



IBM Software Group

DB2 Database Version 9

Vom Backup zu High availability




Joachim Stumpf
IM Technical Sales
IBM Information Management
joachim.stumpf@de.ibm.com



IT Symposium 2007 | Date : 18.4.2007 | DB2 Availability


© 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software



Agenda

- **Overview of terminology**
- **Introduction**
- **Basic extensions**
- **Scenarios**



DB2 | High availability |

© 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software IBM

Terminology

- Cluster usage
 - f* Availability - maximize database uptime
 - f* Scalability - maximize performance
- Availability focus
 - f* Disaster recovery
 - emphasizes protection against loss
 - Backups, log retention ,advanced storage backup
 - f* High availability
 - emphasizes data availability
 - f* both use similar technology


DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software IBM


Financial impact of an outage

Industry	Downtime Cost per Hour
Retail Brokerage	\$6.45 M
Credit card Authorizations	\$2.6 M
Catalog sales center	\$90 K
Airline reservations	\$90 K
ATM service providers	\$15 K

Source: Contingency Planning Research, Inc.





DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Agenda


- **Overview of terminology**
- **Introduction**
- **Basic extensions**
- **Scenarios**

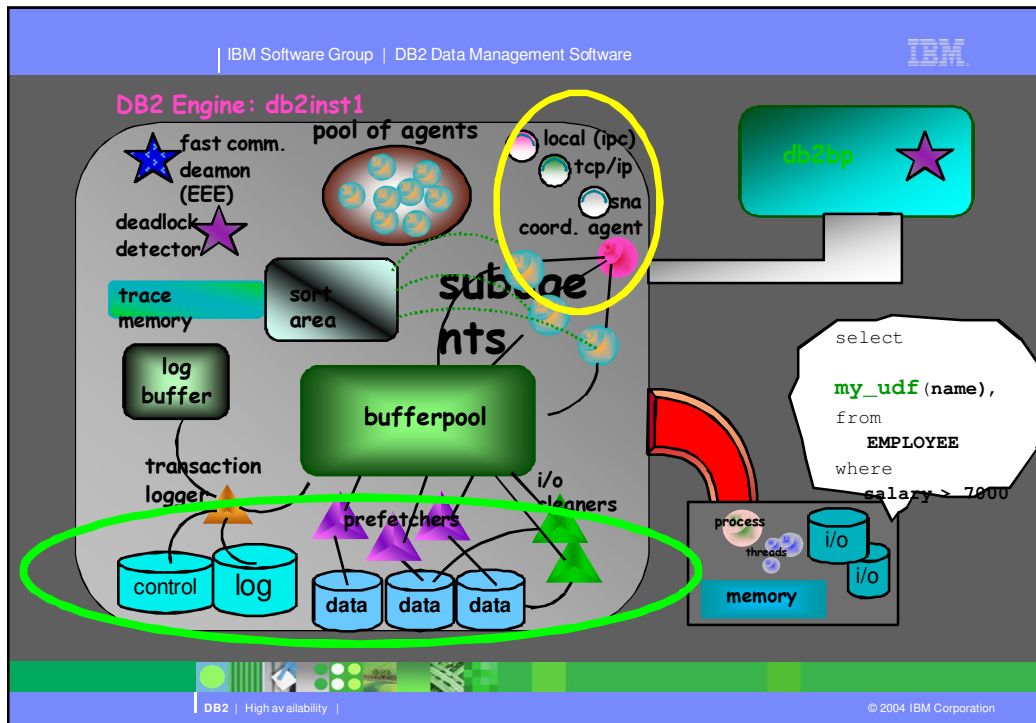
DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

High Availability Requirements

- Increasing requirement to have continuous 24x7 availability
 - f* business to business procurement
 - f* e-commerce portal
 - f* Application Service Provider markets
- Reduce
 - f* planned downtimes
 - Runstats
 - Reorg
 - Backup
 - Load, Index management
 - Configuration changes
 - f* unplanned downtimes
 - disk failure
 - machine failure
 - Network devices broken

DB2 | High availability |  © 2004 IBM Corporation



Critical database resources

- machine (memory , CPU ,board)
 - f database manager (communication / IP)
 - f database (Bufferpool,..)
- disks
 - f Logfiles (drives/storage)
 - f Data (tablespaces->drives/storage)

IBM Software Group | DB2 Data Management Software


DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Hardware support a High availability solution


- RAID
- Intelligent storage subsystems and infrastructure
 - local FlashCopy / mirroring / replication
 - synchronous Peer-to-Peer Remote Copy (PPRC)
 - support high availability and fault tolerance through mirroring techniques
- Redundant network infrastructure
- Redundant computer hardware


DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Software support a High availability solution

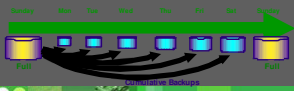
- Operating system (LVM)
 - Filesystem mirroring
- 'Heartbeat' software
 - Software to control and manage Resources
- Database extensions (mirroring / HADR)
 - Logical copy of data

DB2 | High availability |  © 2004 IBM Corporation


IBM Software Group | DB2 Data Management Software 

Base Characteristics of the database for HA

- **Reliable, Fast Database Engine**
 - ✓ Support for Flashcopy
 - ✓ Parallel & scalable design
 - ✓ Dual logging
 - ✓ Fast crash recovery
 - ☞ Efficient page cleaning
 - ☞ Parallel recovery
- **Space management**
 - ✓ Online reconfiguration
 - ✓ Dynamic Space Reclamation (container)
- **HA support tested with**
 - ✓ Steeple Lifekeeper
 - ✓ HACMP
 - ✓ MSCS
 - ✓ HP MC/ServiceGuard
 - ✓ Legato
 - ✓ Veritas
- **LOAD**
 - ✓ Fast
 - ✓ Parallel
 - ✓ Automated restart
- **Backup / Recovery**
 - ✓ Online
 - ☞ Database
 - ☞ Tablespace
 - ☞ incremental




DB2 | High availability | © 2004 IBM Corporation


IBM Software Group | DB2 Data Management Software 

Agenda

- **Overview of terminology**
- **Introduction**
- **Basic extensions**
- **Scenarios**





DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Suspend / Resume command

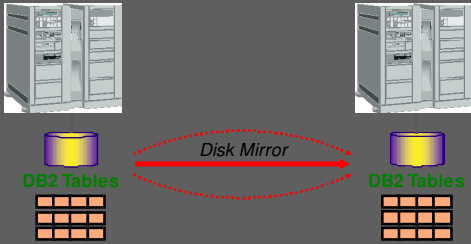
- Suspend I/O on database
 - db2 set write suspend for database
 - Suspending I/O writes will put all table spaces into a new state SUSPEND_WRITE state
 - Writes to the logs are also suspended by this command (commits are delayed)
- Resume I/O on database
 - db2 set write resume for database

DB2 | High availability |  © 2004 IBM Corporation


IBM Software Group | DB2 Data Management Software 


I/O Suspend for Split/Mirror/Flashcopy

- On-line Split Mirror
 - f* split off a consistent mirror of a database while OLTP keeps on running on the primary (live) database
 - only a limited performance hit (no commit possible)
 - f* recovery on the split off database is acceptable for consistency
 - customer cannot afford doing Off-line or On-line Backups on a 1 TB live database
 - Backups and System copies can be done from a mirror image




The diagram illustrates a disk mirror configuration. Two server racks are shown, each with a 'DB2 Tables' icon below it. A red double-headed arrow labeled 'Disk Mirror' connects the two racks, indicating that data is mirrored between them.


DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

DB2INIDB Command


- **snapshot - for Queries (cloning)**
 - results in crash recovery being performed thus making the database consistent (a new log chain is started)
 - `db2inidb <db_name> as snapshot`
- **standby - for takeover (standby)**
 - used for Standby version on a separate machine or directory
 - `db2inidb <db_name> as STANDBY relocate using config_file`
 - places the database in rollforward pending state
- **mirror - for recovery purposes** (similar to a backup)
 - mirrored copy will be used to replace the original database
 - `db2inidb <db_name> as mirror`
 - the database is placed in rollforward pending state

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Using a Split Mirror as a Standby Database

- **First four steps the same as making a clone database**
 - f* Suspend I/O on primary system
 - `db2 set write suspend for database`
 - f* Split the mirror
 - f* Resume I/O on primary system
 - `db2 set write resume for database`
 - f* Attach the mirrored database to another instance.
- **Copy logs**
 - f* set up a user exit program to retrieve log files from the primary system so that the latest logs will be available for this mirrored database
- **Initialize copy**
 - `db2inidb <db_name> as standby`
 - f* to place the mirrored database in rollforward pending state, remove the suspend write state and rollforward the database to end of logs
- **Continue this process (copy+roll-forward) until the primary database is down**

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software IBM

Making a Clone Database

- Suspend I/O on primary system:
 - db2 set write suspend for database
- Split the mirror / Start Flashcopy
 - f use operating system level command to split the mirror from the primary database
- Resume I/O on primary system
 - db2 set write resume for database
 - f database on the primary system should be back to a normal state
- Copy image to another machine
- Start the DB2 crash recovery
 - db2inidb <db_name> as snapshot
 - f You can also use this for an OFFLINE backup, but this backup, if restored on the primary system, cannot be used to rollforward because the log chain will not match

DB2 | High availability | © 2004 IBM Corporation

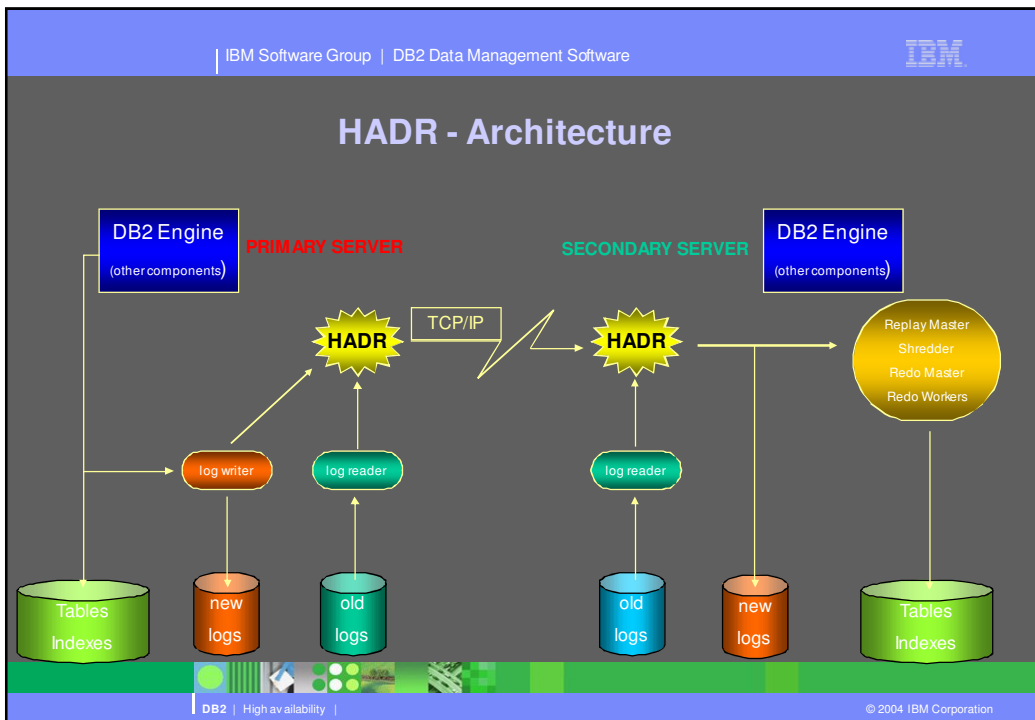
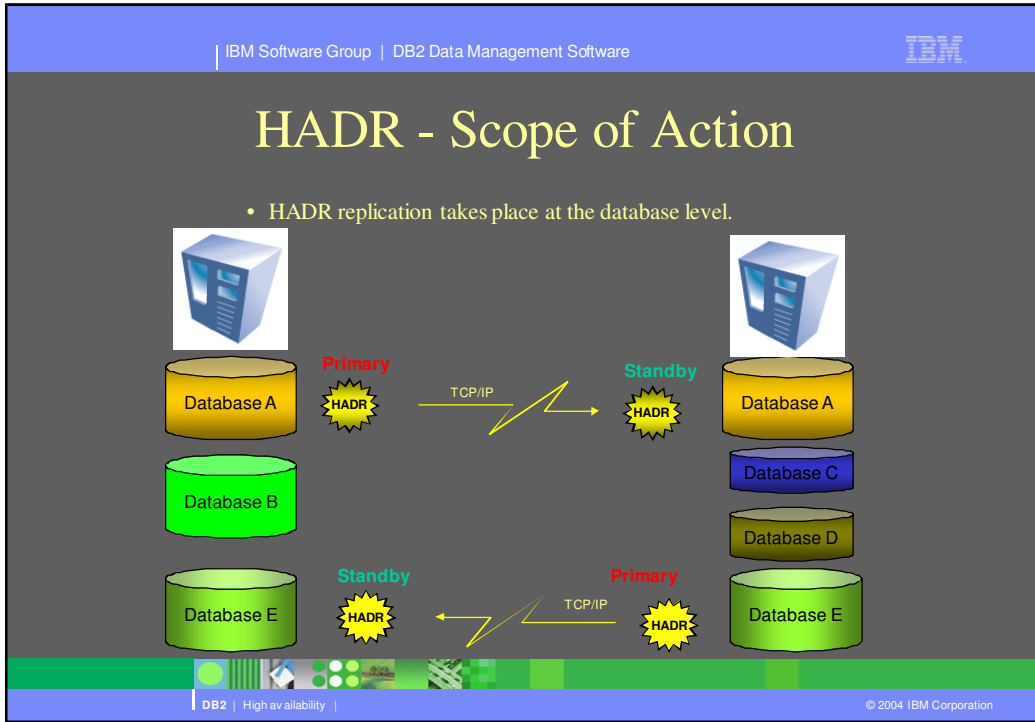
IBM Software Group | DB2 Data Management Software IBM

