



Sdružení uživatelů Informix & DB2

Konfigurace vysoké dostupnosti HADR

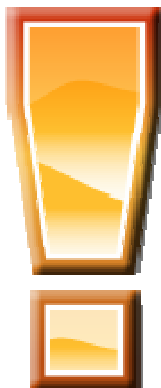
Ing. Jan Musil, IBM ČR
Community of Practice for
CEEMEA

Základní přehled vlastností

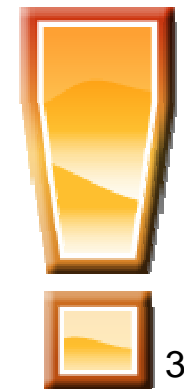
- Implementace Informix Dynamic Server HDR technologie
- Poskytuje řešení vysoké dostupnosti v případě částečného nebo úplného selhání databázového serveru
- Chrání před ztrátou dat replikováním změn v datech ze zdrojové databáze na cílovou databázi
 - Zdrojová databáze = primární
 - Cílová databáze = záložní
- Aplikace mohou pracovat pouze s primární databáze
- Změny do záložní databáze se provádí na základě roll forward žurnálů generovaných na primárním serveru
- Řešení umožňuje přepnutí na záložní server během několika sekund
 - Bez HADR může zpětná synchronizace databáze po selhání trvat minuty (částečné selhání) nebo i podstatně déle (úplné selhání)
- Přepnutí aplikace na záložní server lze provést automaticky
- V případě havárie primárního serveru lze záložní server okamžitě přepnout na primární s plnou funkcionalitou
- Po odstranění havárie původního primáru lze převést zpracování zpět (fail back)

Předpoklady pro použití HADR

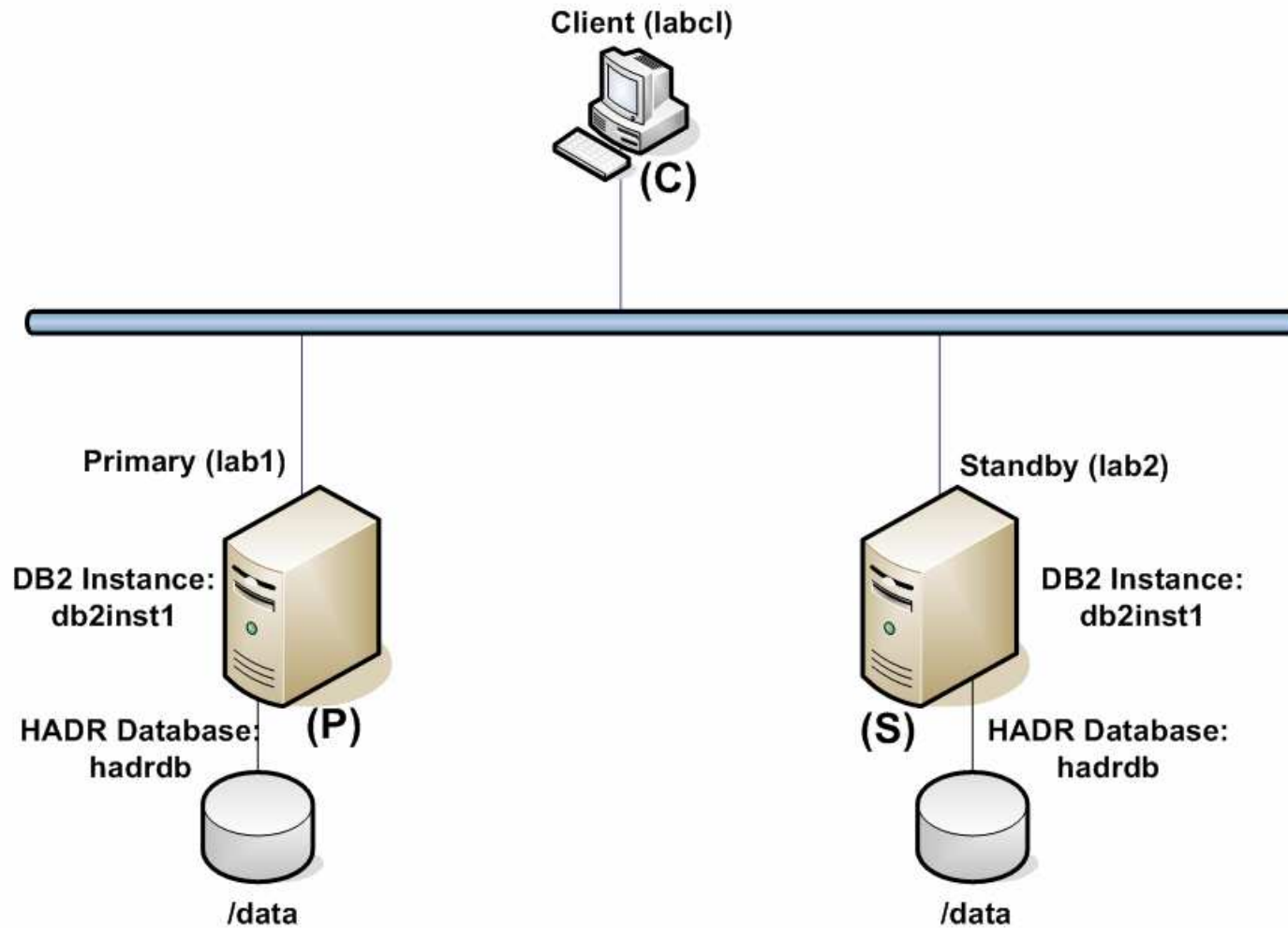
- Spolehlivé TCP/IP síťové propojení primárního a záložního serveru
- Identická verze operačního systému na obou serverch
- Identická verze DB2 (např. DB2 9) a stejná „bit“ architektura (32 nebo 64 bit)
- Stejně jméno DB2 databáze
- DB2 table spaces musí mít stejnou velikost a identický typ, definice kontejnerů musí být identická (cesta, velikost a typ souborového systému)
- Velikost buffer poolu musí být stejná
- Přidělený prostor pro žurnálovací soubory musí být stejný



