9387: Setting up DB2 data sharing… the easy way

Jeff M. Sullivan
IBM systems and Technology Group Lab Services

Wednesday, August 10, 2011: 9:30 AM-10:30 AM
Europe 6 (Walt Disney World Dolphin Resort)

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon 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 given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Presentation topics

- ✓ What is it….really?
- ✓ Data sharing implementation example
- ✓ What works well

Why data sharing?

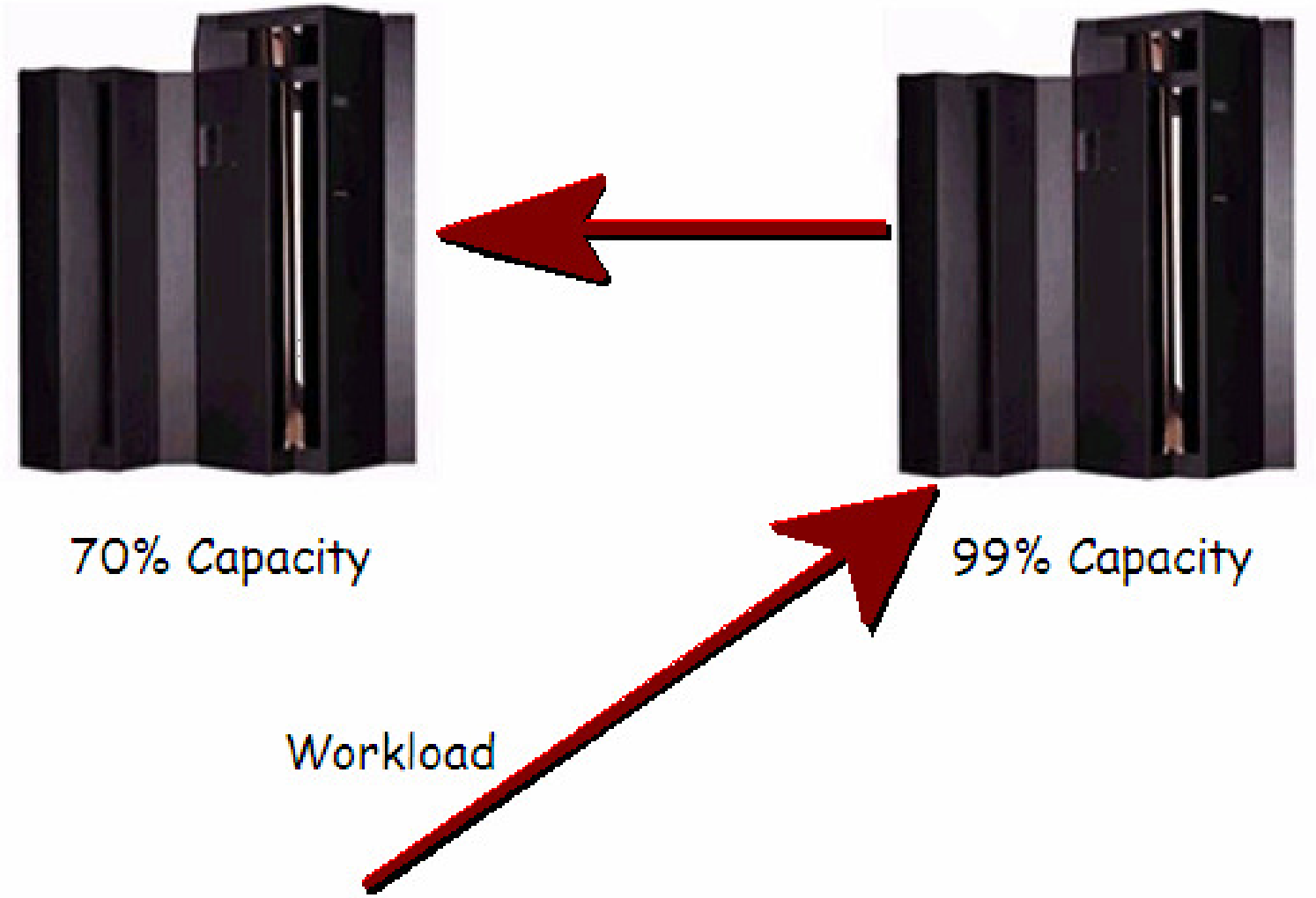
Most common reasons:

- Outgrow single system capacity
- Need higher availability
- Easier growth accommodation
- Dynamic workload balancing
- System consolidation for easier systems management

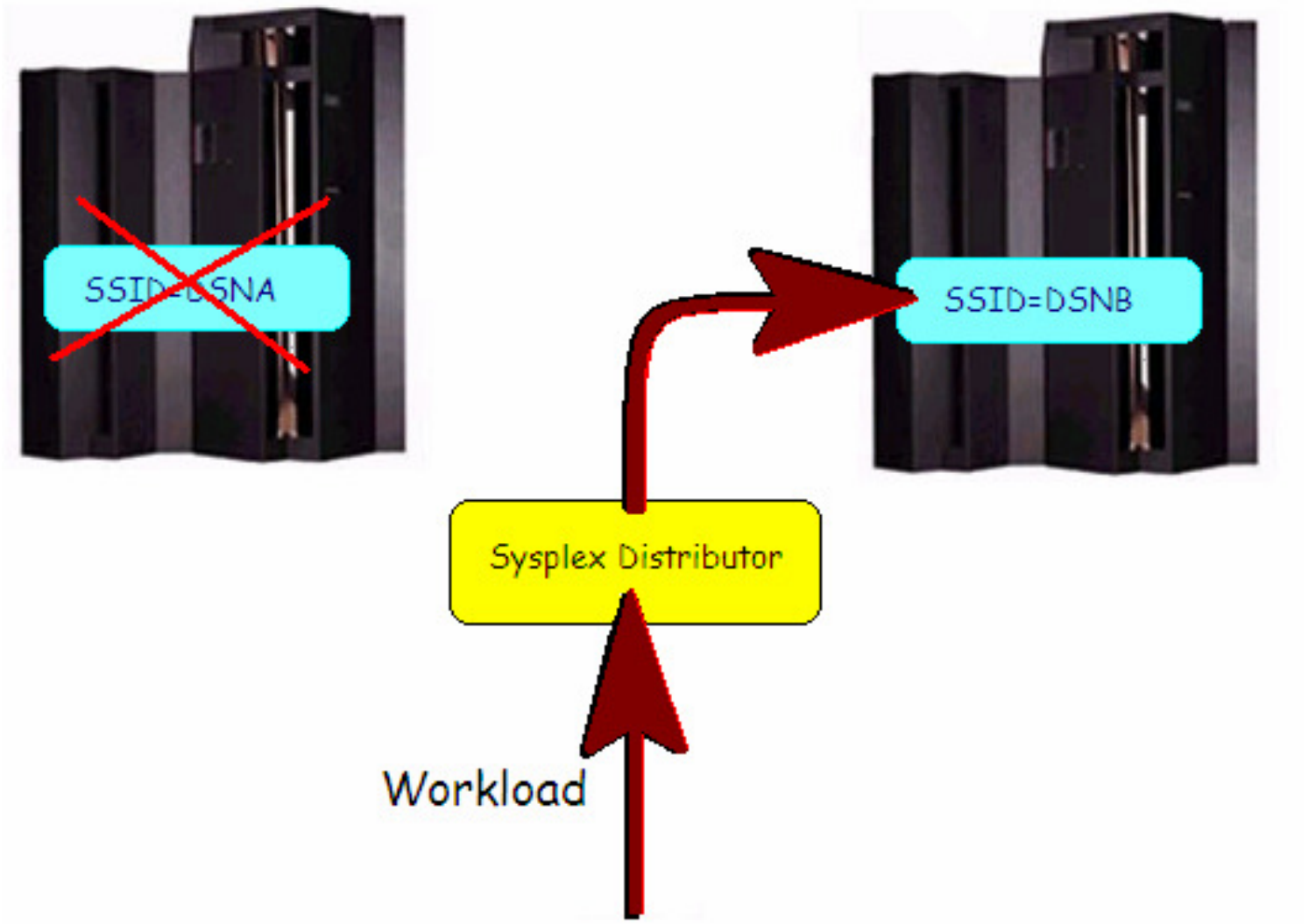
Cool feature

- Application interfaces require no changes

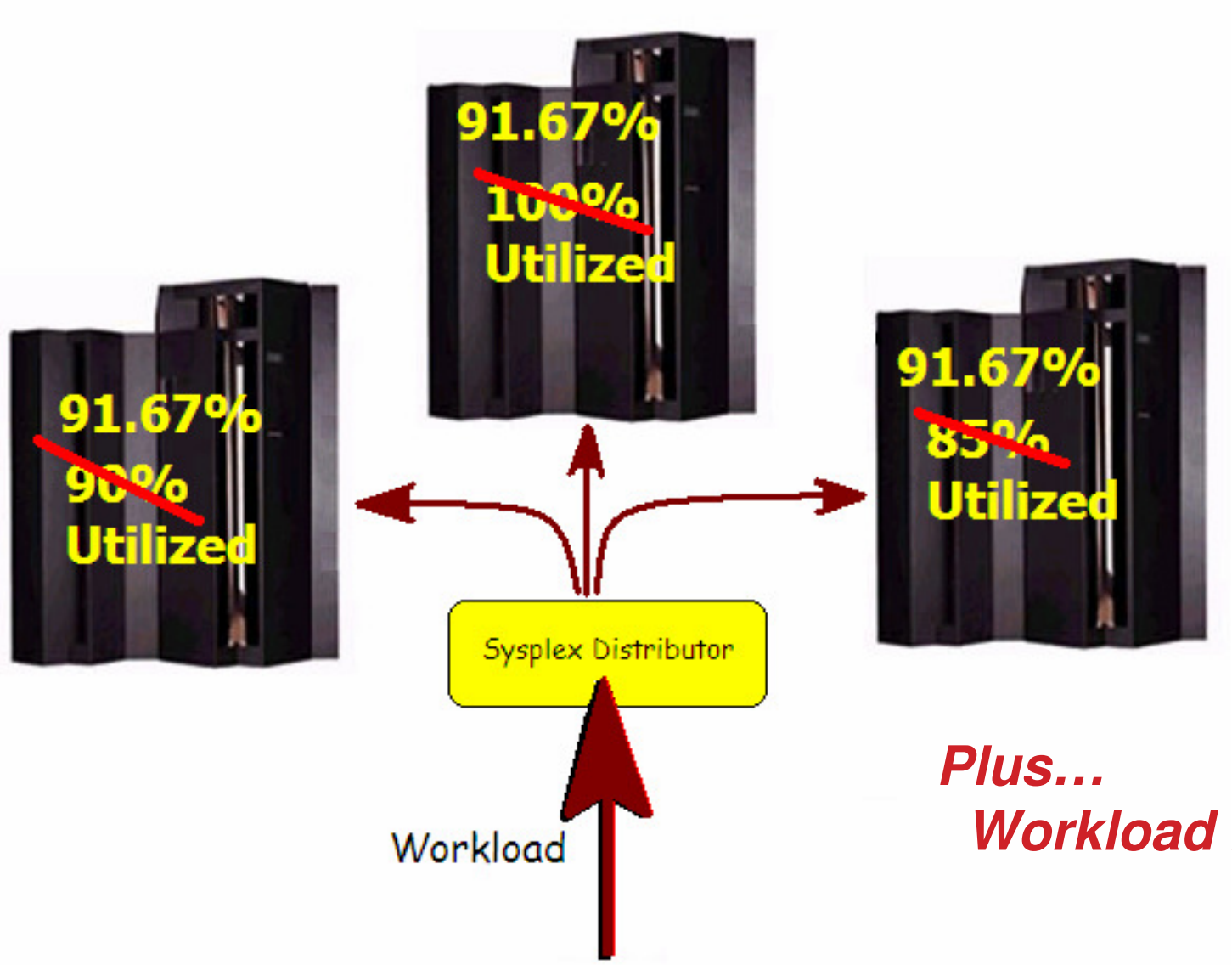
Availability



Continuous availability



Scalability



Plus…
Workload balancing

What it provides:

- Continuous availability
- Scalability

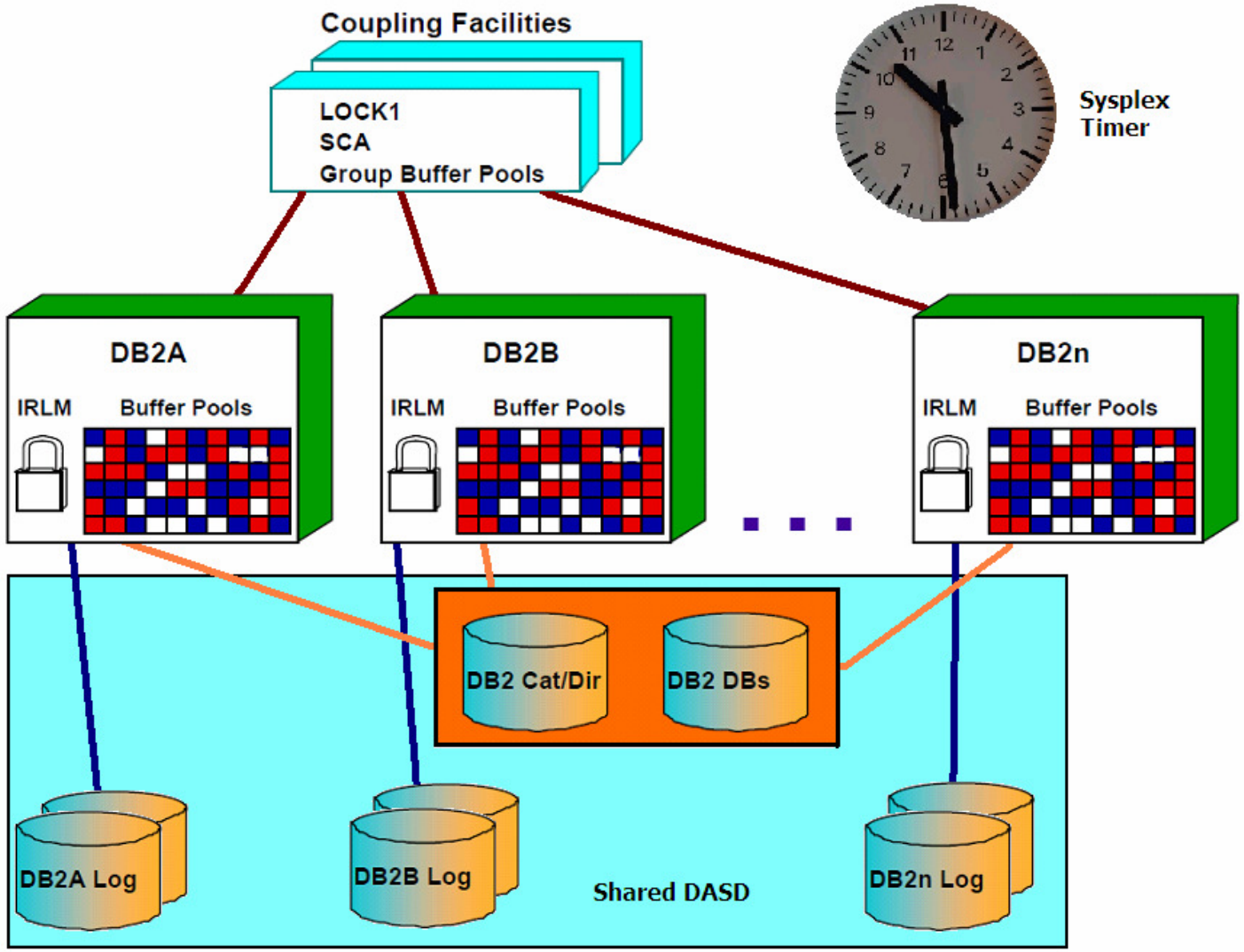
What it does not provide:

- Disaster recovery

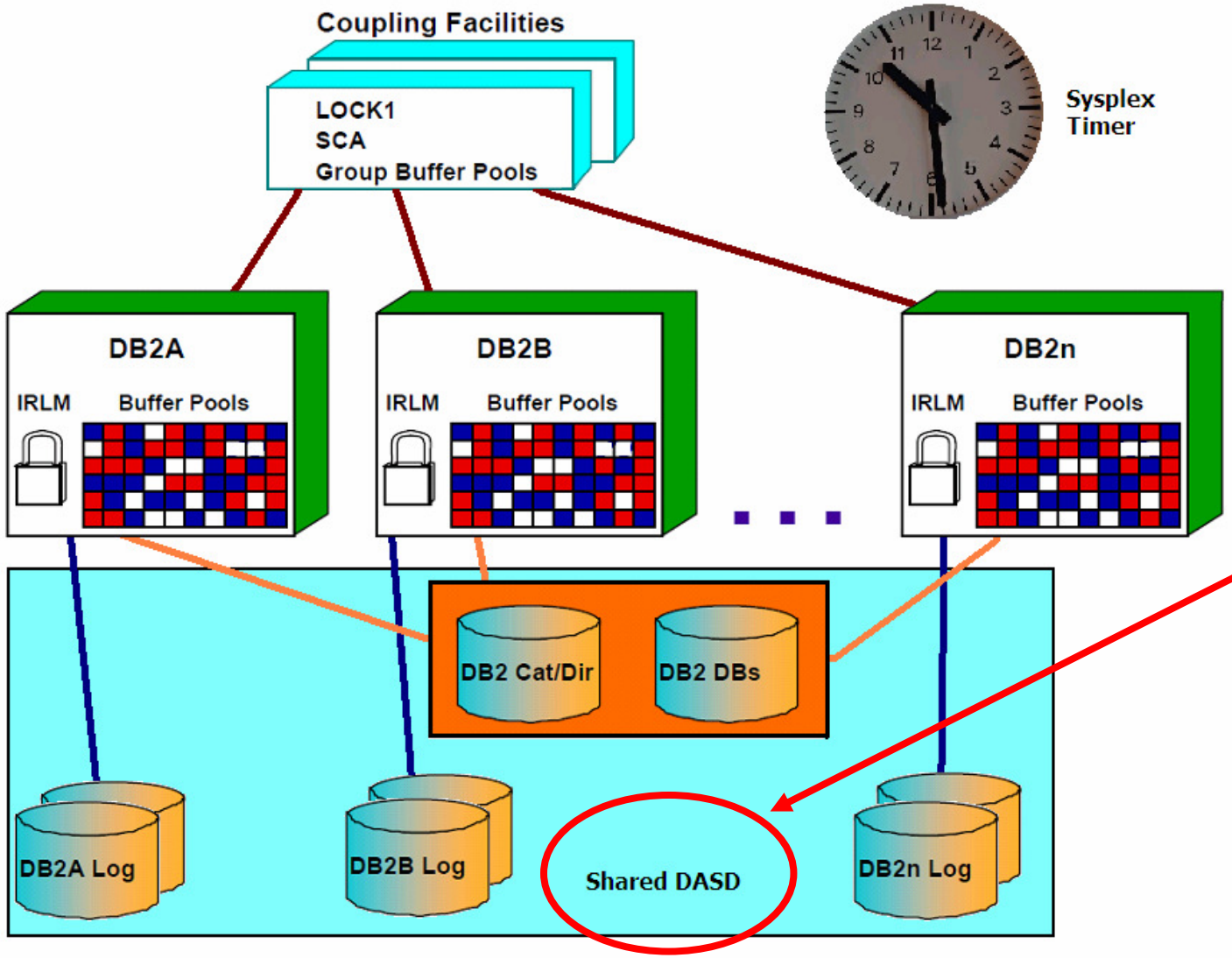
…next, we'll see why

Hint: Bad Idea!

Data sharing architecture components

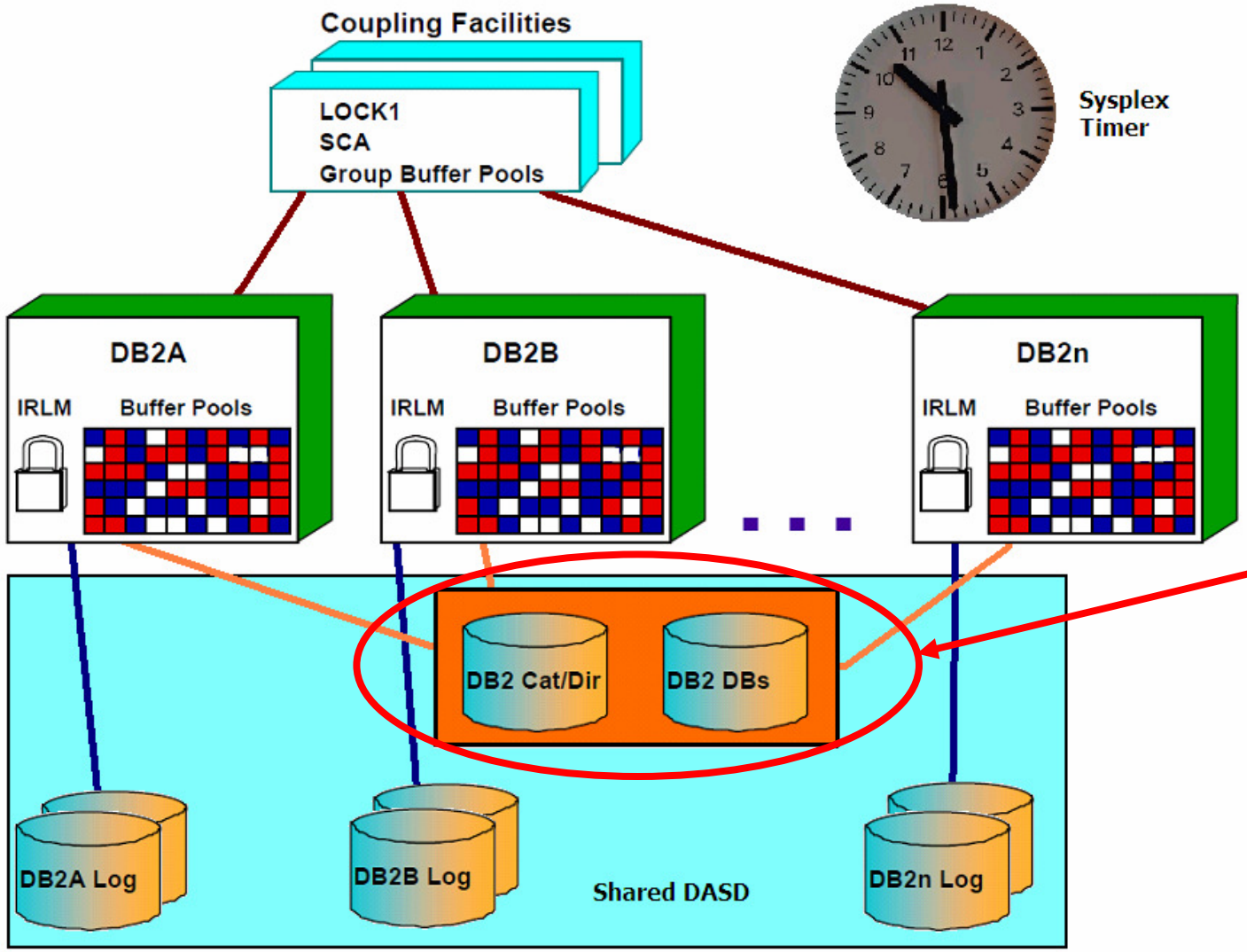


Data sharing architecture components



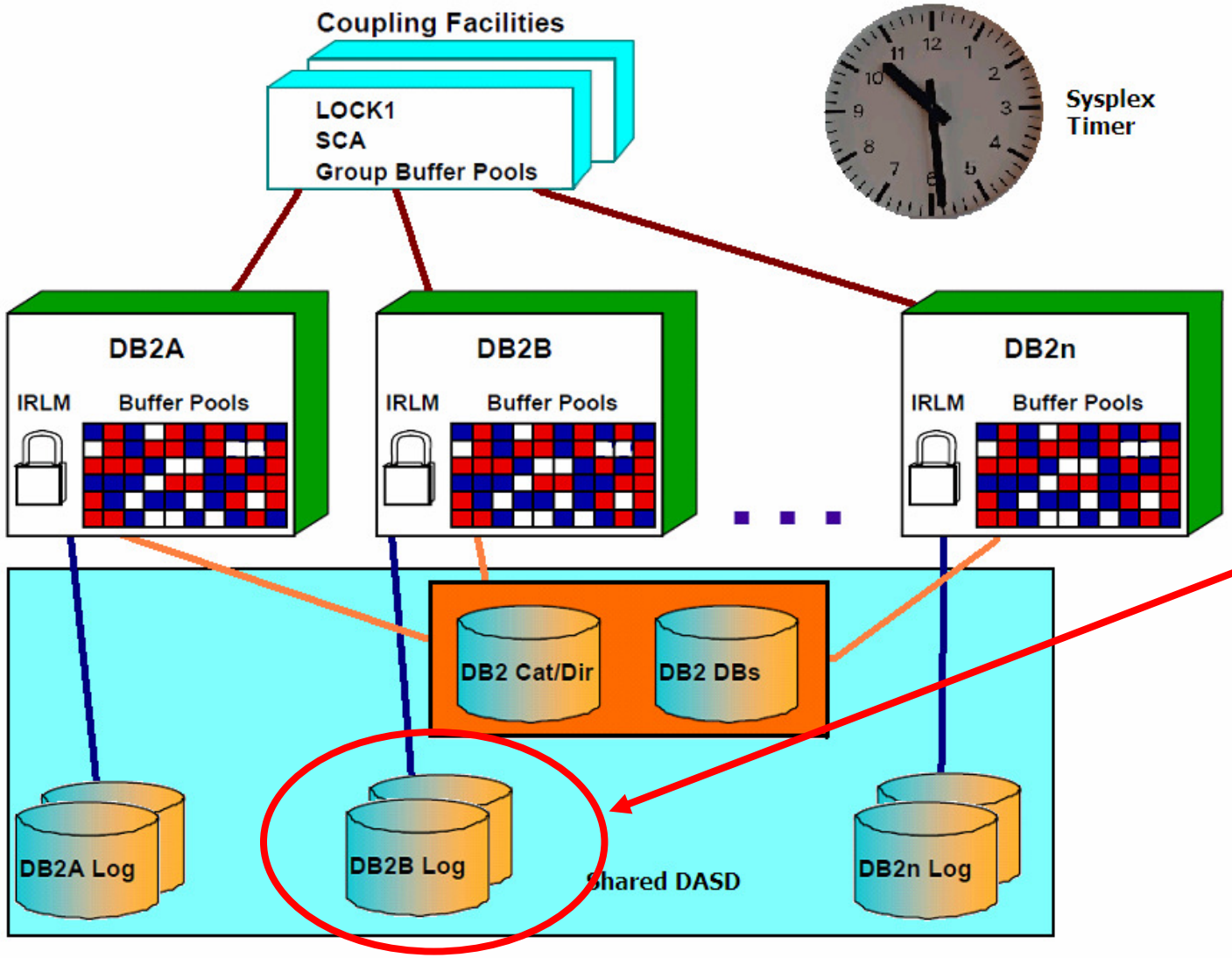
Shared DASD… by "sysplex" definition

Data sharing architecture components



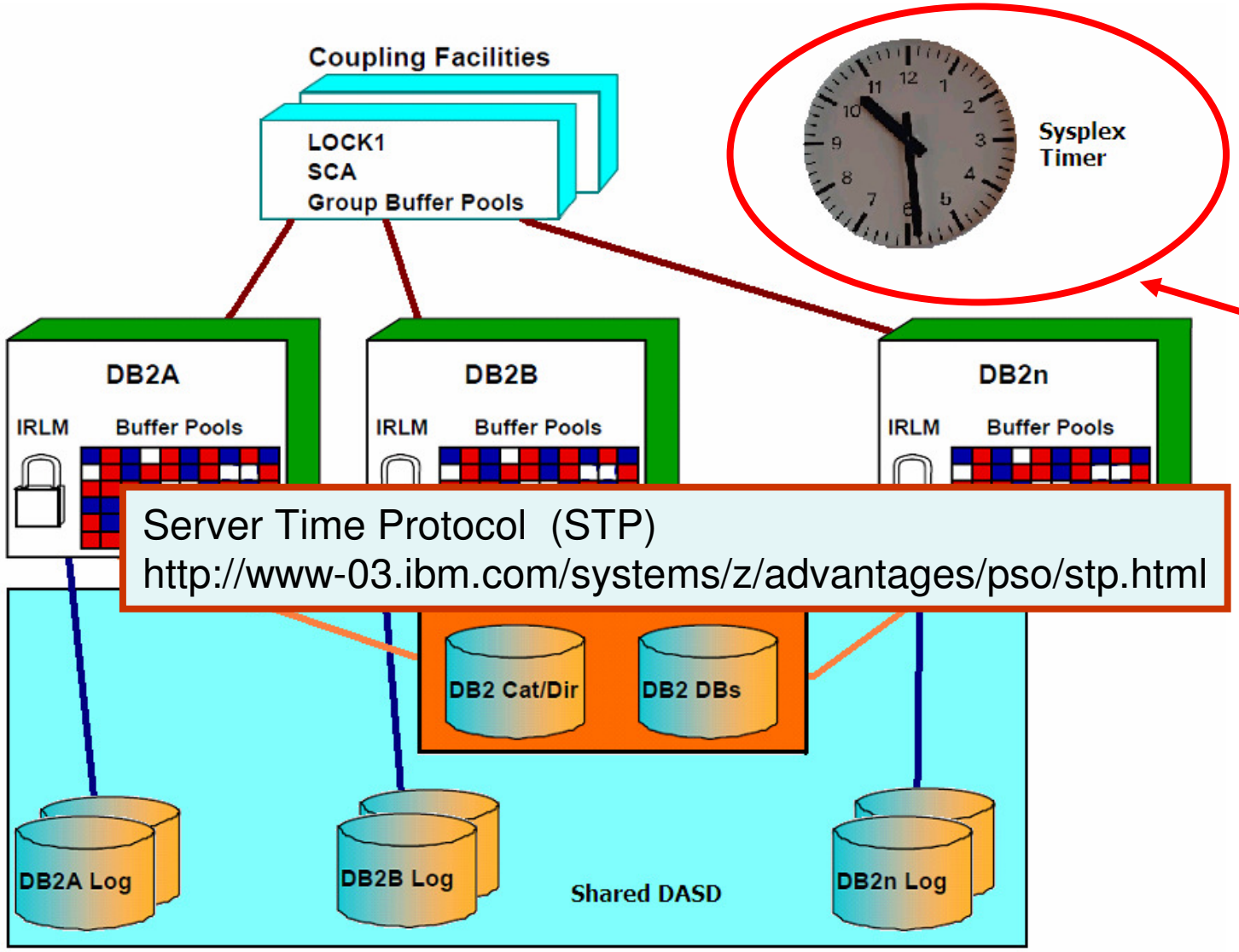
DB2 Catalog and DB2 Databases are shared between members