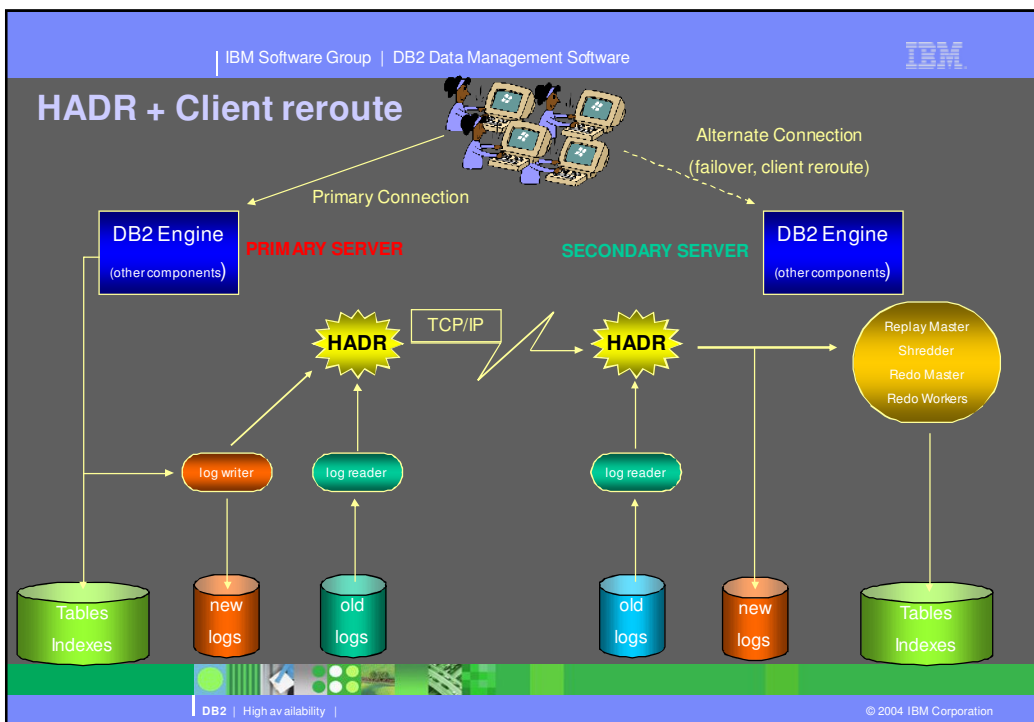
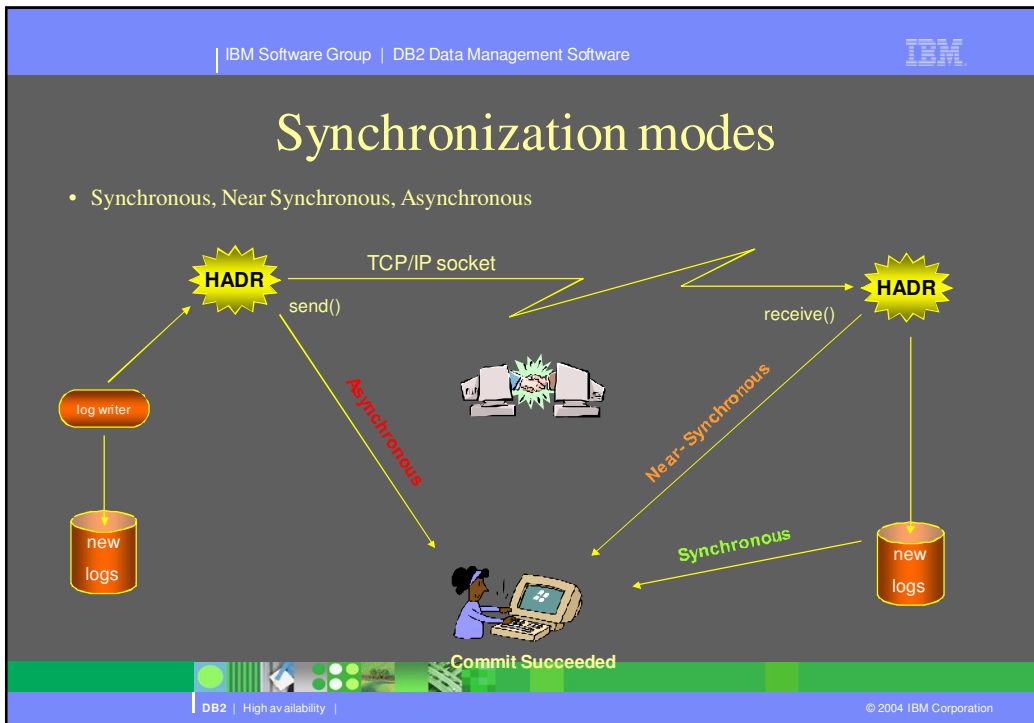
Automatic Client Reroute

- Automatic, transparent connection to alternate server when primary connection fails
 - If there is a currently executing SQL statement, it will fail with sqlcode -30108
 - Transaction can then be re-driven without re-establishing a connection
- Alternate information Stored on client
 - System database directory
 - alternateDataSource property (Java Type 4 driver)
- Works with HADR, EE/ESE, EEE/DPF, Replication

db2 update alternate server for database <dbname> using hostname <hhh> port <nnn>