- Identické počítače (stejný výrobce a typ)
- Vysoká rychlost a kapacita síťového propojení
- Primární a záložní server mají k dispozici stejnou velikost paměti
- Konfigurační parametry databází a instancí jsou stejné



Typická architektura



Módy synchronizace

- Plně synchronní (synchronous)
- „Téměř“ synchronní (Near synchronous)
- Asynchronní (asynchronous)

- Definují se konfiguračním parametrem HADR_SYNCMODE
 - SYNC (počítače fyzicky spolu)
 - NEARSYNC (LAN, WAN)
 - ASYNC (WAN)

Plně synchronní mód



Garantuje nejvyšší stupeň ochrany proti ztrátě transakcí



Delší transakční odezva

1. (P) Posílá transakční žurnály na záložní server a blokuje jejich zápis na disk
2. (S) Zapisuje transakční žurnály na disk a posílá potvrzení primárnímu serveru
3. (P) Po přijetí potvrzení zapisuje transakční žurnály na disk

Téměř synchronní mód (default)



Kratší transakční odezva, nejlepší kompromis mezi spolehlivostí a výkonem



Menší míra ochrany proti ztrátě transakcí (ztráta pouze v případě, kdy oba servery selhávají současně a záložní server nestačí zapsat transakční žurnály na disk)

1. (P) Posílá transakční žurnály na záložní server a blokuje jejich zápis na disk
2. (S) Posílá potvrzení primárnímu serveru ihned po přijetí transakčních žurnálů (nečeká na jejich zápis na disk)
3. (P) Po přijetí potvrzení zapisuje transakční žurnály na disk

