

MYSQL CRASH RECOVERY

*Dinko Korunić @ CARNet
(Grupa za izradu paketa)*

OMG! MySQL je super!!!!11ž



Reality check

- backup:
 - **potpunost** backup procedure i sadržaja?
 - redovno **testiranje** restore procedure?
 - **brzina** restorea u slučaju katastrofičnog incidenta?
- master-slave replikacija
 - postoji li nadzor? - postoje desetine Nagios/ZenOSS pluginova za lag check!
 - usporedba master-slave tablica:
mk-table-checksum

Uobičajene metode backupa (1)

- redovni **mysqldump**:
 - dnevni/satni/itd, selektivni ili cjelokupni
 - flush tables with read lock
 - MyISAM: --read-locks
 - InnoDB: --single-transaction
 - spor proces - moguće **inkrementalno** kroz binlog:
--single-transaction --flush-logs --master-data=2 --all-databases

Uobičajene metode backupa (2)

- druge backup metode:
 - **Bacula**: Client Run Before Job
 - **AutoMySQLBackup**
- replikacija:
 - multi-master, master-slave kombinacije (ili nešto treće) + mysqldump slavea
- Percona XtraBackup / InnoBackupEx:
 - opis: hot backup / binlog-based backup
 - streaming, komprimirani backupovi
 - paralelni backupovi, inkrementali

Uobičajene metode backupa (3)

- Percona XtraBackup / InnoBackupEx:
 - innobackupex --stream=tar ./ | ssh user@desthost "cat - > /data/backups/backup.tar"
 - backup InnoDB kroz binlog, MyISAM standardno
 - restore: kompletno (from the scratch) ili parcijalno kroz binlog (Point In Time)
 - gasi slave thread prije i aktivira poslije
 - preporučena zamjena za mysqldump!

Moguća oštećenja

- MyISAM corruption:
 - --myisam-recover=backup, force
 - *REPAIR TABLE ... USE FRM;*
- InnoDB corruption:
 - u praksi prevencija SEGV...
 - innodb_force_recovery=4
 - moguće opcije: **1, 2, 3, 4, ... 5, 6**
- binlog corruption
- FRM corruption

Alati za oporavak

- **myisamchk:** offline
 - myisamchk -e *.MYI
 - myisamchk --safe-recover *.MYI
- **mysqlcheck (mysqlrepair):** online!
 - *CHECK TABLE, REPAIR TABLE, ANALYZE TABLE, OPTIMIZE TABLE*
 - mysqlcheck -A -r
- **alternative:**
 - ako je server stabilan: dump and reload
 - trikovi: *ALTER TABLE ENGINE=...;*

Alati za ekstrakciju podataka

- **innoinfo**
 - zaostao i neupotrebljiv
- **percona-innodb-recovery-tool:**
 - bzx branch lp:percona-innodb-recovery-tool
 - sjajan offline alat za InnoDB (korupcija, neželjeni drop, itd.)
 - priprema: osnove C, format redaka (**COMPACT ili REDUNDANT**), **ibdata1 ili *.idb**, definicija tablica, raspon **fizičkih stranica željenog primarnog ključa**

Priprema laboratorija

- mysql server u recovery načinu:
 - “laboratorij”: kompilirati posljednju inačicu, pripremiti da radi samo u tkućem/radnom direktoriju
 - oporavak tipično iz primarnih ključeva
 - praktičan problem:
 - ikakav select/update/merge uzrokuje SEGV
 - undo log purge – uzrokuje finalno brisanje
 - innodb_force_recovery=4
 - mysqld_safe --skip-grant-tables &

Ekstrakcija stranica (1)

- format / tip fizičkih redaka
 - **COMPACT**: default MySQL >= 5.0.3
 - **REDUNDANT**: starije inačice
 - identifikacija: *SHOW TABLE STATUS*
 - opcionalno: **constraints_parser**
- **page_parser** - ekstrakcija svih stranica iz grupnog ibdata1:
 - `page_parser -5 -f ibdata1`
 - spremi sve stranice (i ID) po individualnim direktorijima za svaku tablicu