DB2 | High availability | © 2004 IBM Corporation





IBM Software Group | DB2 Data Management Software

IBM

HADR Setup: Step 1

- Clone the primary
 - Enable LOGARCHIVEMETHOD (LOGRETAIN)
 - Backup primary database
 - Restore db on standby
 - After the restore, the standby is placed in "perpetual rollforward" mode
- Configure HADR

Primary

Standby

Split Mirror
Flash Copy
db2inidb

backup

Backup database db2_hadr on <path>

Restore database db2_hadr

DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software

IBM

HADR Setup: Step 2

- Clone the primary
- Configure HADR
 - Configure the database on both sides
 - HADR_LOCAL_HOST ALPHA1
 - HADR_LOCAL_SVC Port_ALPHA1
 - HADR_REMOTE_HOST ALPHA2
 - HADR_REMOTE_SVC Port_ALPHA2
 - HADR_REMOTE_INST DB2CTLSV
 - HADR_SYNCMODE NEARSYNC
 - HADR_TIMEOUT 120
 - LOGINDEXBUILD ON
 - Update services file

Primary

Standby

Update db cfg for db using

Update db cfg for db using

DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software

IBM

HADR Operation start: Step 1

- Start the standby
 - Command
 - Start **hadr on database <db2_hadr> as standby**
 - Local catch-up
 - Replay locally available log files (if any)
 - Replay log files that may be retrieved at the primary and shipped to the standby
 - Remote catch-up (primary must be available)
 - Replay log pages from the primary's **archived** logs
 - Replay log pages from the primary's **active** log to the standby until the standby is caught up to the tail of the log.
 - Replay log pages from the primary's **in-memory** log buffer to the standby whenever the primary flushes those pages to disk
 - Peer State (primary must be available)
 - Primary and standby work in sync
- Start the primary

DB2 | High availability | © 2004 IBM Corporation


IBM Software Group | DB2 Data Management Software

IBM

HADR Operation Start : Step 2

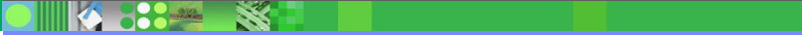
- Start the standby
- Start the primary
 - Command
 - Start **hadr on database <db2_hadr> as primary (by force)**
 - If the standby does not show up after a waiting period (configurable), HADR will not start.
 - This behavior avoids that two systems could work as primary at the same time (split brain).
 - Unless started "by force", the HADR primary waits for the standby to contact it.


DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Two flavors of TAKEOVER


- Emergency TAKEOVER (by force)
 - The standby does NOT wait for any acknowledgement from the primary
 - Standby takes primary role
- Normal TAKEOVER (no force)
 - Primary and Standby switch roles
 - i.e. to run Rolling upgrades

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Scenarios

- Failover / Disaster Recovery
 - f* Backup / Restore
 - f* Log shipping
 - f* Replication
 - f* HA software
- continuous availability

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software

IBM

Backup / Restore Procedure

- Enable LOGARCHIVEMETHOD (LOGRETAIN)
- Backup database regularly
- Restore db on new machine to End of LOGS

The diagram illustrates the backup and restore process. On the left, a server icon and a yellow cylinder represent the source database. A green arrow points from the source to a yellow box labeled 'Backup + Logs'. Above this arrow is a horizontal bar with colored segments (red, orange, green, blue, purple) and a red arrow pointing right. From the 'Backup + Logs' box, another green arrow points to a server icon and a yellow cylinder on the right, labeled 'Standby' and 'Restore database dbx'. Below the source server is the text 'Backup database dbx on <path>'. Below the destination server is the text 'Restore database dbx'. At the bottom left of the slide, it says 'DB2 | High availability |'. At the bottom right, it says '© 2004 IBM Corporation'.

DB2 | High availability |

© 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software

IBM

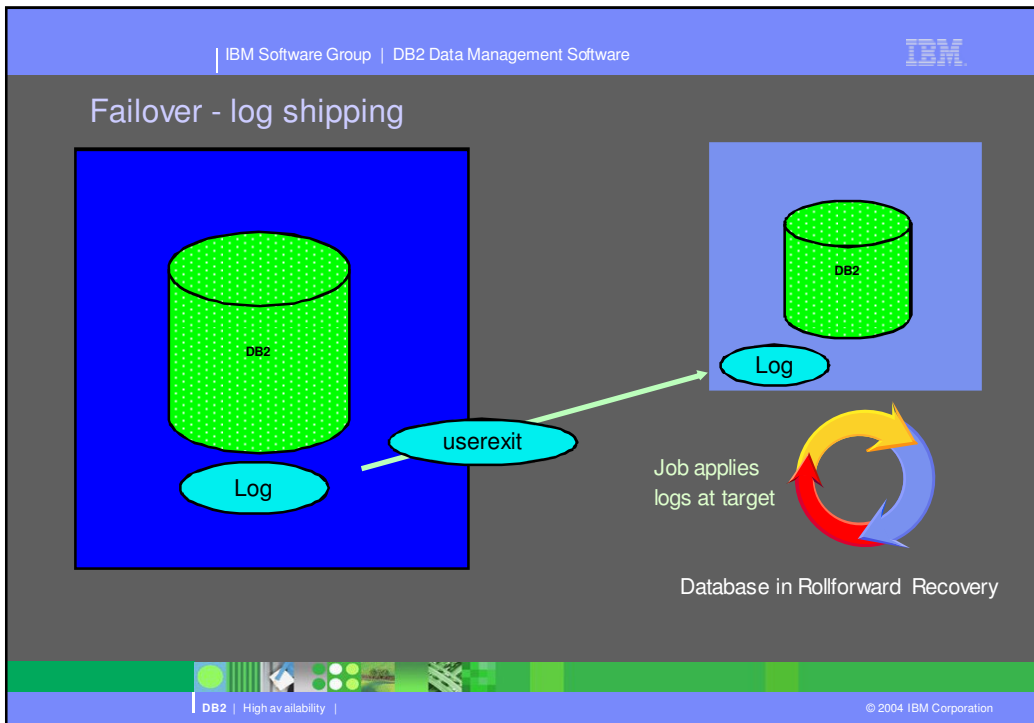
Backup / Restore

- Backup images made on regular base and shipped
- Logs are archived and shipped
- Scripts for install, create and configure are available

Advantage	Disadvantage
no extra software ?	Additional procedures or software
less distance limitations	No automatic failover
System created on Demand	Time consuming