Asynchronní mód



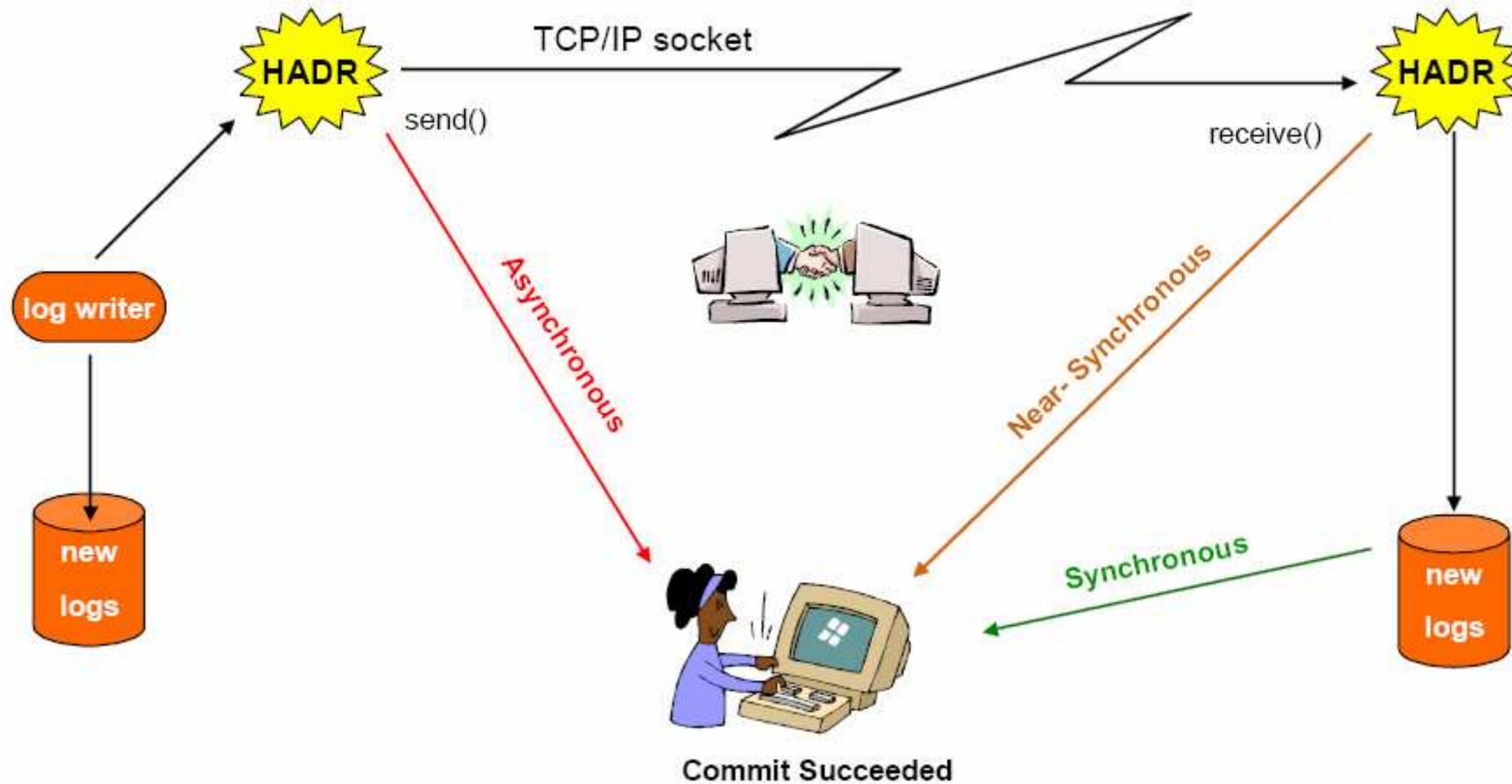
Krátká transakční odezva



Nejvyšší míra transakční ztráty v případě selhání primárního serveru („commit“ transakce na primárním serveru negarantuje „commit“ transakce na záložním serveru)