Data sharing architecture components



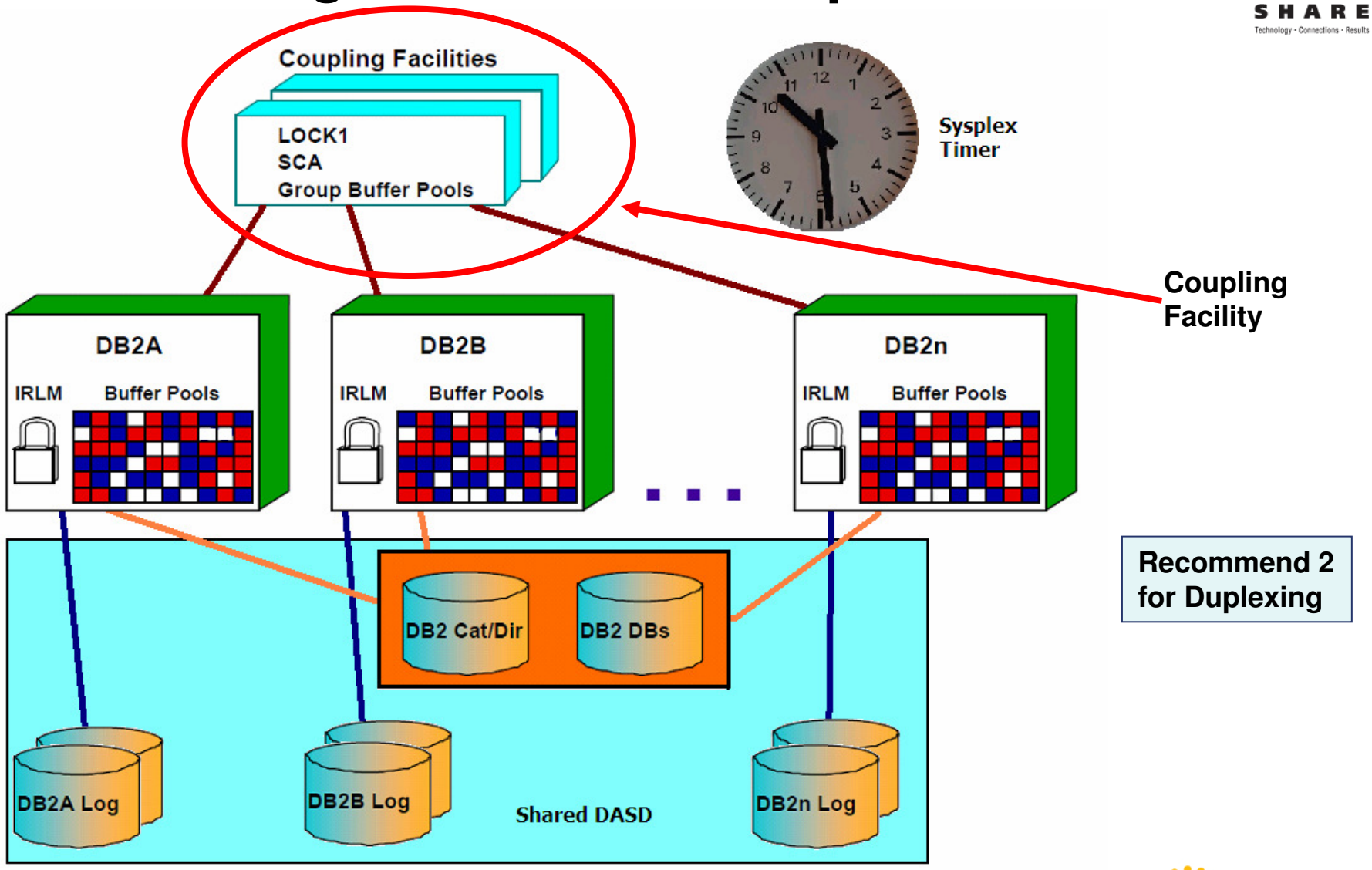
DB2 Logs are not shared between members

Data sharing architecture components

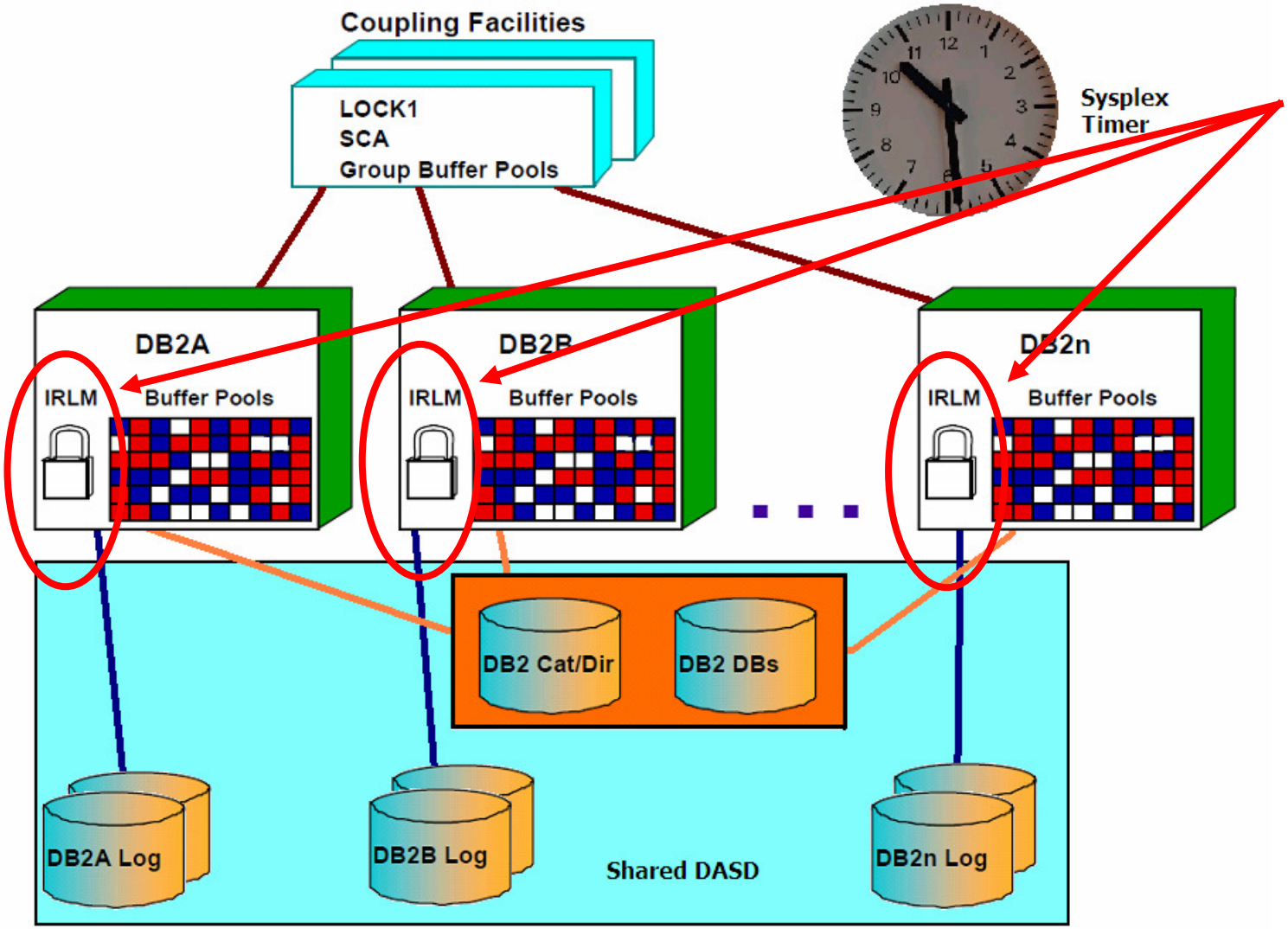


Common timing in data sharing

Data sharing architecture components

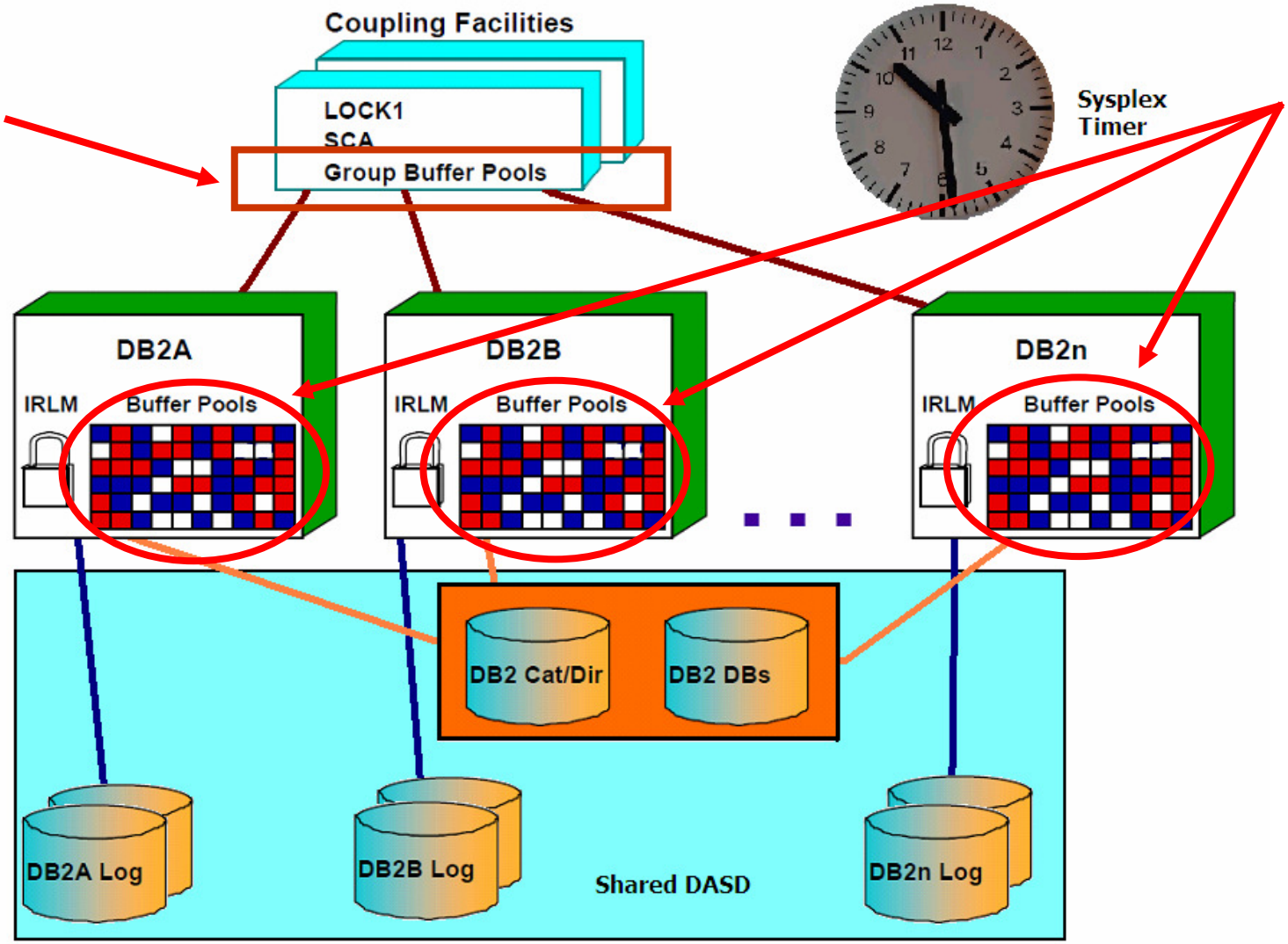


Data sharing architecture components



Locking at the subsystem (member) level

Data sharing architecture components

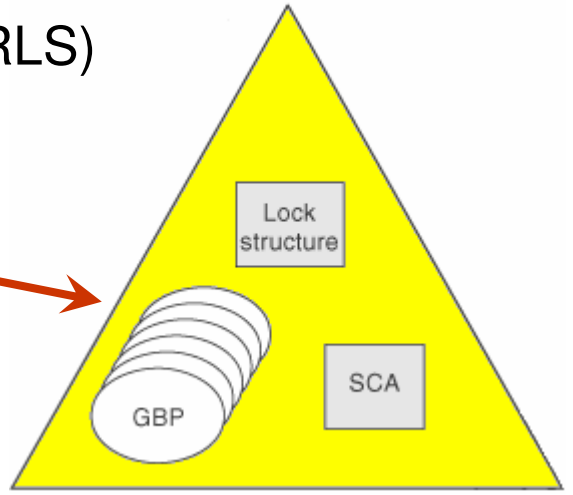


Each member has a bufferpool plus the group bufferpool

Sysplex vs Parallel Sysplex

- Sysplex gave us…
  - A standard way to communicate between systems
  - The support for cluster data sets containing member status
  - A common time source in the cluster
- Parallel sysplex gave us…
  - Ability to process data workload processing
  - DB2, IMS, and CICS record-level sharing (RLS)

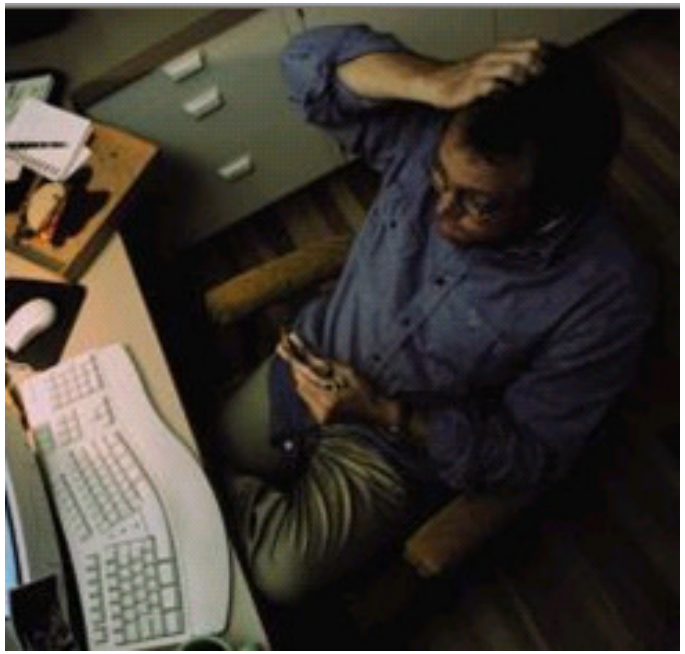
Coupling Facility structures with DB2 data sharing:
- Shared Communication Area (SCA)
- Lock
- Group buffer pools (GBPs)