DB2 | High availability |

© 2004 IBM Corporation



IBM Software Group | DB2 Data Management Software

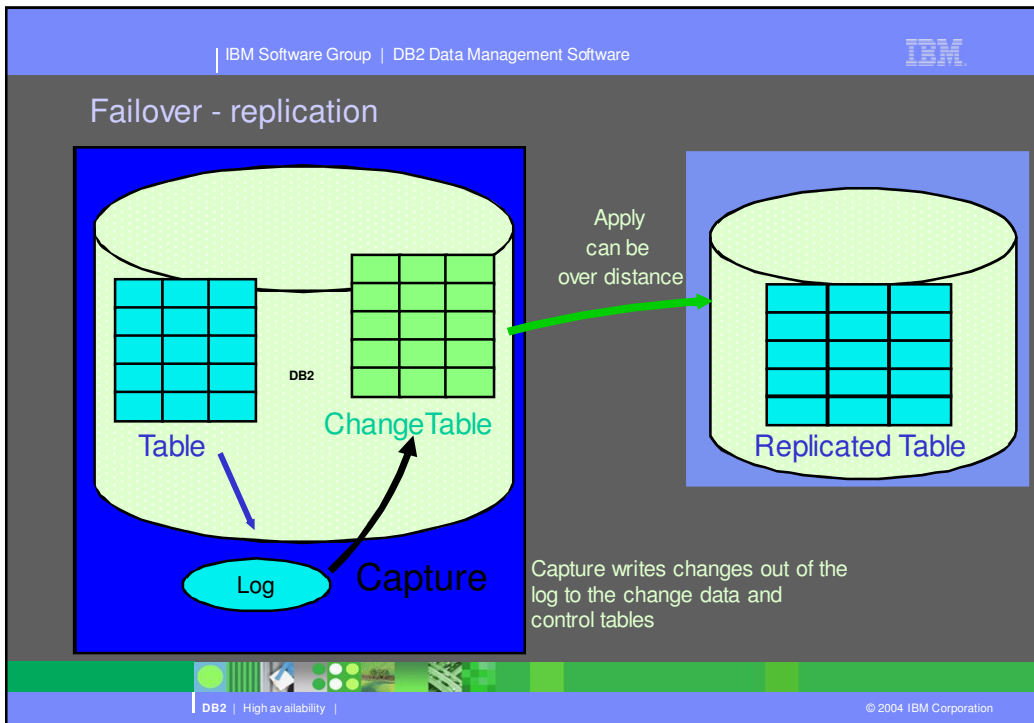
IBM

Failover - log shipping

- image of the database is held in parallel to the default system in rollforward recovery
- can be done with with procedures to get and apply logs
 - f* redirected restore
 - f* db2inidb
 - f* or third Party software like Libelle

Advantage	Disadvantage
no extra software ?	Additional procedures or software
less distance limitations	No automatic failover
	second System is passive

DB2 | High availability | © 2004 IBM Corporation



IBM Software Group | DB2 Data Management Software

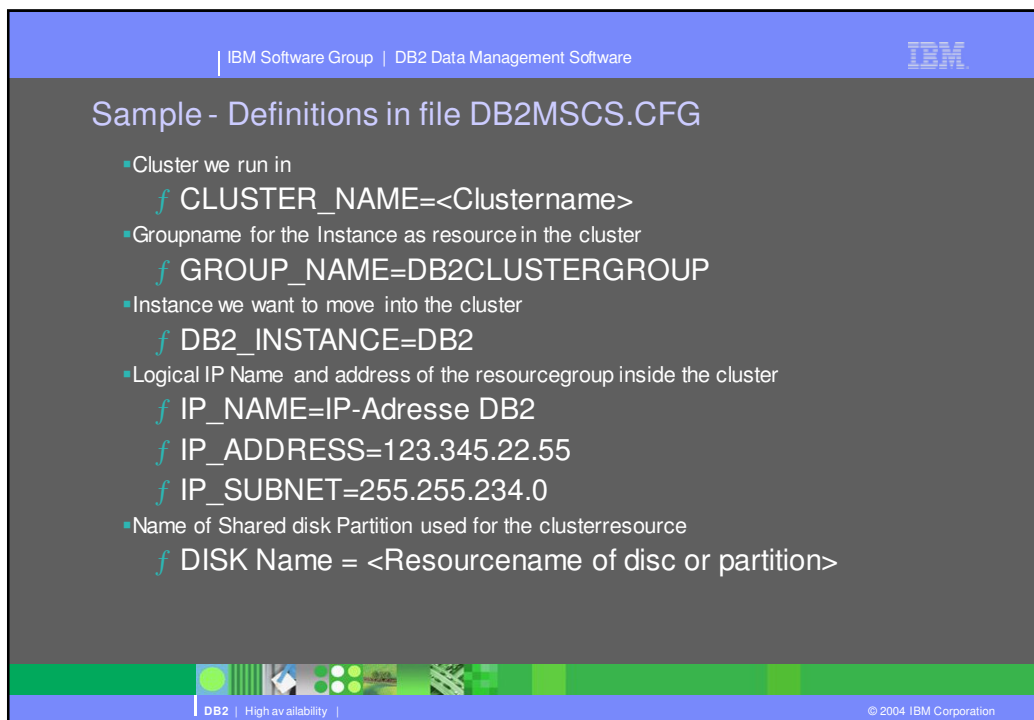
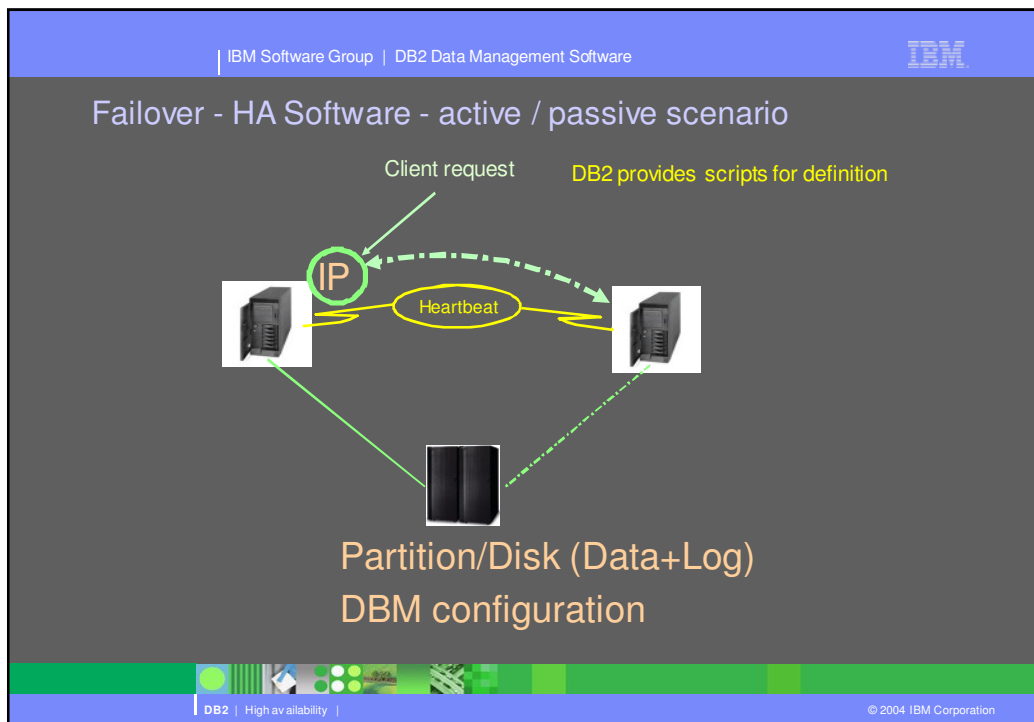
IBM