1. (P) Posílá transakční žurnály na záložní server a okamžitě je zapisuje na disk
2. (S) Přijímá a zapisuje transakční žurnály na disk

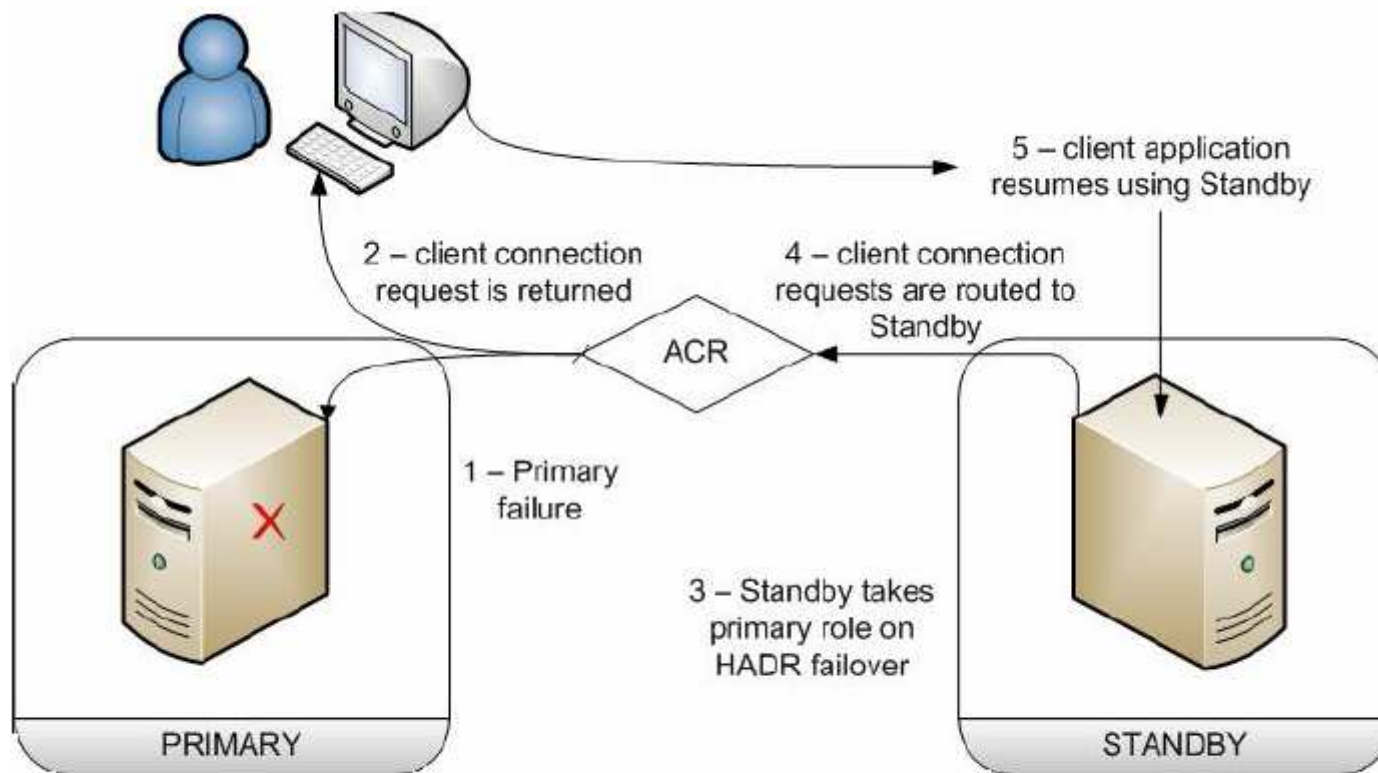
Shrnutí jednotlivých typů synchronizace



Fáze konzistence HADR serverů

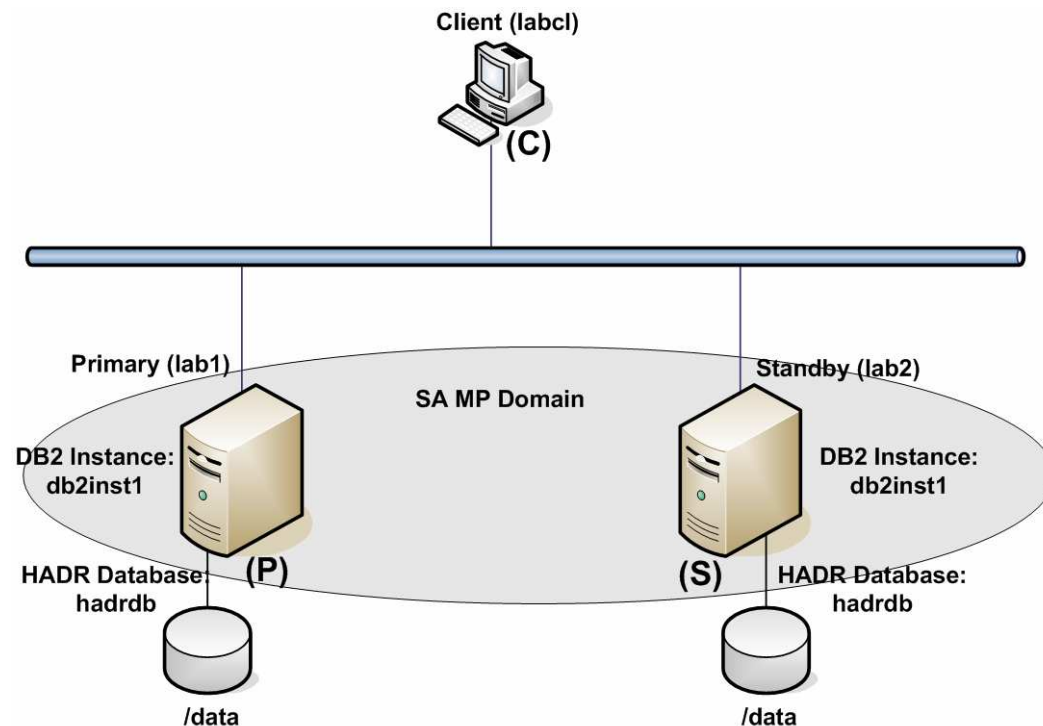
- „catch-up“ fáze
 - Primární server posílá záložnímu serveru všechny logické transakční žurnály, které jsou potřeba k uvedení záložního server do stejného okamžiku konzistence, jako je primární server
- „peer“ fáze
 - Primárního a sekundárního serveru jsou datově konzistentní (důsledek úspěšné catch-up fáze)

Přesměrování klientů



Architektura s IBM Tivoli SA MP

- IBM Tivoli SA MP = IBM Tivoli System Automation for Multiplatforms
 - Součástí instalace DB2
- Tivoli SA MP zajišťuje
 - Automatické monitorování
 - Automatické fail over



Demonstrace

- Nastavení a konfigurace
- Rolling upgrade
- Přepnutí při selhání

Konfigurace HADR prostřednictvím CLP (1/2)

- (P) db2 update database configuration for sample using LOGRETAIN recovery
db2 update database configuration for sample using LOGINDEXBUILD ON
- (P) db2 backup database sample to /usr/db2/backup
- (S) db2 restore database sample from /usr/db2/backup taken at 20061011141321
replace history file
- (P) db2 update alternate server for database sample using hostname polonium
port 50001
- (S) db2 update alternate server for database sample using hostname lead port
50001
- (P)(S) /etc/services

```
DB2_HADR_1    55001
DB2_HADR_2    55002
```

Konfigurace HADR prostřednictvím CLP (2/2)

