	'LINKGRP2'	'446'	' '

## SYSIBM.IPNAMES

LINKNAME	SECURITY_OUT	IPADDR	USERNAMES
'LINKGRP1'	'R'	's1.ibm.com'	' '
'LINKGRP2'	'R'		

## SYSIBM.IPLIST

LINKNAME	IPADDR
'LINKGRP2'	'db2a.s1.ibm.com'
'LINKGRP2'	'db2b.s1.ibm.com'

DB2AB is a LOCATION ALIAS defined at DB2A and DB2B.




What about non-DB2 UDB for z/OS requesters?

# Subsetting DB2 Data Sharing Groups (V8)

## From Any DRDA TCP/IP Requester

- Enabled via alias-port specification on ALIAS parameter of DDF BSDS statement

DDF ... ALIAS=  , ...

*f* specify same alias-name and **alias-port** combination on each member to participate in the same alias "subset"



# Subsetting DB2 Data Sharing Groups (V8)

## From Any DRDA TCP/IP Requester ...

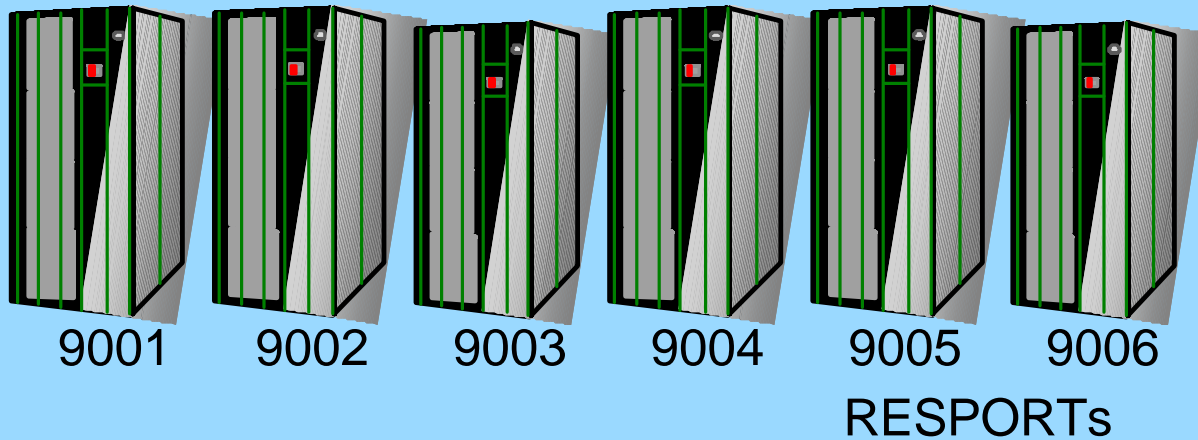
- DDF will perform the following actions on startup:
  - f* register with WLM its IP address and **alias-port** for each **alias-port** specified alias-name location group
  - f* extend the TCP/IP SQL listener to accept requests on each specified **alias-port**
    - if using sysplex distributor, initial connection requests to distributed DVIPA and **alias-port** will be sent to members participating in subset
- Connection to alias-name location and **alias-port** contains only "subsetted" server list



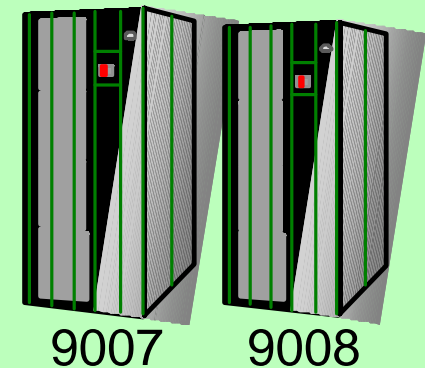
# V8 Subsetting Example

DDF LOCATION=DB0XGROUP,PORT=9000,

ALIAS=DB0XLEGACY:9010



ALIAS=  
DB0XNEW:9011



- Connect request to DB0XGROUP and port 9000 will be accepted by any available member
  - f* server list containing all available members of group will be returned
- Connect request to DB0XLEGACY/port 9010 or DB0XNEW/port 9011 will be accepted by only those members with alias/port definitions
  - f* server list containing only available members of targetted subset will be returned