Failover - replication

- image of the database is held in parallel on a separate system
- changes on the leading system are applied through Apply process

Advantage	Disadvantage
no extra software	Asynchronous process, potential loss of data
less distance limitations	No automatic failover
mirror usable	extra workload on primary system

DB2 | High availability | © 2004 IBM Corporation



IBM Software Group | DB2 Data Management Software

IBM

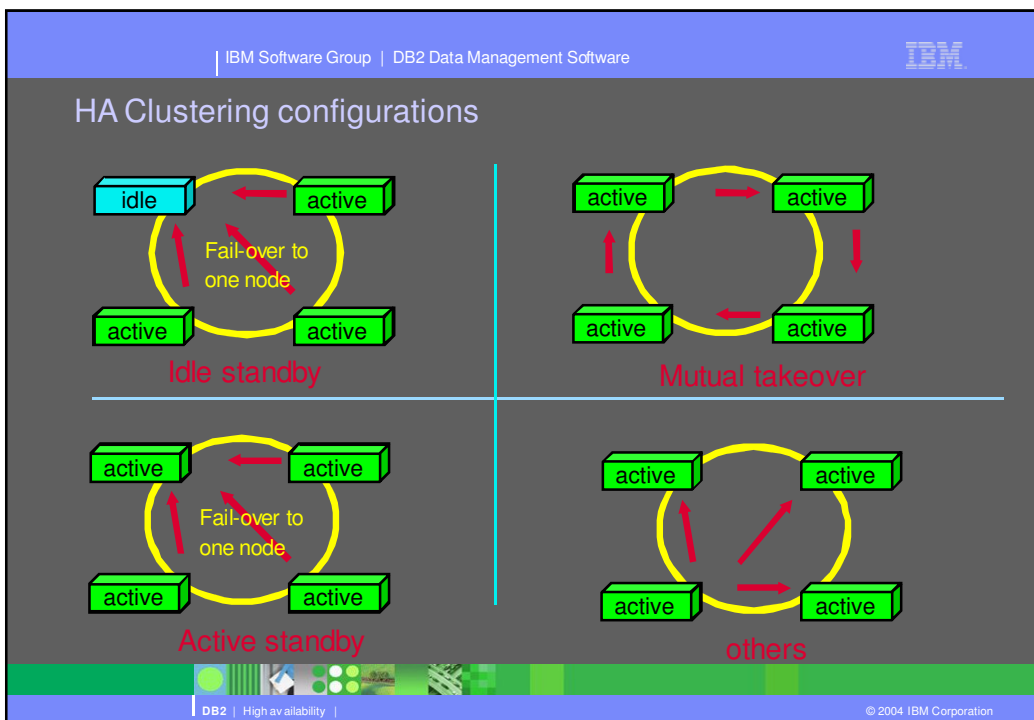