- (P) db2 update db cfg for sample using HADR_LOCAL_HOST LEAD
 db2 update db cfg for sample using HADR_LOCAL_SVC DB2_HADR_1
 db2 update db cfg for sample using HADR_REMOTE_HOST POLONIUM
 db2 update db cfg for sample using HADR_REMOTE_SVC DB2_HADR_2
 db2 update db cfg for sample using HADR_REMOTE_INST DB2INST1
 db2 update db cfg for sample using HADR_SYNCMODE SYNC
 db2 update db cfg for sample using HADR_TIMEOUT 3
 db2 connect to sample
 db2 quiesce database immediate force connections
 db2 unquiesce database
 db2 connect reset

- (S) db2 update db cfg for sample using HADR_LOCAL_HOST POLONIUM
 db2 update db cfg for sample using HADR_LOCAL_SVC DB2_HADR_2
 db2 update db cfg for sample using HADR_REMOTE_HOST LEAD
 db2 update db cfg for sample using HADR_REMOTE_SVC DB2_HADR_1
 db2 update db cfg for sample using HADR_REMOTE_INST db2inst1
 db2 update db cfg for sample using HADR_SYNCMODE SYNC
 db2 update db cfg for sample using HADR_TIMEOUT 3

- (S) db2 deactivate database sample
 db2 start hadr on database sample as standby
- (P) db2 deactivate database sample
 db2 start hadr on database sample as primary

Základní operace prováděné v HADR prostředí

- Spouštění

 - (S) db2 start hadr on database sample as standby

 - (P) db2 start hadr on database sample as primary [by force]

- Zastavování

 - (P) db2 deactivate database sample

 - db2 stop hadr on database sample

 - (S) db2 deactivate database sample

 - db2 stop hadr on database sample

Přepínání (take over)

- Záměna rolí

(S) db2 takeover hadr on database sample

- 1.(S) informuje (P), že dojde k přepnutí
- 2.(P) ukončí všechna připojení a nedovolí nová připojení
- 3.(P) odroluje všechny otevřené transakce zašle zbytek žurnálu na (S)
- 4.(S) přehraje obdržené transakční logy
- 5.(P) se stává novým záložním serverem
- 6.(S) se stává novým primárním serverem

- Failover

(S) db2 takeover hadr on database sample by force

- 1.(S) zašle požadavek, aby se (P) ukončil
- 2.(S) nečeká na žádná potvrzení o stavu (P)
- 3.(S) zastavuje přijímání žurnálů od (P) a stává se novým primárním serverem

Obnovení HADR z failover

- Obnova původního primárního serveru a převedení do on-line stavu
- Nastartování původního primárního serveru jako nový standby
(P) db2 start hadr on database *sample* as standby
- Přepnutí rolí
(P) db2 takeover hadr on database *sample*

DB2 HADR a Q replication

Points of comparison	High availability disaster recovery	Q replication
What is the scope of the setup procedure?	Entire DB2 database.	Tables within a DB2 database or subsystem.
How is data propagated to the standby?	Log operations are shipped to the standby database and replayed continuously through forward recovery.	Committed transactional data are captured from the DB2 recovery log and applied to target tables.
Is synchronous replication possible?	Yes.	No.
Is asynchronous replication possible?	Yes.	Yes.
Are client applications automatically rerouted to the standby?	Yes.	Yes.
Which operating systems are supported?	Linux [®] , UNIX [®] , and Windows [®] .	Linux, UNIX, Windows, z/OS [®] .
Can applications read data from the standby?	No.	Yes.
Can applications write data to the standby?	No.	Yes.
Is SQL Data Definition Language (DDL) automatically replicated in addition to SQL Data Manipulation Language (DML)?	Yes.	No.
On what hardware, operating system, and version of DB2 can the standby run?	Hardware, operating system, and version of DB2 must be identical to those where the source database is located.	Hardware, operating system, and version of DB2 can be different from those where the source database is located.
Are there tools for monitoring performance?	Yes.	Yes.
Is network compression and encryption built in?	No.	Yes.
Can the source and standby databases be partitioned?	No.	Yes.

Děkuji za pozornost

jan_musil@cz.ibm.com