Ekstrakcija stranica (2)

- koji nam raspon stranica treba?
 - *CREATE TABLE innodb_table_monitor (id int) ENGINE=InnoDB;*
 - MySQL error.log ili konzola
 - *INDEX: name PRIMARY, id 0 286, fields 1/11, type 3*
 - bitan je PK: tražimo 0-286 direktorij
 - u slučaju *.ibd datoteka, najčešće ne treba raditi split:
 - *innodb_file_per_table=1*

Spajanje stranica

- spajanje u jednu datoteku:
 - uspješno pročitane stranice potrebno **spojiti** kao pripremu za provjeru konzistencije i ekstrakciju ispravnih redaka:
 - ```
find pages-1246363747/0-286/ -
type f -name '*.page' | sort -n
| xargs cat >
tablica_pages_concat
```

# Definicije / ograničenja (1)

- **create\_defs.pl** - definicije tablica:
  - korak nakon ekstrakcije stranica, za identifikaciju **ispravnih** redaka
  - potreban aktivan (laboratorij) MySQL servis
  - ```
./create_defs.pl --  
host=localhost --user=root --  
password=lozinka --db=baza --  
table=tablica >  
include/table_defs.h
```

Definicije / ograničenja (2)

- tipični primjer SQL definicije tablice:
 - *CREATE TABLE `t1` (`ID` int(11), `NAME` varchar(120), `N_FIELDS` int(10), PRIMARY KEY (`ID`), KEY `NAME`(`NAME`)) ENGINE=InnoDB DEFAULT CHARSET=latin1*
- PK:
 - ID, DB_TRX_ID, DB_ROLL_PTR, NAME, N_FIELDS
- SK:
 - NAME, ID (PK!)

Definicije / ograničenja (3)

```
{ /* int(11) unsigned */      { /* varchar(120) */  
name: "ID",                      name: "NAME",  
type: FT_UINT,                     type: FT_CHAR,  
fixed_length: 4,                   min_length: 0,  
has_limits: TRUE,                  max_length: 120,  
limits: {                          has_limits: TRUE,  
can_be_null: FALSE,                limits: {  
uint_min_val: 0,                  can_be_null: TRUE,  
uint_max_val:                 char_min_len: 0,  
        4294967295ULL                char_max_len: 120,  
},                                char_ascii_only: TRUE  
can_be_null: FALSE                },  
},                                can_be_null: TRUE  
},
```

Definicije / ograničenja (4)

- interni zapisi (nije potrebno mijenjati):
 - **D_TRX_ID** – ID zadnje transakcije koja je uzrokovala taj zapis
 - **DB_ROLL_PTR** – pokazivač na undo record gdje je prethodna verzija retka
 - **DB_ROW_ID** – prvo polje u tablicama bez primarnog ključa

Dohvat ispravnih redaka

- što preciznije definirati ograničenja:
 - type, fixed_length, can_be_null
 - has_limits:
 - uint_min_val, uint_max_val
 - char_min_len, char_max_len, char_ascii_only
- koristi se za **constraints_parser**
 - definicije se statički kompiliraju u alat
 - kod promjene potrebno rekompilirati alat!
 - `./constraints_parser -5 -f tablica_pages_concat`
 - izlaz: ime tablice i retci

Kratki repetitorij

- safe_mysqld, innodb_recovery=4
- format: REDUNDANT ili COMPACT
- raspon ID-jeva stranica same tablice
- kompiliranje
- ekstrakcija: page_parser
- spajanje u jednu datoteku: find + cat
- priprema definicija: create_defs.pl
- rekompiliranje
- constraints_parser

EOF

- hvala na pažnji
- kompleksno područje, malo vremena, previše materijala
- pitanja, diskusija
- vidimo se iduće godine!