Sample Procedure to implement

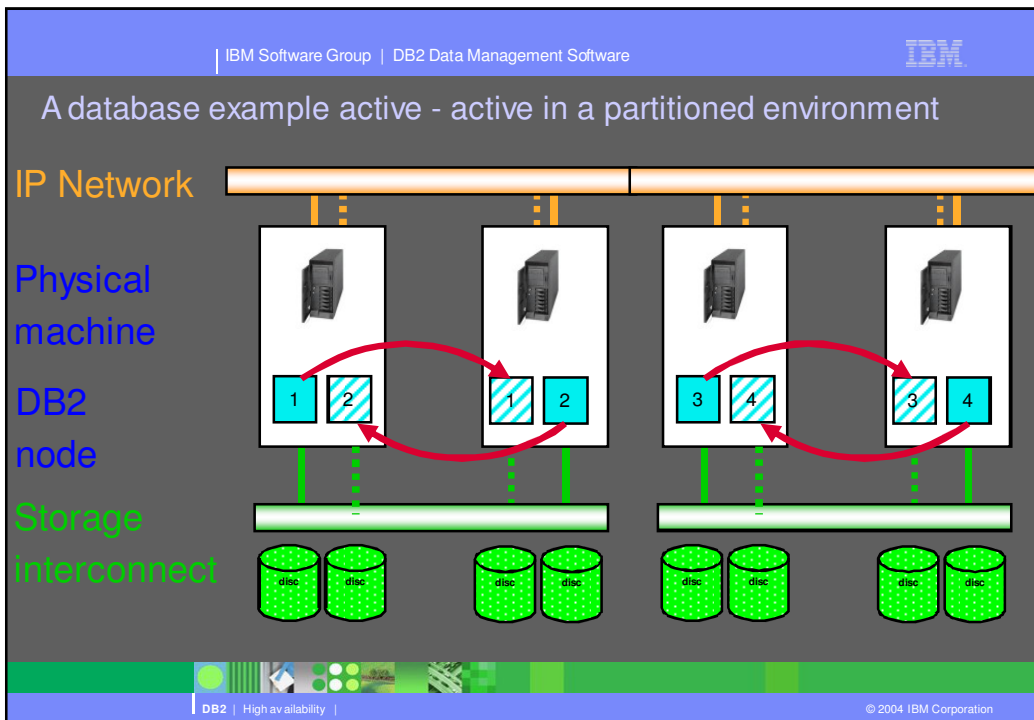
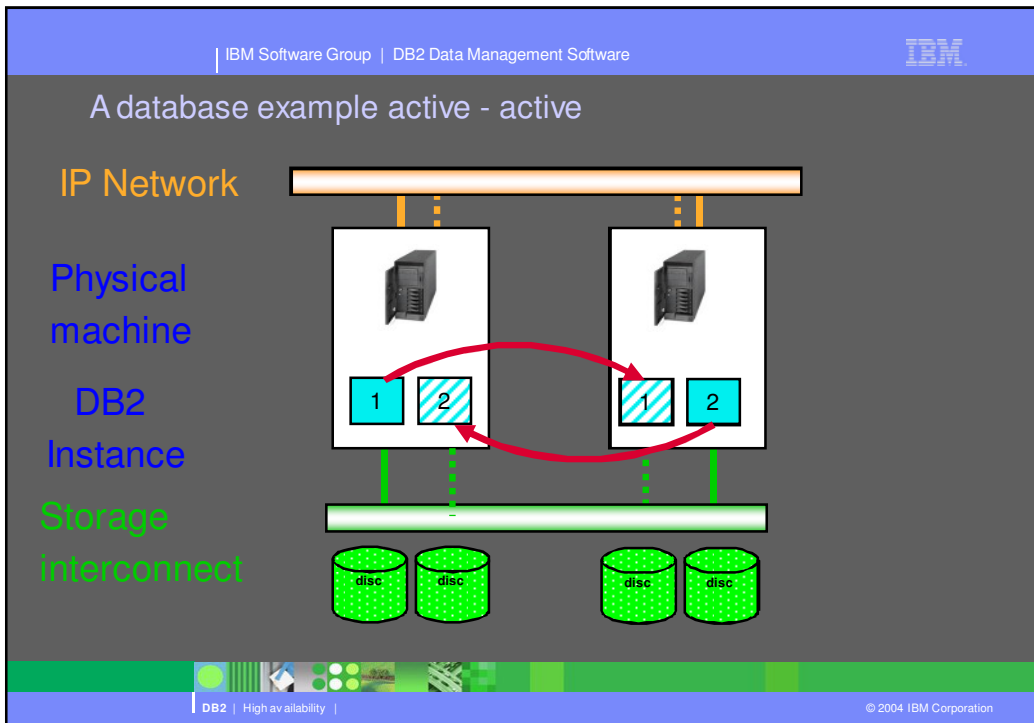
- Install DB2 on both machines on internal disks
- Create an instance on one machine
- Create shared disc resources
- Get info about internal IP Network
- Complete the configuration file DB2MSCS.CFG
- Run DB2MSCS utility

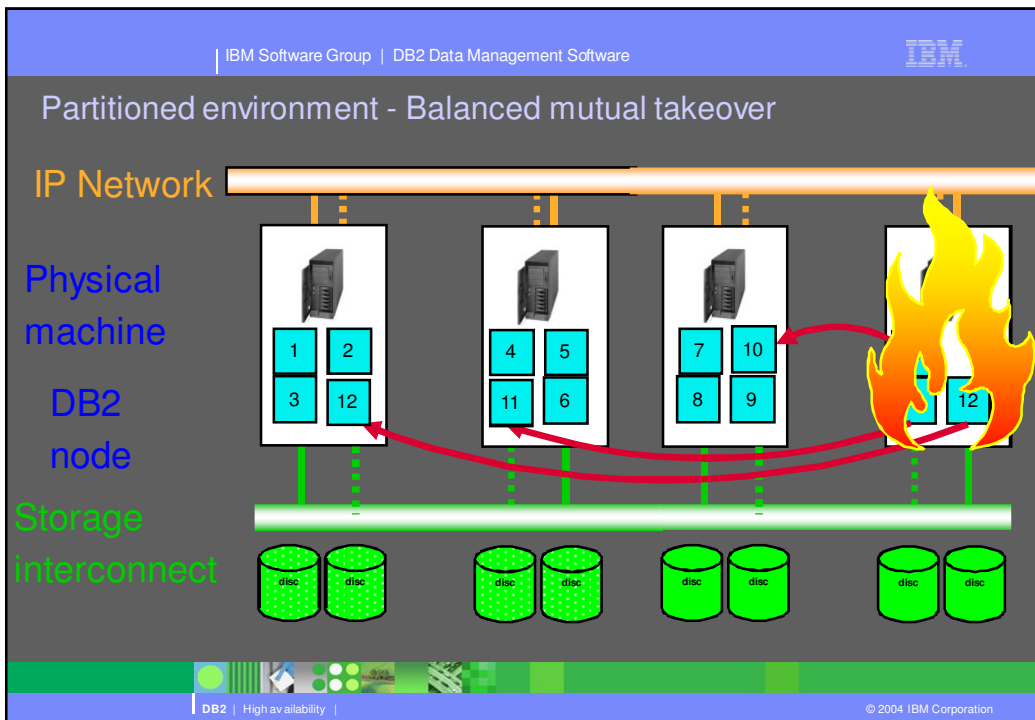
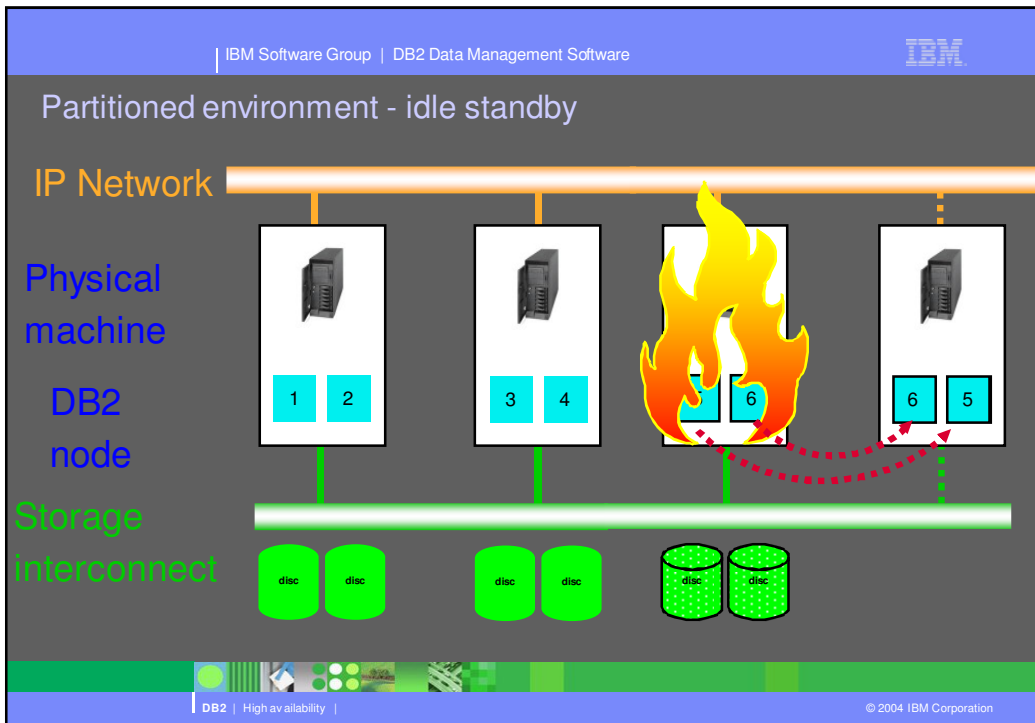
Run with

```
db2mscs -f DB2MSCS.CFG [-d:<tracefilename>]
```