One more thing…

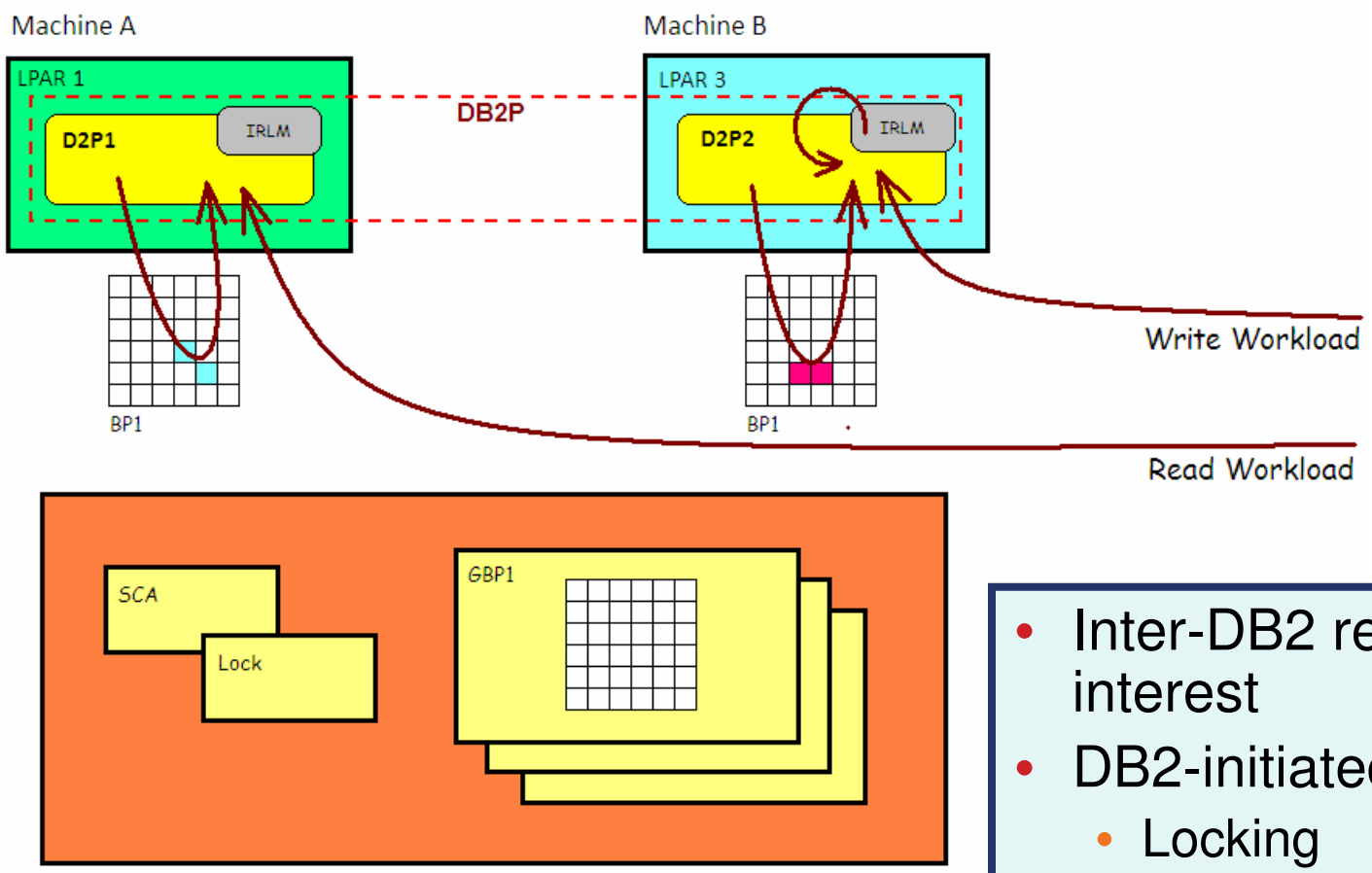
Is it a DB2 subsystem or a DB2 member in a data sharing group?

- It is a subsystem if direct attached
- It is a member in a data sharing group if group attached



Next, a workload example…

How data sharing works



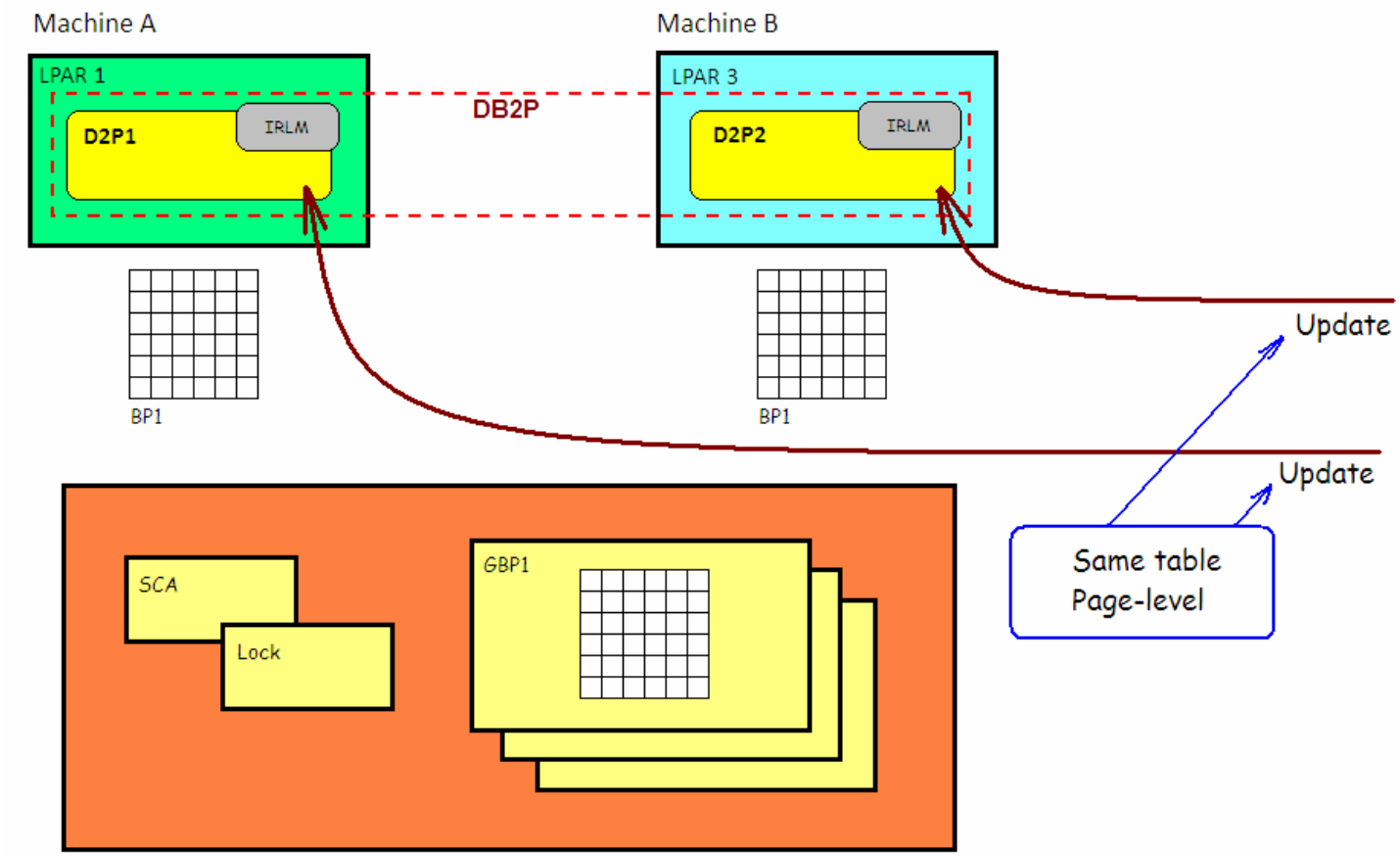
- Inter-DB2 read/write interest
- DB2-initiated
  - Locking
  - Buffer Management
- < 1% Overhead

How data sharing works

Some conditions that might initiate inter-DB2 read/write interest

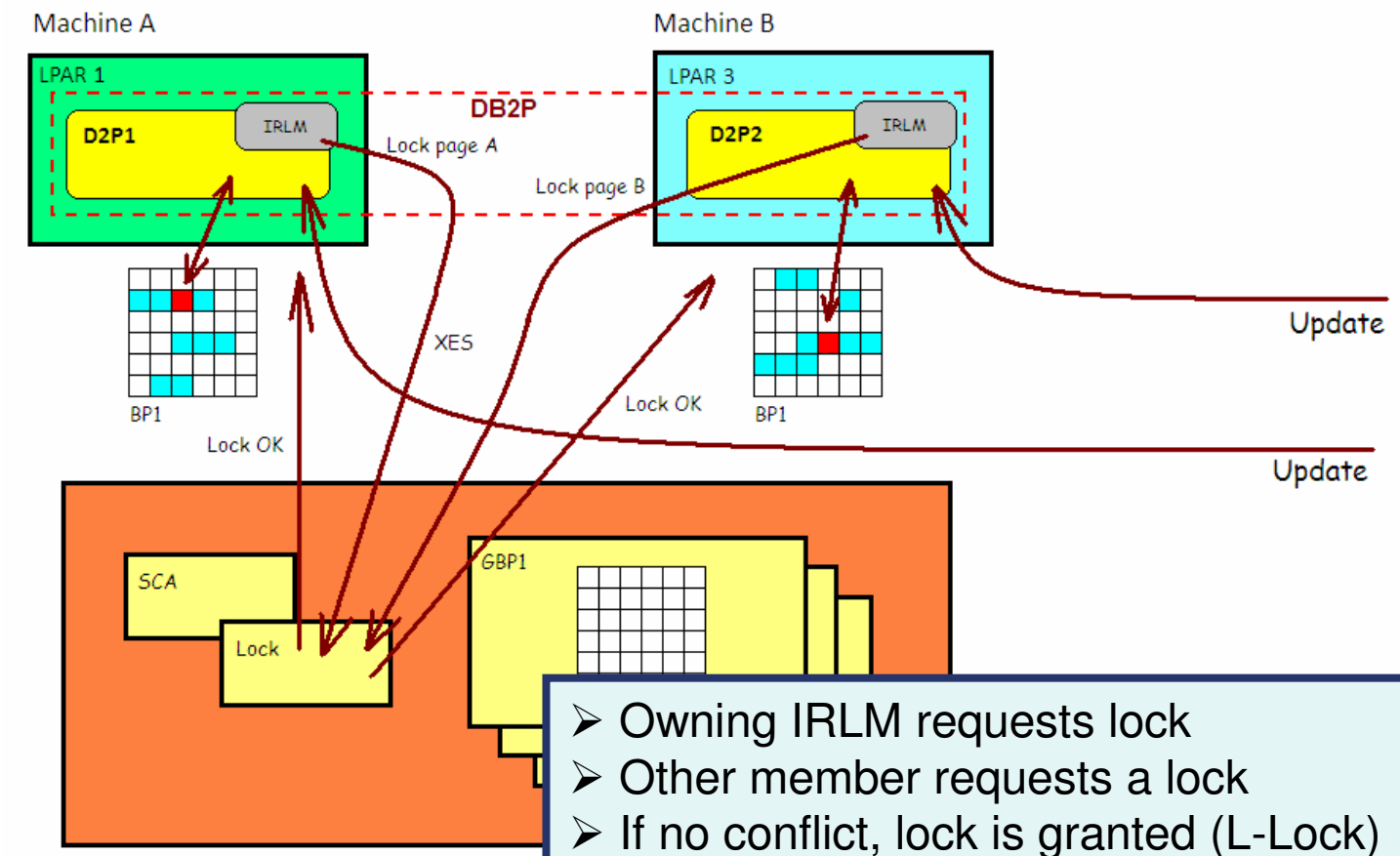
- Possible “write” from 1 or more members against the same tablespace or index
- Locking escalation in one of the members
- Classic models
  - Transactional tables (history, auditing, appl-controlled persistent data)
  - Highly used tables
  - Tables needing a REORG (most favorable page algorithm)

How data sharing works



We have 2 UPDATE workloads for the same pageset (tablespace)….

Locking



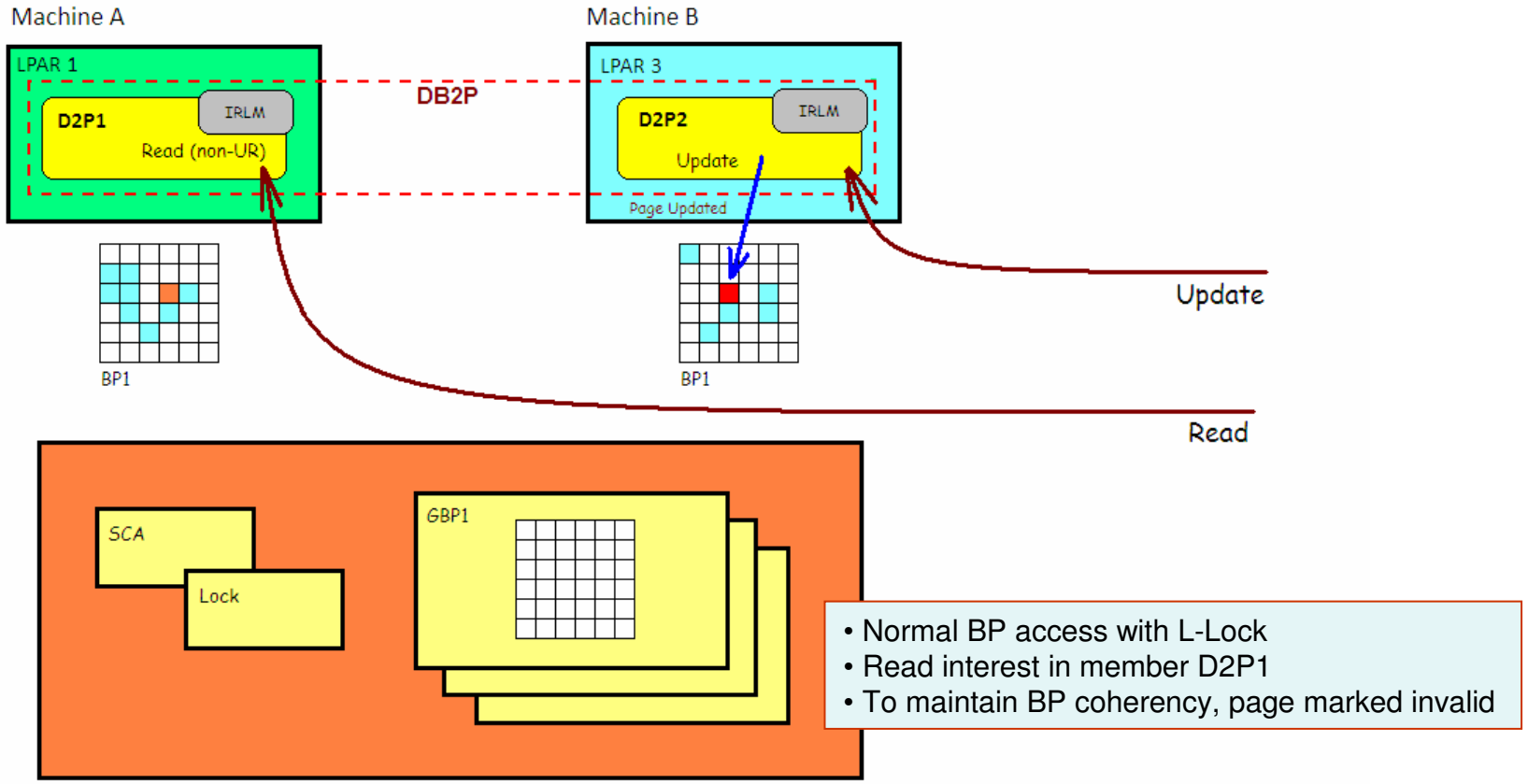
- Owning IRLM requests lock
- Other member requests a lock
- If no conflict, lock is granted (L-Lock)
- Each member manages their own updated page

Locking

- L-Locks
  - Locks held by transactions (UOW-held)
  - Traditional locks DB2 has always used
- P-Locks
  - Manage buffer coherency

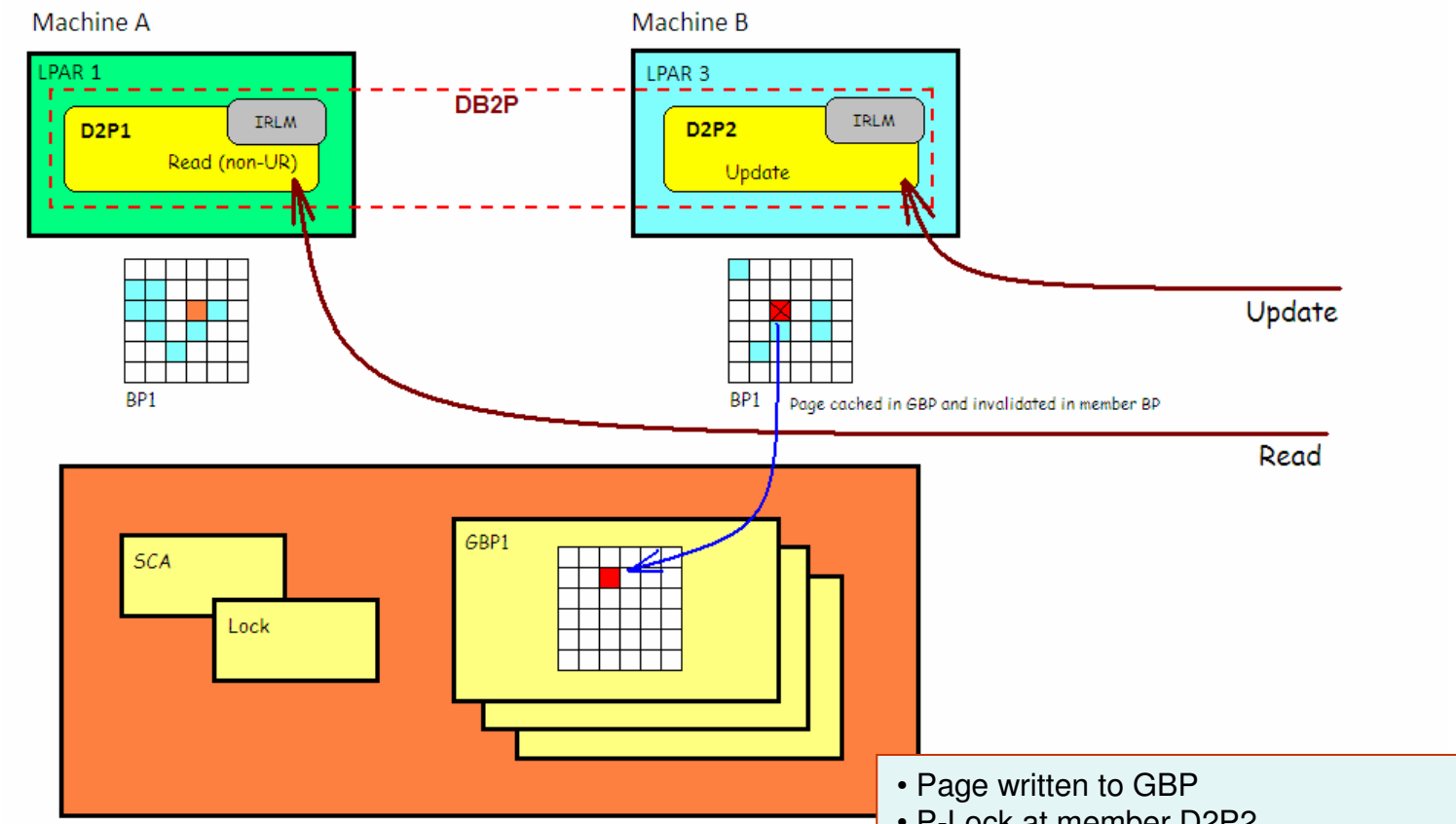
- DB2's goal – minimize all these locks
  - Keep it local
  - P-Locks over more than member is member-held, not UOW-held.
  - Only most restrictive lock is escalated to CF lock

Consider the following….



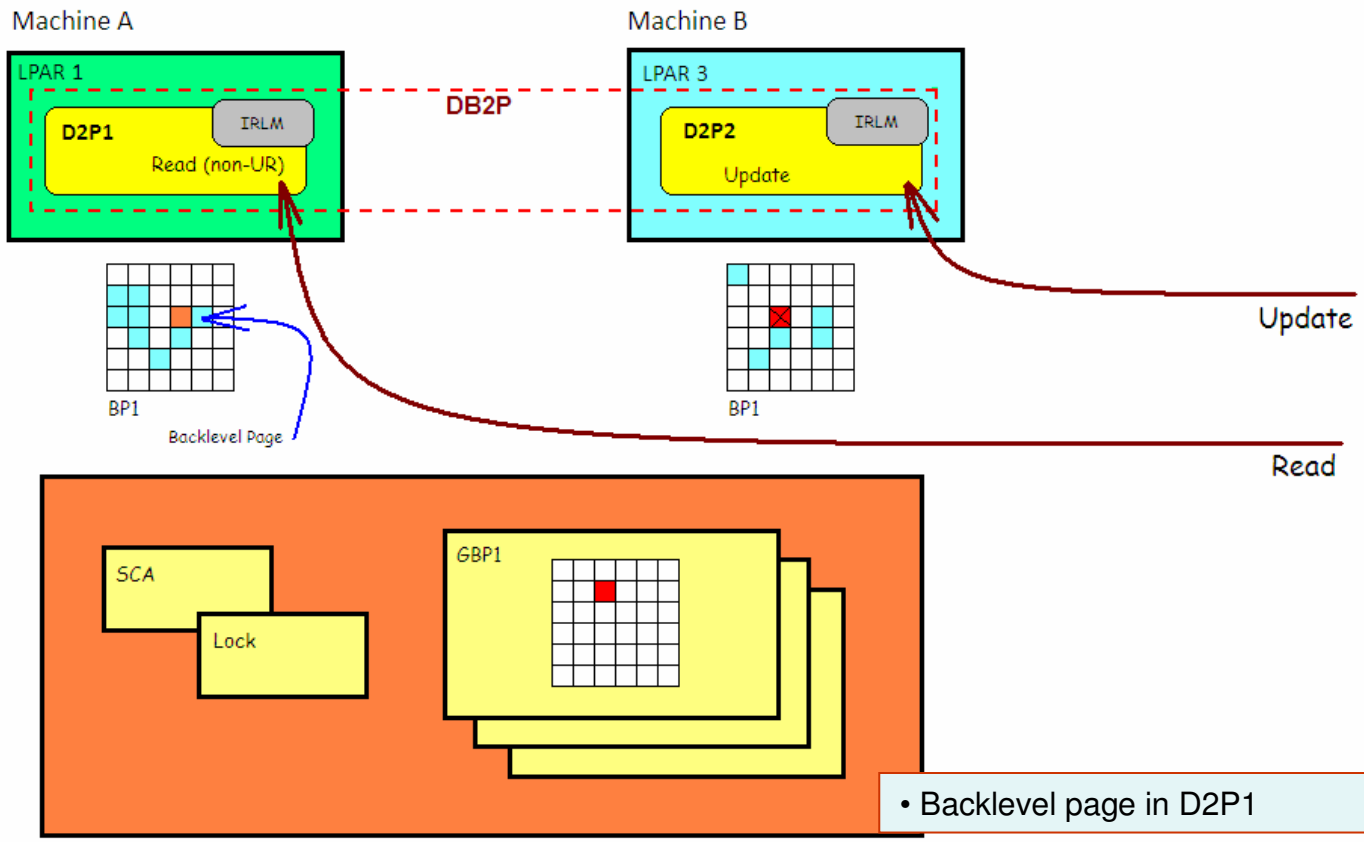
We have 1 UPDATE workload and 1 SELECT workload for the same pageset (tablespace)….

Cache page to GBP



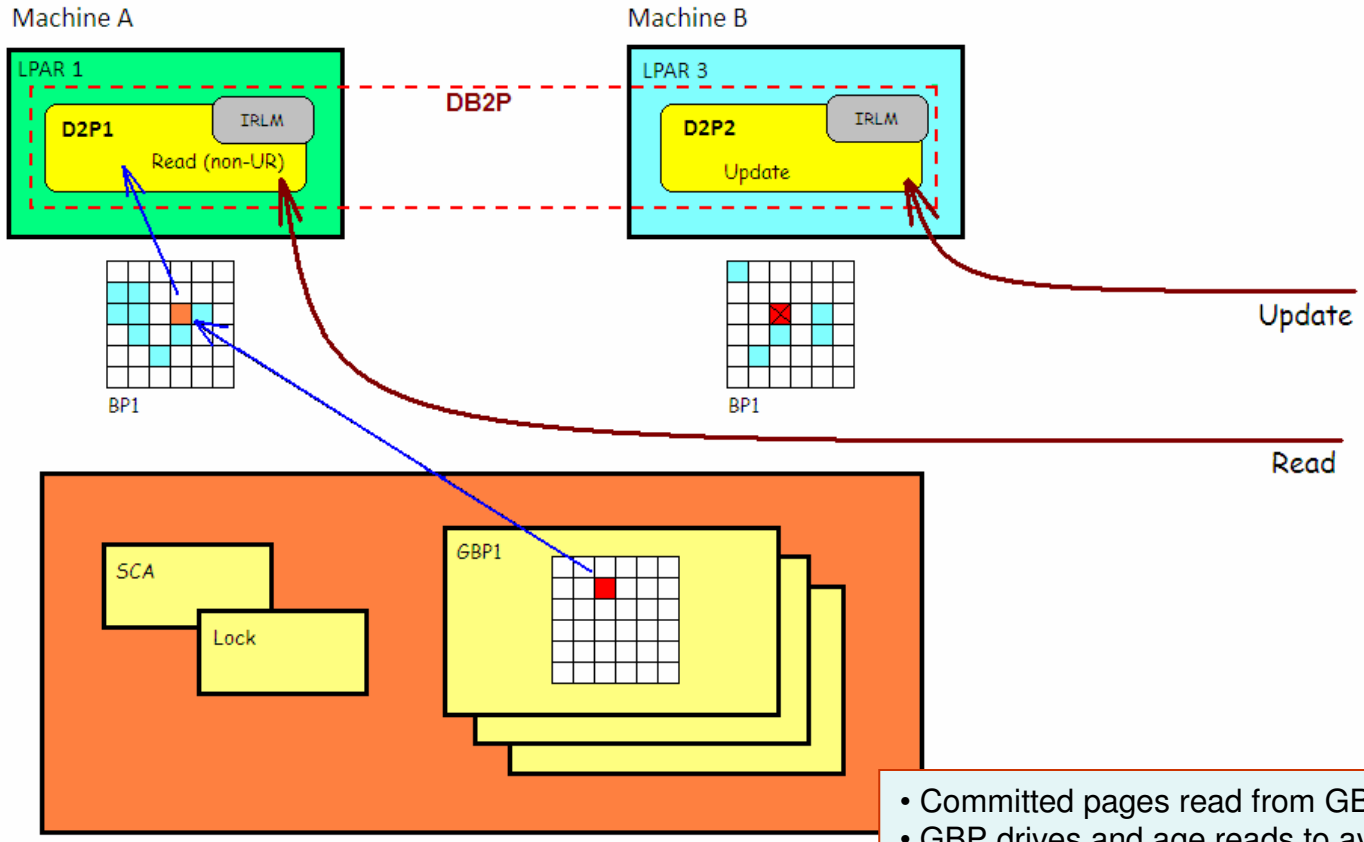
- Page written to GBP
- P-Lock at member D2P2
- Read in member D2P1 may or may not take place depending on WITH UR

Other member page read



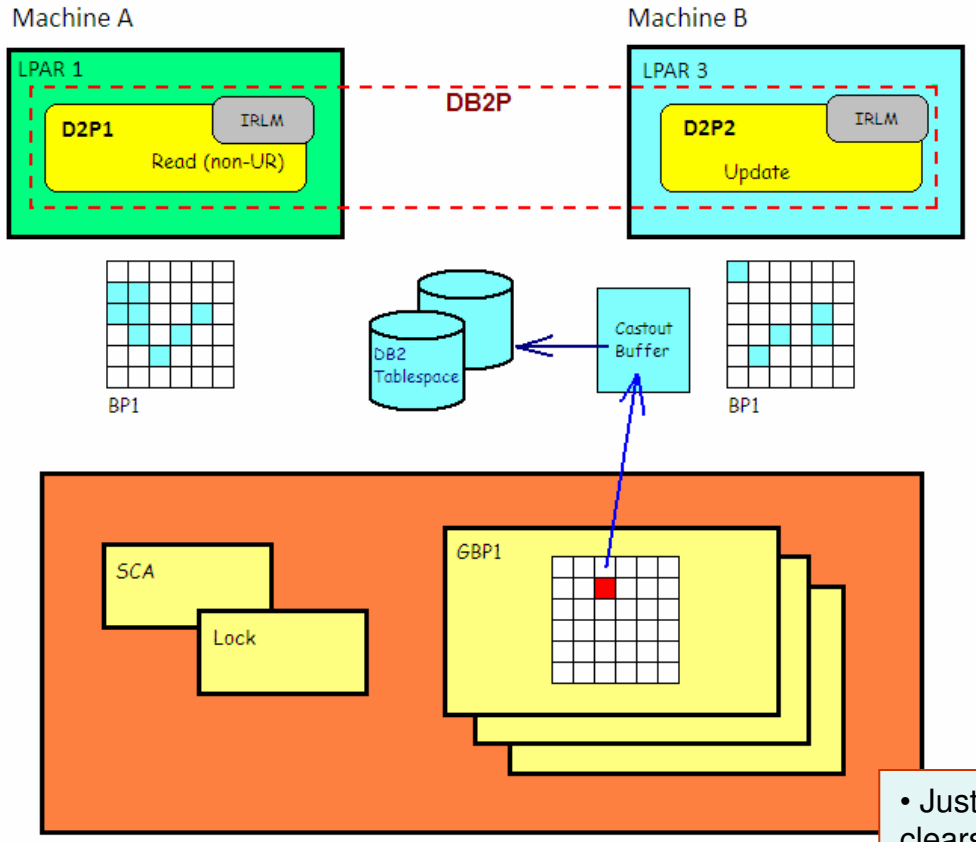
- Backlevel page in D2P1

Page used from GBP



- Committed pages read from GBP
- GBP drives and age reads to avoid I/O

Castout

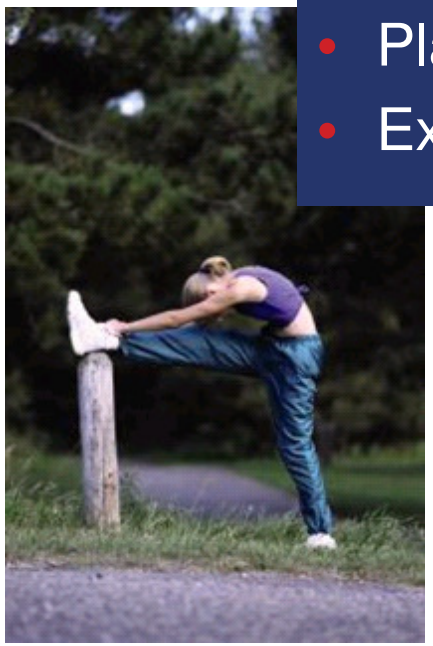


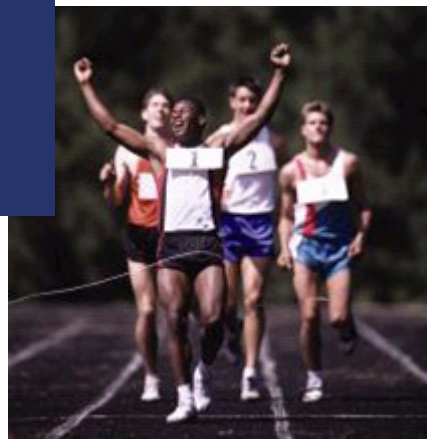
- Asynchronous
- Adjustments
  - CLASST
  - GBPOOLT