DB2 | High availability | © 2004 IBM Corporation







IBM Software Group | DB2 Data Management Software

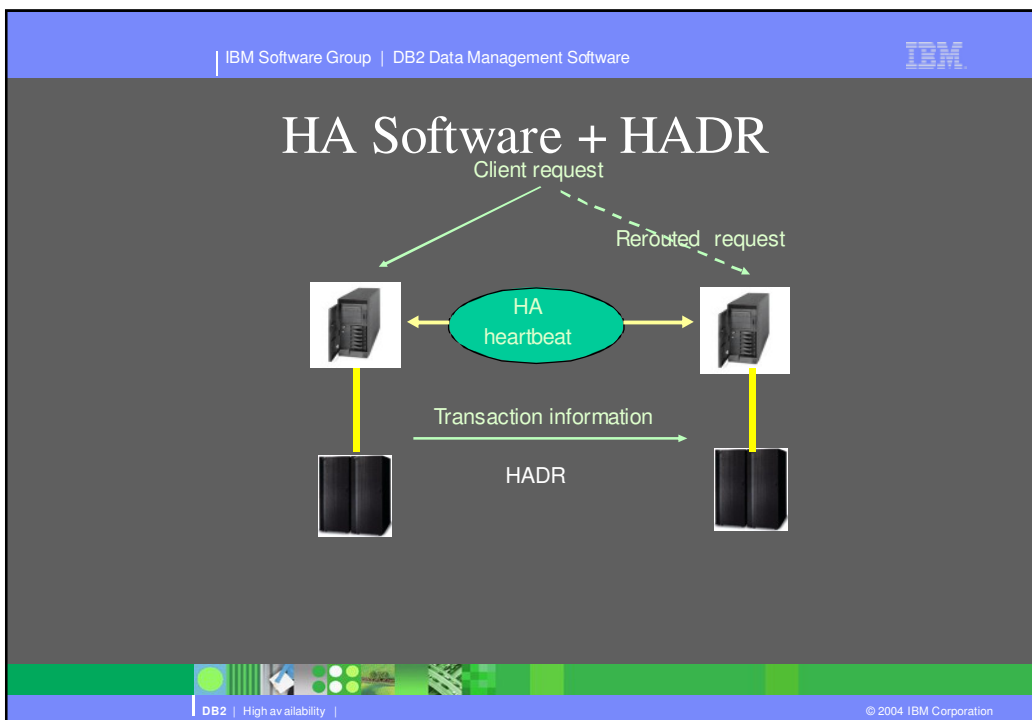
IBM


Failover - HA Software

- database has to reside on an external disk system
- machine should have two network adapters
- resources have to be defined in the HA software
- HA Software:
 - f AIX – HACMP ,TSA
 - f HP MC/Service Guard
 - f Windows - Microsoft cluster server or Steeleye lifekeeper
 - f Linux - Steeleye lifekeeper , TSA ,
 - f Sun - Sun cluster

Advantage	Disadvantage
automatic takeover	Filesystem takeover maybe with add. checks
short delay	Open transactions are rolled back
	Application should be HA aware

DB2 | High availability | © 2004 IBM Corporation





IBM Software Group | DB2 Data Management Software 

Failover – HADR + HA Software

- resources have to be defined in the HA software
 - Definition requires less information
- HA Software:
 - f AIX – HACMP , TSA (Policy available)
 - f Windows - Microsoft cluster server or Steeleye lifekeeper
 - f HP MC/ServiceGuard
 - f Linux - Steeleye lifekeeper , TSA (Policy available)
 - f Sun - Sun cluster


Advantage	Disadvantage
Data duplicated on logical level	Application should be HA aware
In case of takeover only crash recovery needed (FAST)	Open transactions are rolled back Only information logged is shipped

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Policy content

- Definition scripts
- Start script
- Stop script
- Monitoring scripts

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software IBM

Continuous Availability

Client request
↓
TM
↓
[Server 1] [Server 2] [Server 3]
↓ ↓ ↓
[Storage 1] [Storage 2] [Storage 3]

Requests send to both or more machines
Solution needs a Transaction manager / Coordinator

DB2 | High availability | © 2004 IBM Corporation


IBM Software Group | DB2 Data Management Software IBM

Continuous Available

- SQL transaction takes place on two machines
- redundant / additional hardware
 - f Marathon solution
 - f Avioka, XKoto
- additional software (transaction manager)
- Programs are HA aware


Advantage	Disadvantage
nearly 100% Uptime	Additional coding
Rolling version upgrades	Requires add. Hardware
	Different API

DB2 | High availability | © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

Referenzen

- DB2 Developer Domain
 - [f http://www.ibm.com/developerworks/db2/library/techarticle/0304wright/0304wright.html](http://www.ibm.com/developerworks/db2/library/techarticle/0304wright/0304wright.html)
 - [f ftp://ftp.software.ibm.com/ps/products/db2/info/vr9/pdf/letter/en_US/db2hae90.pdf](ftp://ftp.software.ibm.com/ps/products/db2/info/vr9/pdf/letter/en_US/db2hae90.pdf)
 - [f http://www-306.ibm.com/software/data/hp/library.html](http://www-306.ibm.com/software/data/hp/library.html)

DB2 | High availability |  © 2004 IBM Corporation

IBM Software Group | DB2 Data Management Software 

The End



DB2 | High availability |  © 2004 IBM Corporation