• Just like LRU algorithm in DB2, the GBP clears itself – CASTOUT
• CLASST is like VDWT
• GBPOOLT is like DWT

Data sharing project

- Educate
- Planning
- Execution



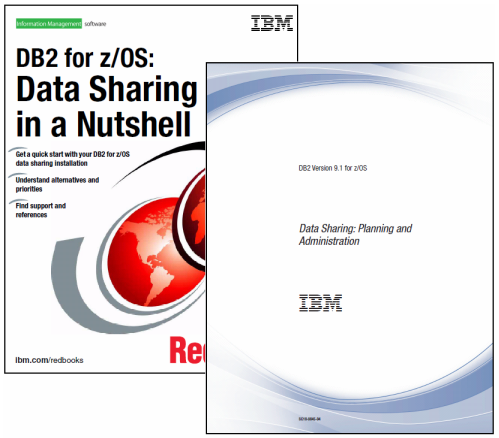


Main reference documentation

DB2 for z/OS: Data Sharing in a Nutshell

(SG24-7322) Oct 2006

http://www.redbooks.ibm.com/abstracts/sg247322.html



DB2 9 for z/OS Data Sharing: Planning and Administration

(SC18-9845) October 2009

http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/topic/com.ibm.db29.doc.dshare/dsndsk14.pdf

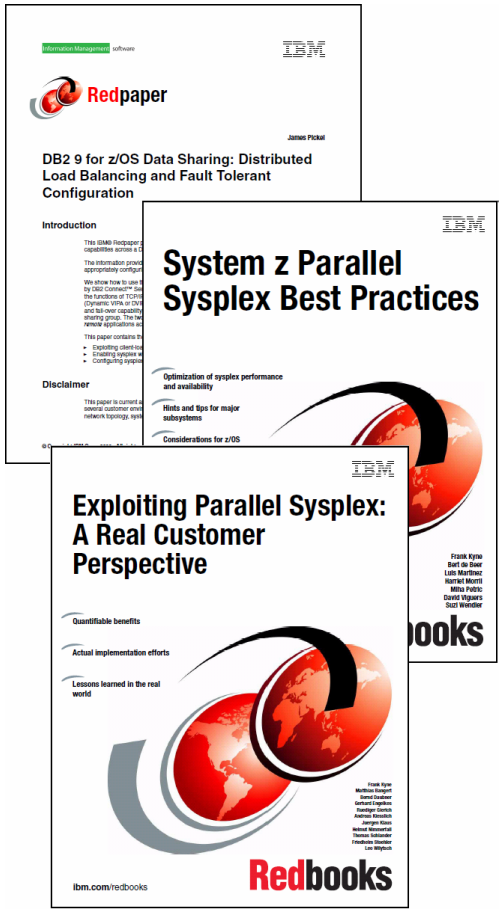
- Educate
- Planning
- Execution

Secondary reference documentation

System z Parallel Sysplex Best Practices
(SG24-7817) January 2011
http://www.redbooks.ibm.com/abstracts/sg247817.html

DB2 9 for z/OS Data Sharing: Distributed Load Balancing and Fault Tolerant Configuration
(RedPaper) 2008
http://www.redbooks.ibm.com/abstracts/redp4449.html

Exploiting Parallel Sysplex:
A Real Customer Perspective
(SG24-7108) October 2006
http://www.redbooks.ibm.com/abstracts/sg247108.html



Classes

Learning / installing:

DB2 9 for z/OS Data Sharing Implementation (CV410)
3-day lecture-only course

DB2 9 for z/OS Data Sharing Implementation Workshop (CV450)
4 1/2 day lab course

Disaster Recovery:

DB2 9 for z/OS Data Sharing Recovery and Restart (CV420)
2-day lecture-only course

DB2 9 for z/OS Data Sharing Recovery/Restart Workshop (CV920)
4 1/2 day lab course

Other Related Courses

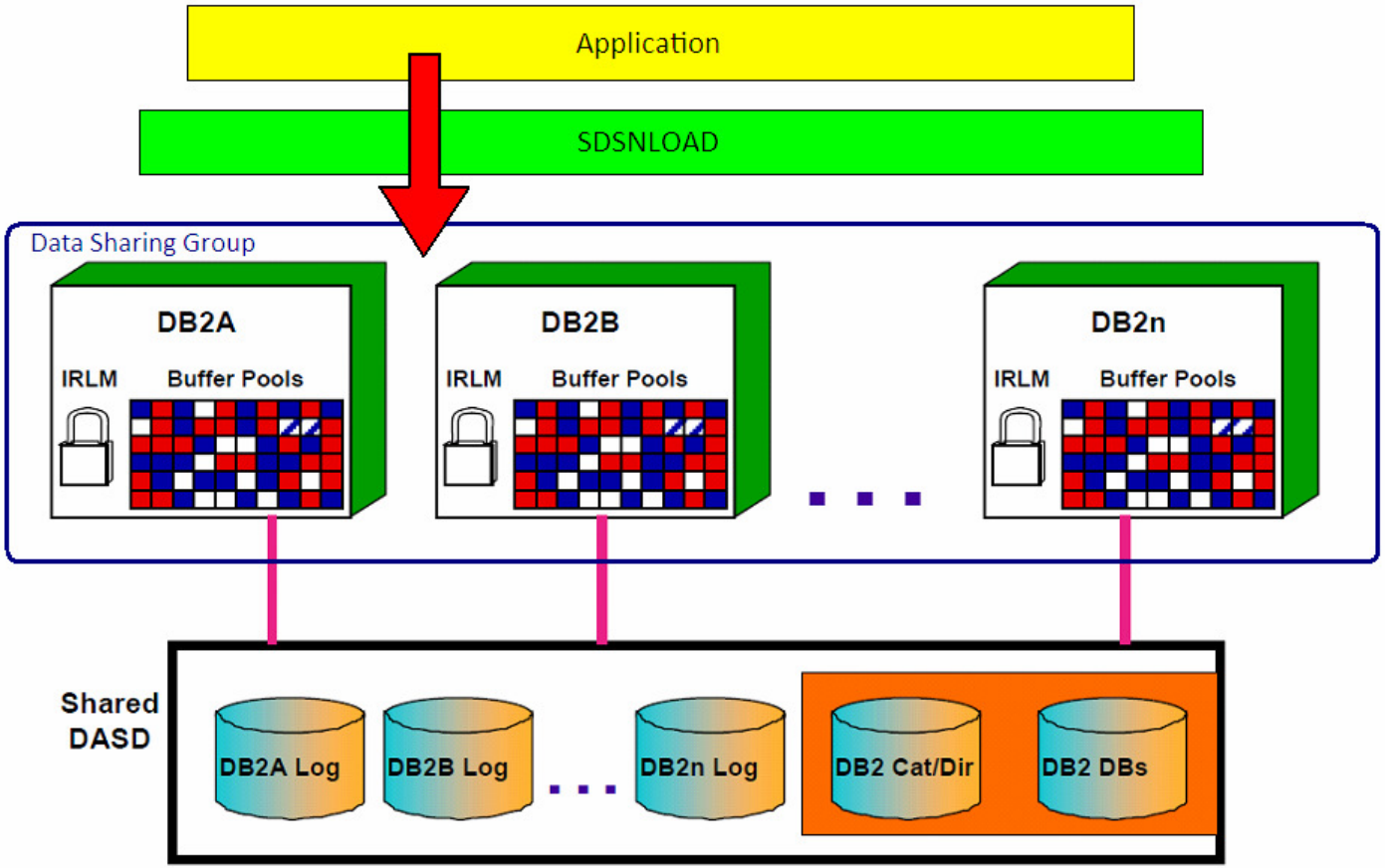
CICS V4.1 CICSPlex System Manager Introduction (WM844)
CICS V4.1 CICSPlex System Manager Administration (WM854)
Parallel Sysplex Operations for the Data Sharing Environment (IMS, DB2, CICS) (SZ970)

Planning

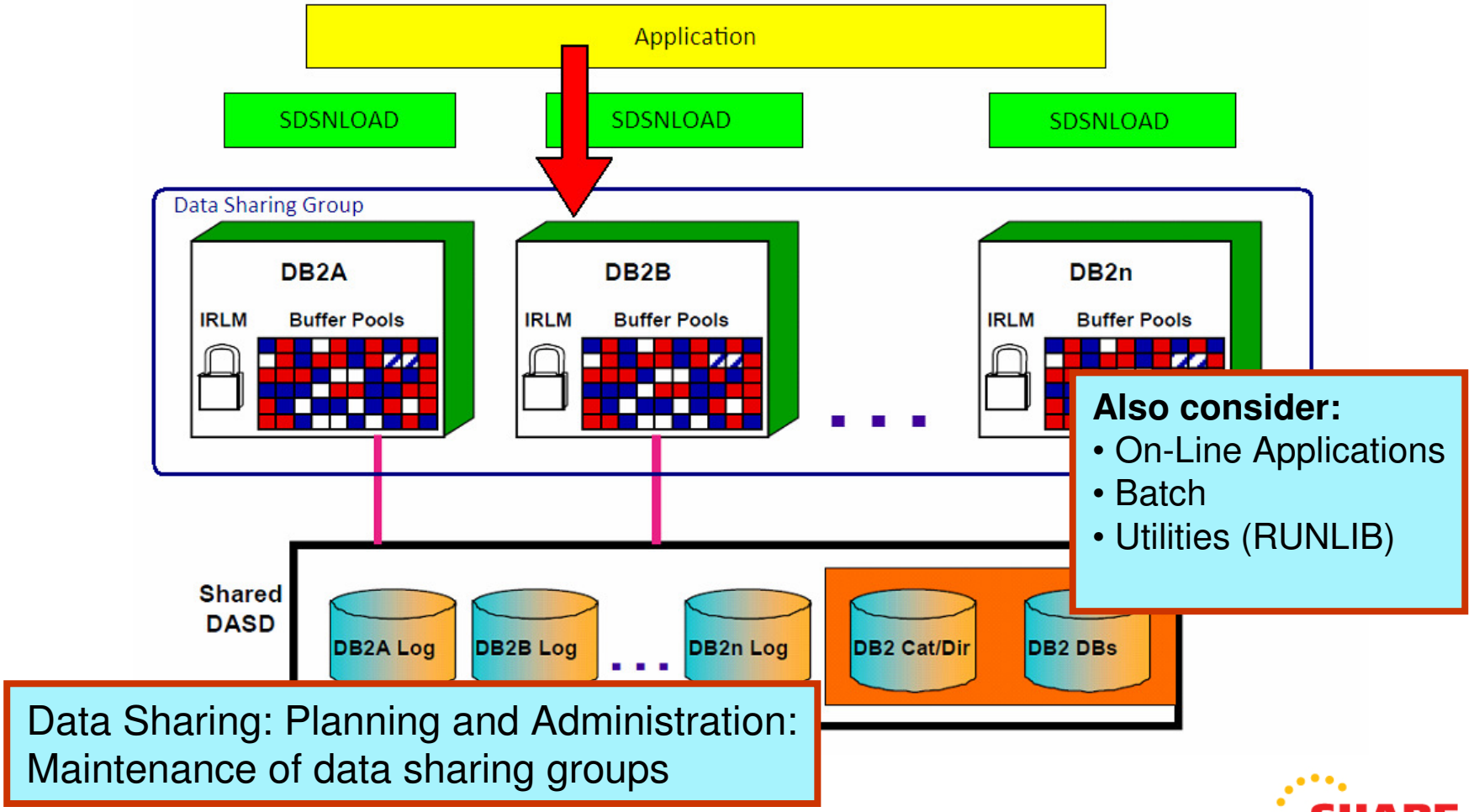
1. Think about the philosophy
   - Maintenance/Upgrade
   - Configuration
2. Name standards
   - Dataset
   - DB2
   - Environmental

- Educate
- Planning
- Execution

Maintenance and the art of application execution



Maintenance and the art of application execution

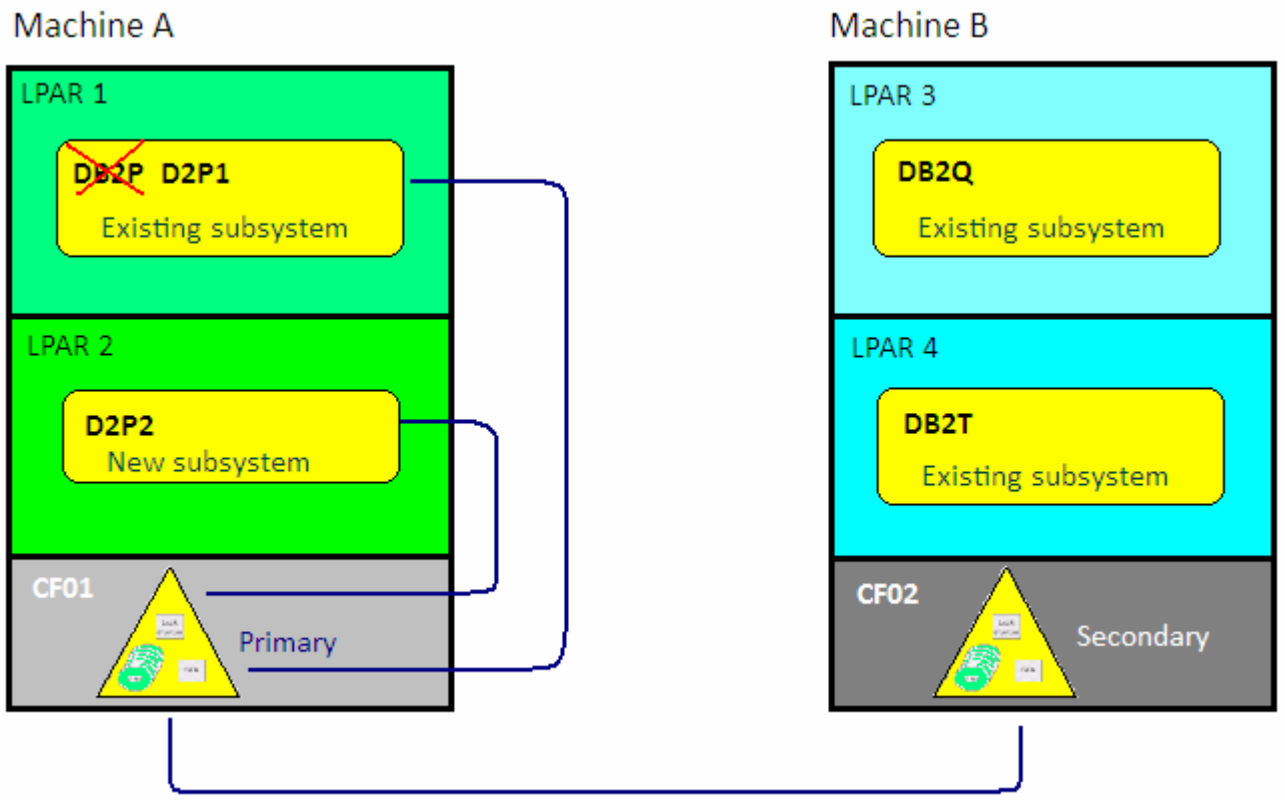


Data Sharing: Planning and Administration: Maintenance of data sharing groups

Configuration quandary

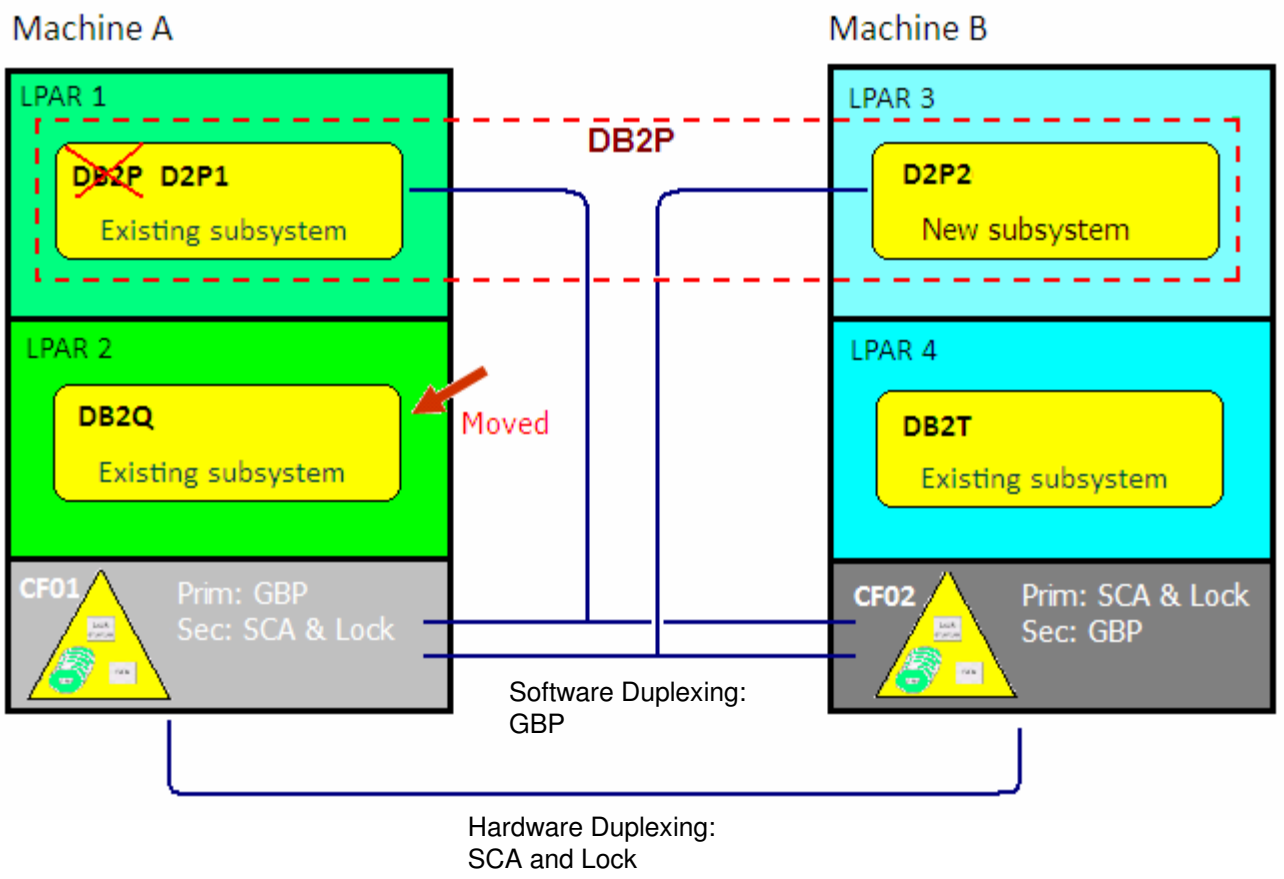
- Data sharing members on the same machine
- CF not configured

What is the point of using data sharing?



Configuration quandary

1. Move the test subsystem from machine B to machine A
2. Split group buffer pool structures from SCA and lock structures
3. Any single point-of-failure will not disrupt service



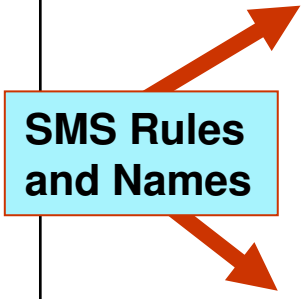
Continuous availability!

Naming

Example of creating a new data sharing group by adding a second DB2 subsystem/member

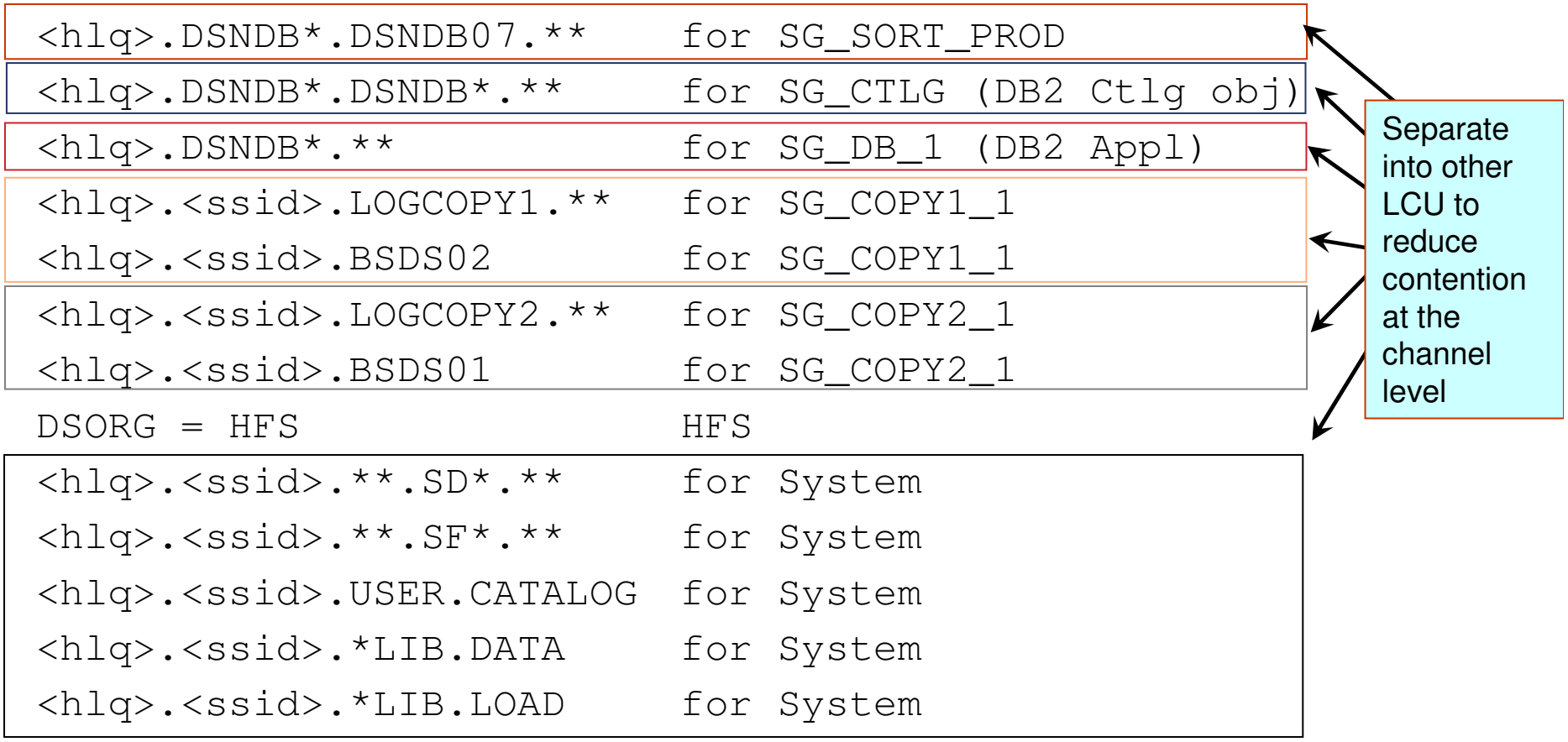
		Production		
		Existing	**New (Existing)**	**New (Adding)**
Unique to each member	HLQ:	DB2P	n/a	n/a
	VCAT:	DB2P	D2P1	D2P1
	SSID/Member:	DB2P	D2P1	D2P2
	ResPort:	453	353	354
	IRLMID:	PRLM	IRP1	IRP2
	LU:	LUDB2P	LUD2P1	LUD2P2
	Location:	PRODLODB2P	PRODLOD2P1	PRODLOD2P2
	Work:	DSNDB07	D2P1DB07	D2P2DB07
	IRLM PROC:	IRLMPROC	IRP1PROC	IRP2PROC
	zParm (DSNTIJUZ):	DSNZDB2P	DSNZD2P1	DSNZD2P2
Common items	Location:		PRODLOD2P1	
	Group/HLQGroup		DSNDB2P	
	Port:		352	
	Group Attach:		DB2P	
	IRLMXCF Group:		DXRDB2P	

Naming – System Managed Storage



		Production		
		Existing	New (Existing)	New (Adding)
Unique to each member	HLQ:	DB2P	n/a	n/a
	VCAT:	DB2P	D2P1	D2P1
	SSID/Member:	DB2P	D2P1	D2P2
	ResPort:	453	353	354
	IRLMID:	PRLM	IRP1	IRP2
	LU:	LUDB2P	LUD2P1	LUD2P2
	Location:	PRODLODB2P	PRODLOD2P1	PRODLOD2P2
	Work:	DSNDB07	D2P1DB07	D2P2DB07
	IRLM PROC:	IRLMPROC	IRP1PROC	IRP2PROC
	zParm (DSNTIJUZ):	DSNZDB2P	DSNZD2P1	DSNZD2P2
Common items	Location:		PRODLOD2P1	
	Group/HLQGroup		DSNDB2P	
	Port:		352	
	Group Attach:		DB2P	
	IRLMXCF Group:		DXRDB2P	

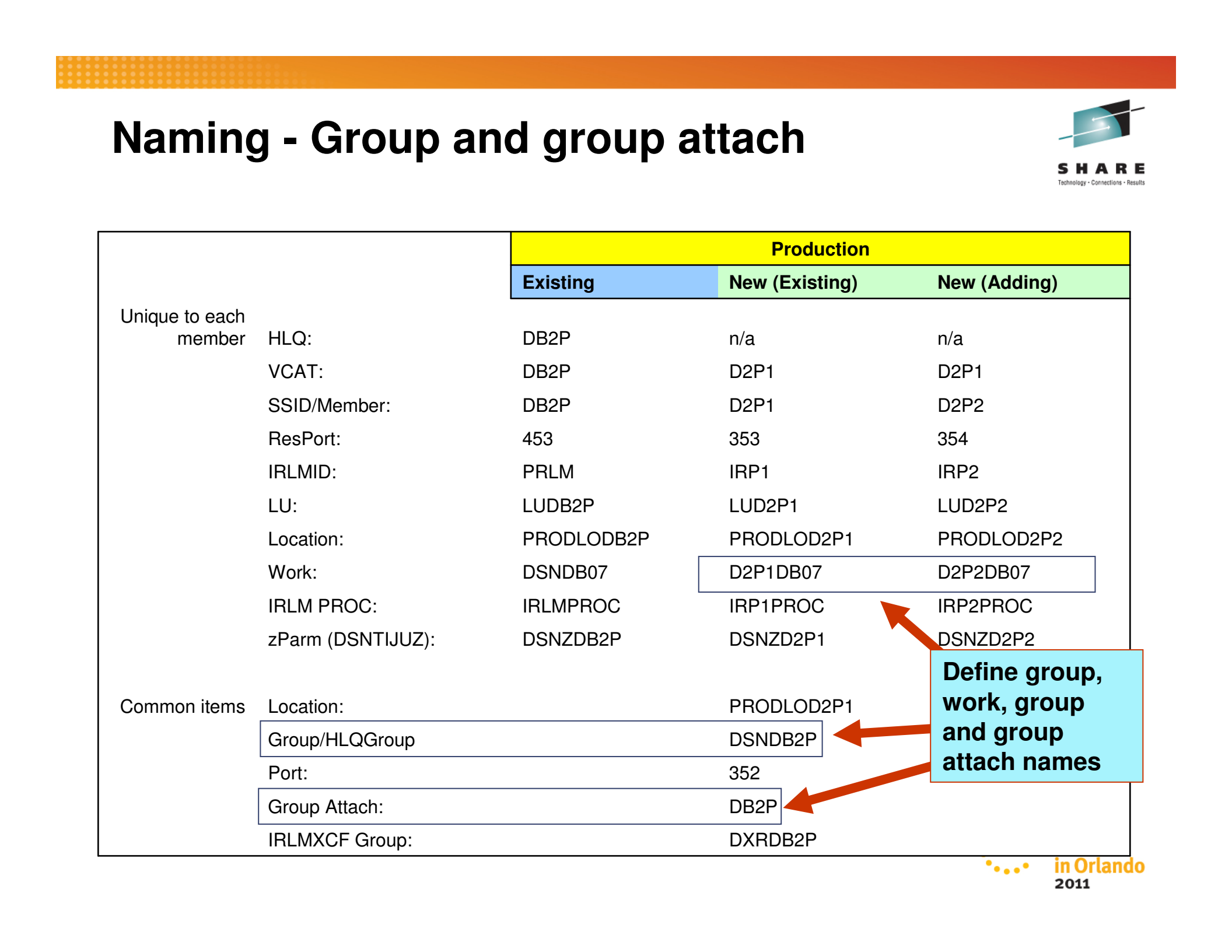
System Managed Storage (SMS)



Redbook: DB2 9 for z/OS and Storage Management
http://www.redbooks.ibm.com/redpieces/abstracts/sg247823.html?Open

Naming - Group and group attach

		Production		
		Existing	**New (Existing)**	**New (Adding)**
Unique to each member	HLQ:	DB2P	n/a	n/a
	VCAT:	DB2P	D2P1	D2P1
	SSID/Member:	DB2P	D2P1	D2P2
	ResPort:	453	353	354
	IRLMID:	PRLM	IRP1	IRP2
	LU:	LUDB2P	LUD2P1	LUD2P2
	Location:	PRODLODB2P	PRODLOD2P1	PRODLOD2P2
	Work:	DSNDB07	D2P1DB07	D2P2DB07
	IRLM PROC:	IRLMPROC	IRP1PROC	IRP2PROC
	zParm (DSNTIJUZ):	DSNZDB2P	DSNZD2P1	DSNZD2P2
Common items	Location:		PRODLOD2P1	
	Group/HLQGroup		DSNDB2P	
	Port:		352	
	Group Attach:		DB2P	
	IRLMXCF Group:		DXRDB2P	



CF structure name - DB2 name relationship

```
                INSTALL DB2 - DEFINE GROUP OR MEMBER
===>

Check parameters and reenter to change:

 1  GROUP NAME     ===> DSNDB2P      Name of the DB2 group
 2  MEMBER NAME    ===> D2P1         Name of DB2 member in group
 3  WORK FILE DB   ===> D2P1DB07     Work file database name for this member
 4  GROUP ATTACH   ===> DB2P         Group attach name for TSO, batch, utilities
```

STRUCTURE NAME(DSNDB2P_GBP0) SIZE(204800)
INITSIZE(102400)
FULLTHRESHOLD(90)
ALLOWAUTOALT(YES)
DUPLEX(ENABLED)
REBUILDPERCENT(1)
PREFLIST(CF02, CF01)
ENFORCEORDER(YES)

STRUCTURE NAME(DSNDB2P_LOCK1) SIZE(262144)
INITSIZE(131072)
REBUILDPERCENT(1)
PREFLIST(CF01,CF02)
ENFORCEORDER(YES)

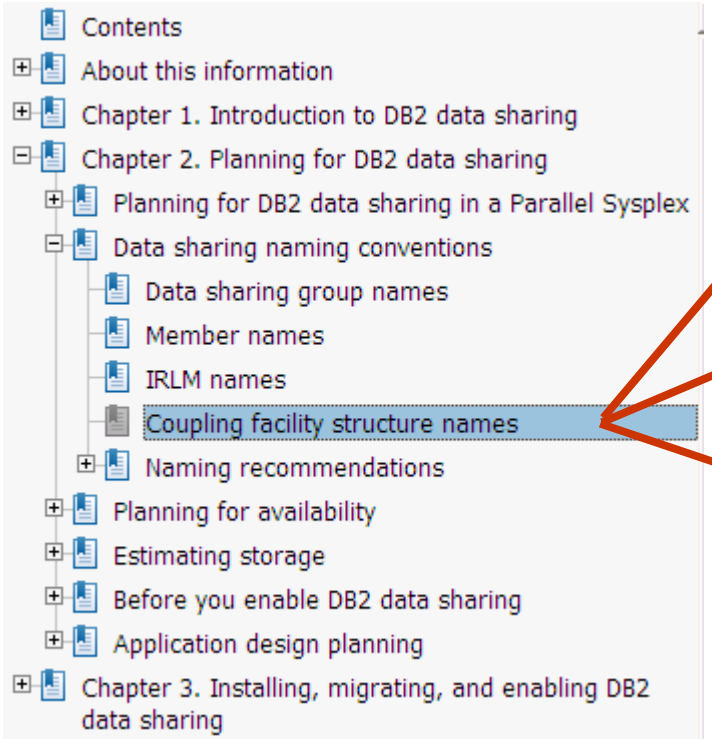
STRUCTURE NAME(DSNDB2P_SCA) SIZE(65536)
INITSIZE(32768)
REBUILDPERCENT(1)
PREFLIST(CF01,CF02)
ENFORCEORDER(YES)

- Shared Communication Area (SCA)
- Lock
- Group buffer pools (GBPs)

Work with your z/OS systems programmer when setting up structures

CF structure name - DB2 name relationship

In “Data Sharing: Planning and Administration”



Lock structure name:

*groupname*_LOCK1

Shared communications area:

*groupname*_SCA

Group buffer pool names:

*groupname*_GBP*xxxx*

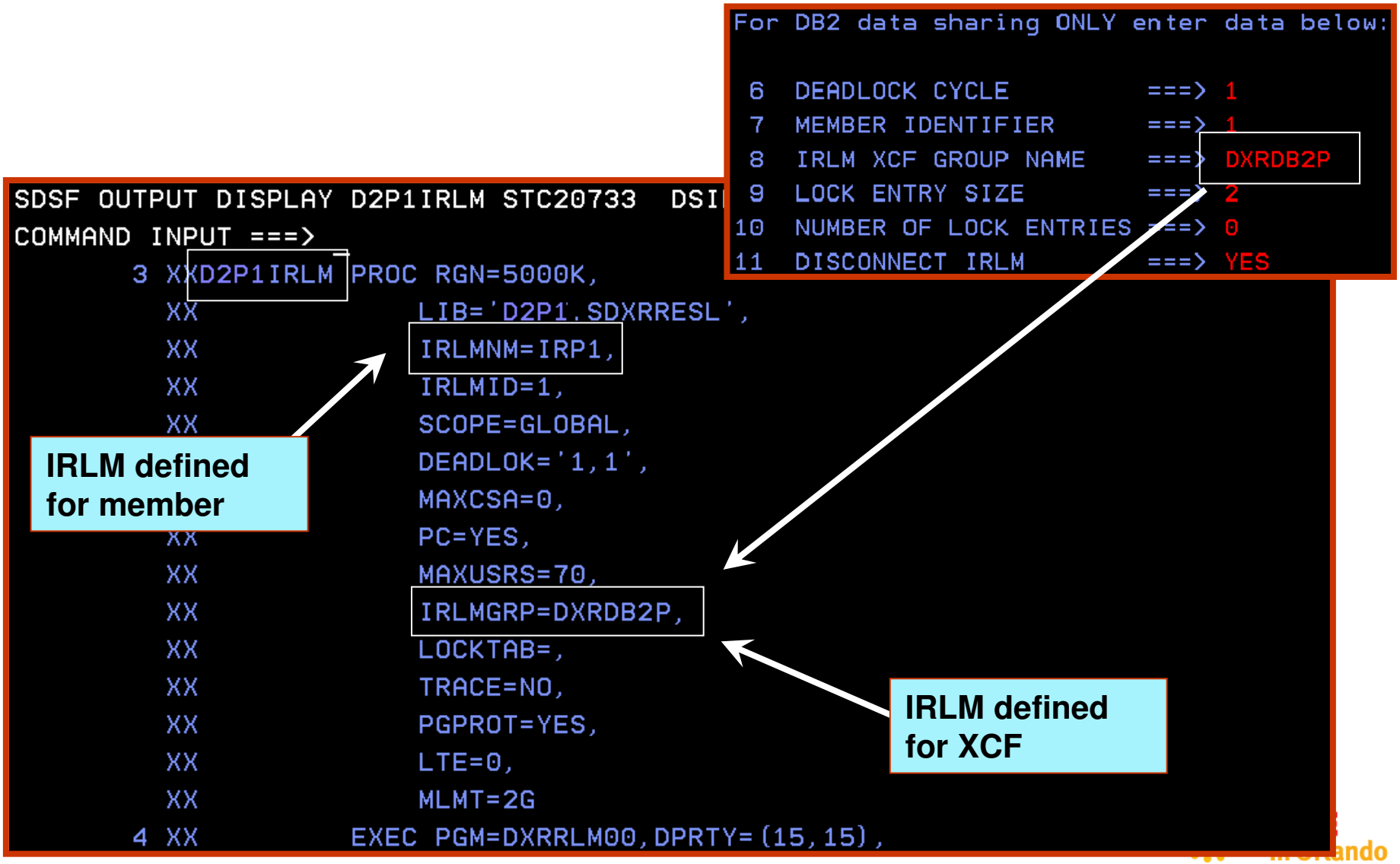
Naming - Locking

		Production		
		Existing	New (Existing)	New (Adding)
Unique to each member	HLQ:	DB2P	n/a	n/a
	VCAT:	DB2P	D2P1	D2P1
	SSID/Member:	DB2P	D2P1	D2P2
	ResPort:	453	353	354
	IRLMID:	PRLM	IRP1	IRP2
	LU:	LUDB2P	LUD2P1	
	Location:	PRODLODB2P	PRODLOD2P1	
	Work:	DSNDB07	D2P1DB07	
	IRLM PROC:	IRLMPROC	IRP1PROC	IRP2PROC
	zParm (DSNTIJUZ):	DSNZDB2P	DSNZD2P1	DSNZD2P2
Common items	Location:		PRODLOD2P1	
	Group/HLQGroup		DSNDB2P	
	Port:		352	
	Group Attach:		DB2P	
	IRLMXCF Group:		DXRDB2P	

IRLM defined for member

IRLM defined for XCF

DB2 definition for IRLM



Naming - Connectivity

		Production		
		Existing	New (Existing)	New (Adding)
Unique to each member	HLQ:	DB2P	n/a	n/a
	VCAT:	DB2P	D2P1	D2P1
	SSID/Member:	DB2P	D2P1	D2P2
	ResPort:	453	353	354
	IRLMID:	PRLM	IRP1	IRP2
	LU:	LUDB2P	LUD2P1	
	Location:	PRODLODB2P	PRODLOD2P1	
	Work:	DSNDB07	D2P1DB07	D2P2DB07
	IRLM PROC:	IRLMPROC	IRP1PROC	IRP2PROC
	zParm (DSNTIJUZ):	DSNZDB2P	DSNZD2P1	DSNZD2P2
Common items	Location:		PRODLOD2P1	
	Group/HLQGroup		DSNDB2P	
	Port:		352	
	Group Attach:		DB2P	
	IRLMXCF Group:		DXRDB2P	

Resync Port

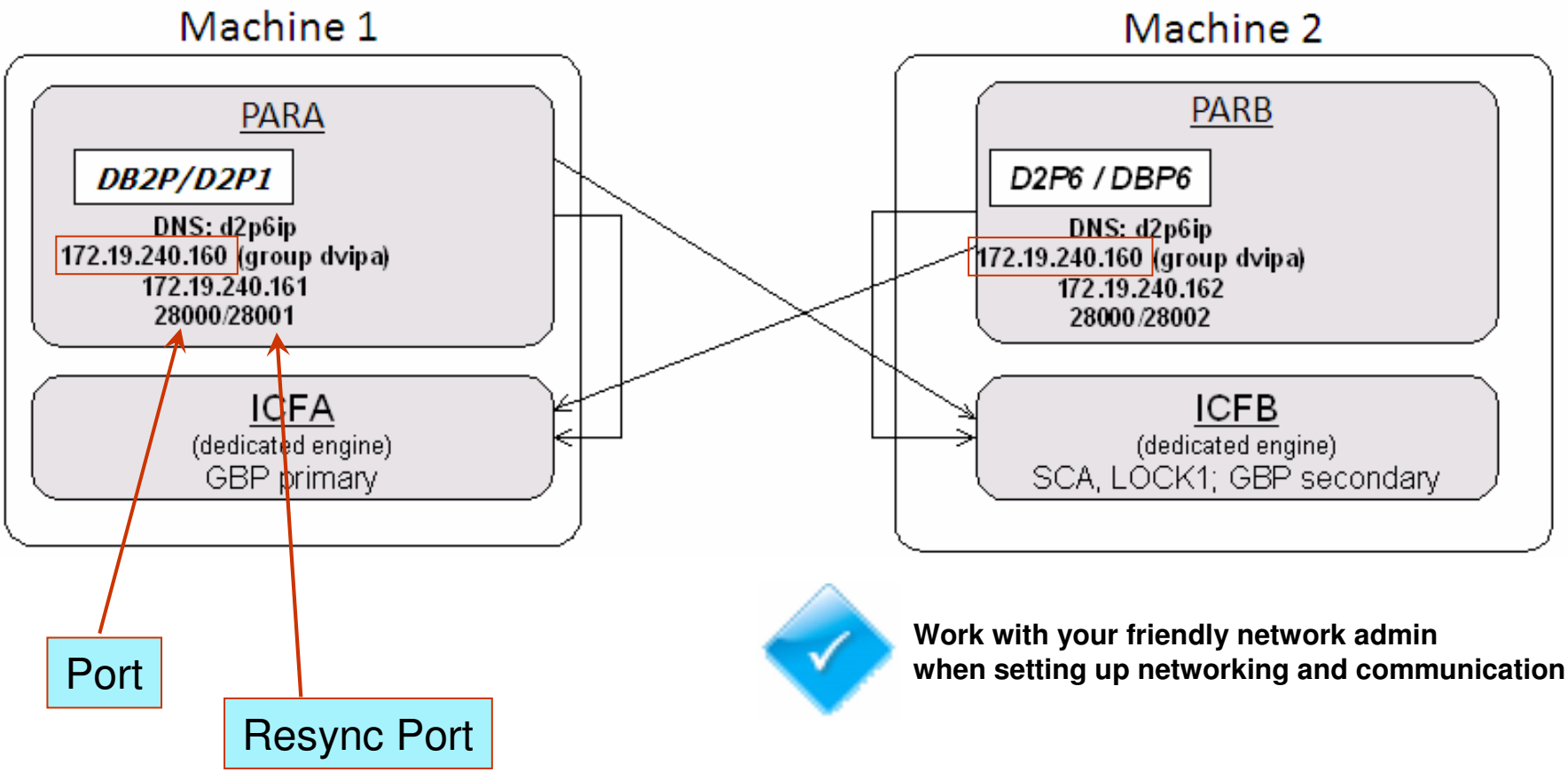
Port

Network miscellaneous items

- Dynamic virtual IP addressing (DVIPA) and Sysplex Distributor
- Location name: DSNTIPR

Redbook: Data Sharing in a Nutshell
Chapter 5 “Dynamic workload balancing”

Port and resync port - example



Work with your friendly network admin when setting up networking and communication

Recap on Planning – Key Items

- Establish naming conventions
- Determine DB2 configuration in the Parallel Sysplex
- SMS rules in place for shared disk for DB2 system and user data sets
- Determine sizes of DB2 structures in the coupling facility (use CFSizer)
- Change specific DSNZPARMs
- Plan for continuous availability

Recap on Planning – Deferral Items

- Workload balancing
- Identify connections
- Scheduling processes
- Review application considerations

Recap on Planning - Operations

- Plan for maintenance
- Modify automation to use new names
- Update procedures
  - Restart-in-place procedures
  - Cross-system restart procedures
    - Retained locks
    - RESTART(LIGHT))
  - Diagnostic procedures
- Train operators in new procedures

Execute the plan

- Change WLM policy definitions
- Define DB2 structures in the Coupling Facility
- Enable data sharing for the first DB2 member
- Add an additional data sharing member

z/OS MVS Planning: Workload Management, SA22-7602
Redbook System Programmer's Guide to: Workload Manager, SG24-6472

- Educate
- Planning
- Execution

Post-implementation

- Test data sharing function.
- Monitor behavior of a data sharing group
- Optionally - add other DB2 subsystem into data sharing group
- Optionally - run two DB2 members from the same data sharing group on the same z/OS image
- Optionally – Set up sysplex query parallelism

DB2 for z/OS Data Sharing - Planning and Administration:
http://publib.boulder.ibm.com/infocenter/dzichelp/v2r2/index.jsp?topic=/com.ibm.db2.doc.dshare/z310pln.htm

What worked well on past data sharing engagements

- Good communications
- Aggressive name standard
- Research, research, research
- Know the tooling in place
- Be mindful of changes
- CPU overhead of LPAR coupling facility
- Identify and avoid Single Points of Failure (SPOF)

What worked well on past data sharing set-ups

- Ironing-out/document naming standards before implementation
- Exploit rolling maintenance and upgrades for higher availability
- Use separate libraries (SDSN*, SDX*) at the member level
- Recovery testing (especially for the SYSPROGs and OPERATORS)

Coupling Facility

- Using CFSizer for CF initial sizing
- Coupling Facility Configuration
  - Optimal is three
  - Plenty of storage
- Sufficient space for all the structures
- Location of structures

What worked well on past data sharing set-ups

Coupling facility

- Hardware duplexing if you are using ICF on the same processors
  - Prefer at least one Standalone CF if possible
  - Allows duplexing of non-software duplex structures such as
    - DB2 LOCK and SCA
    - CICS Name Structure
    - etc.
- Coupling Data Set layouts and formatting
- Adequate DUMP space!
- The number of XCF signaling paths
- The number of physical paths (Coupling Links)
- Sysplex Failure Management (SFM) needs to be setup appropriately
- Additional PARMLIB member

Work with your friendly z/OS systems programmer when setting up the CF for Parallel Sysplex and Data Sharing

Need a jump-start?

Skill gap?

Need it yesterday?

IBM Systems Lab Services and Training

www.ibm.com/systems/services/labservices

stgls@us.ibm.com ● Systemz@us.ibm.com

Jeff M. Sullivan

jeffsull@us.ibm.com

IBM systems and Technology Group Lab Services
http://www-03.ibm.com/systems/services/labservices/