



Percona XtraBackup: install, usage, tricks

Vadim Tkachenko
Percona Inc, co-founder, CTO
Vadim@percona.com

Alexey Kopytov
Percona Inc, Principal Software Engineer
Alexey.Kopytov@percona.com

Why do we talk?

- Vadim Tkachenko
 - Design & architecture of XtraBackup
- Alexey Kopytov
 - Lead developer of XtraBackup

This talk online

- PowerPoint
 - <http://bit.ly/XB-2012>
- PDF
 - <http://bit.ly/XB-2012-pdf>

Do I/you need a backup?

Yes. A story from 2009

<http://gnolia.com/>



“Ma.gnolia experienced every web service’s worst **nightmare: data corruption and loss**. For Ma.gnolia, this means that the service is offline and members’ bookmarks are unavailable, both through the website itself and the API. As I evaluate recovery options, I **can’t provide a certain timeline** or prognosis as to when or to what degree Ma.gnolia or your bookmarks will return; only that this process **will take days, not hours.**”

The site was never recovered

Backups do not give you

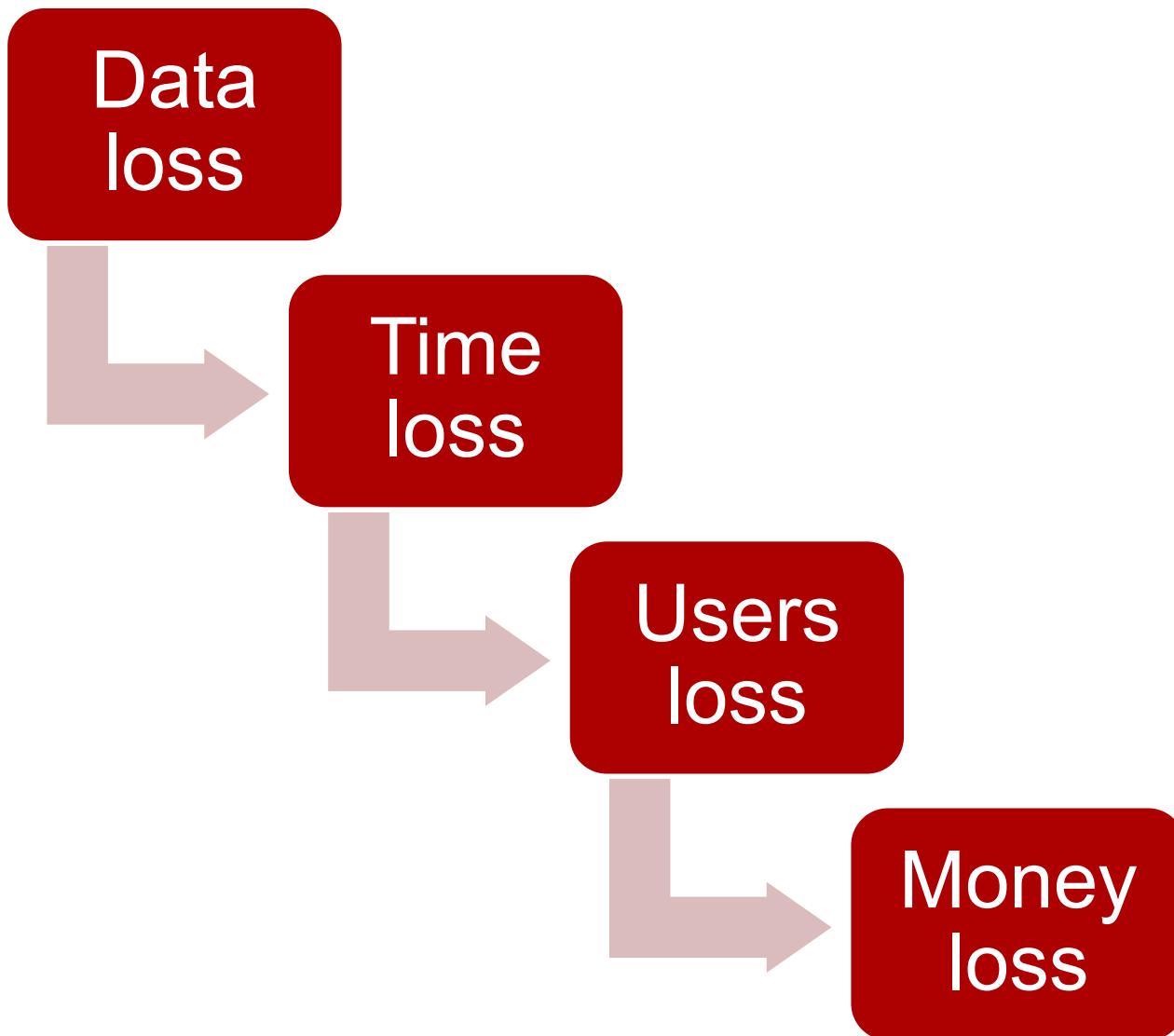
More users

More
money

Increase
productivity

Backups are boring

Not having backup



Recovery time matters

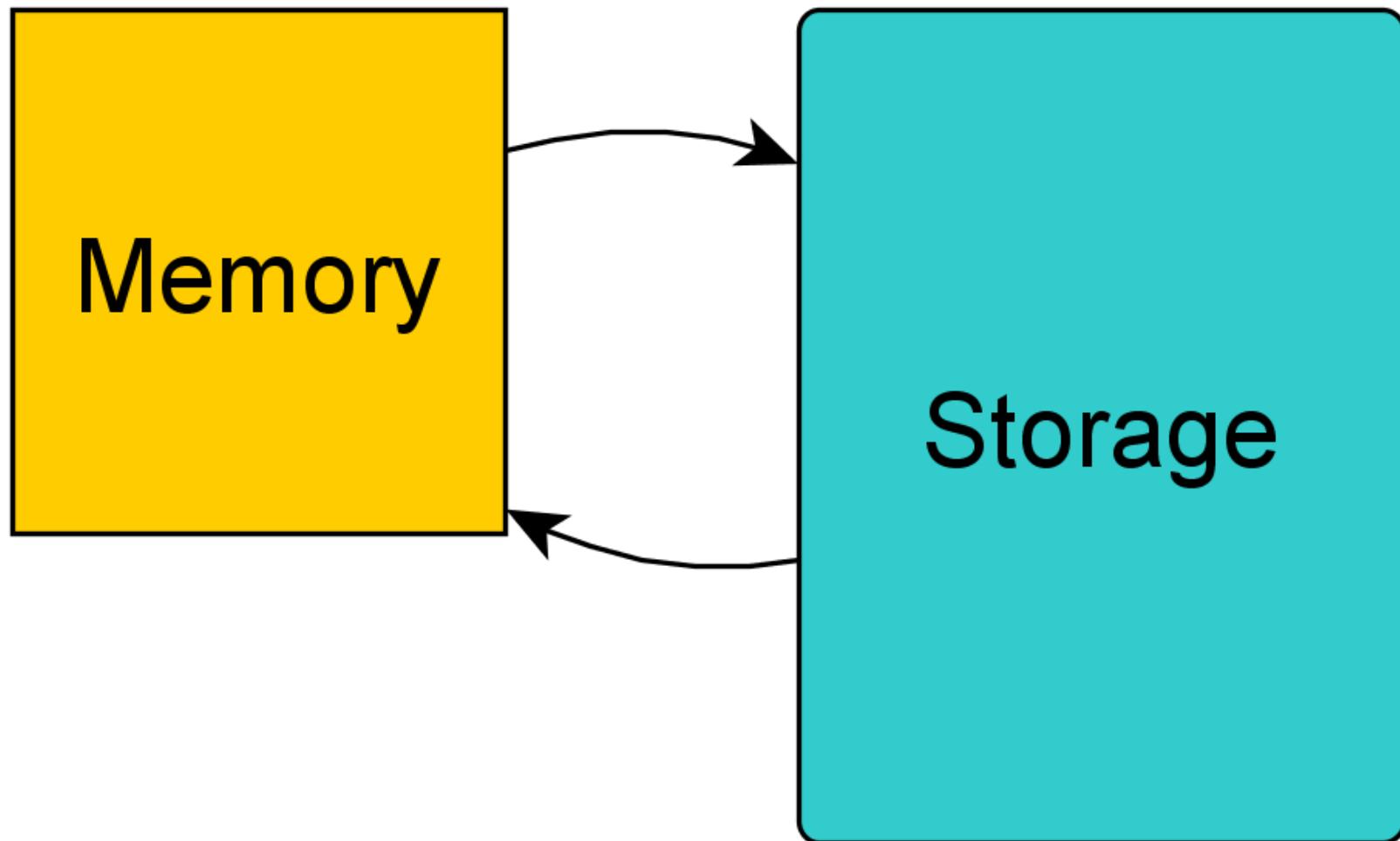
Regularly test your recovery procedure

You may get surprised

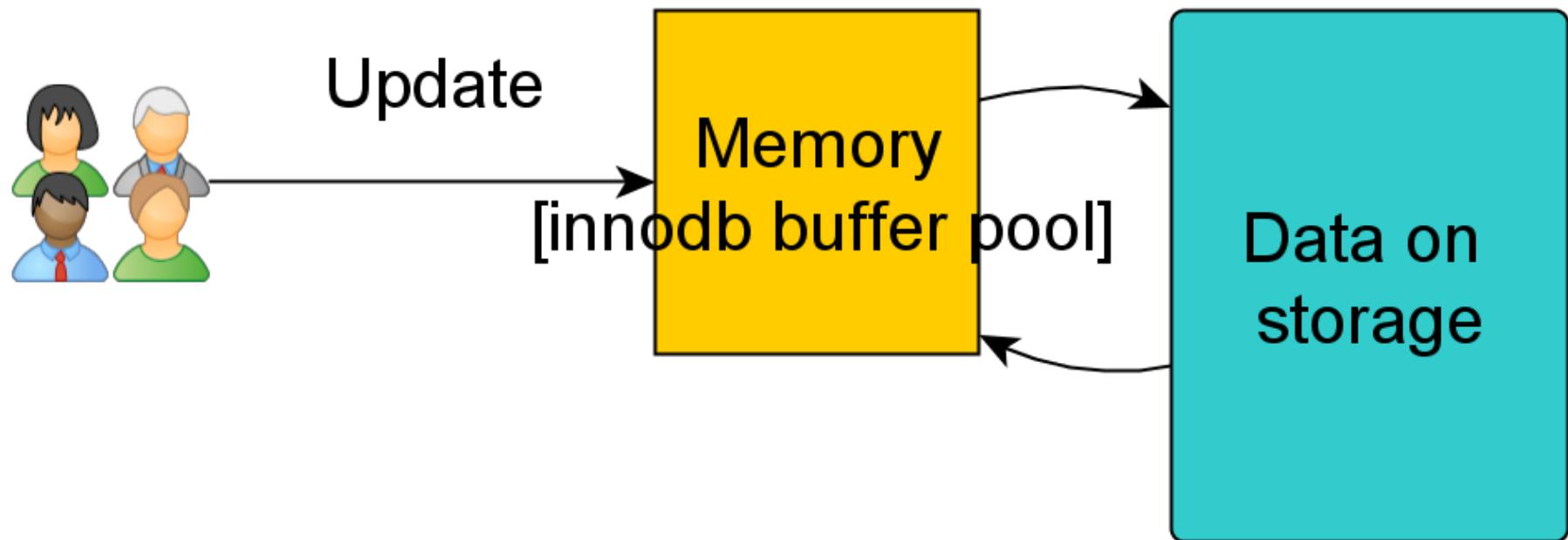
Backup – copy data

Why can't we use just “cp” or “rsync”?

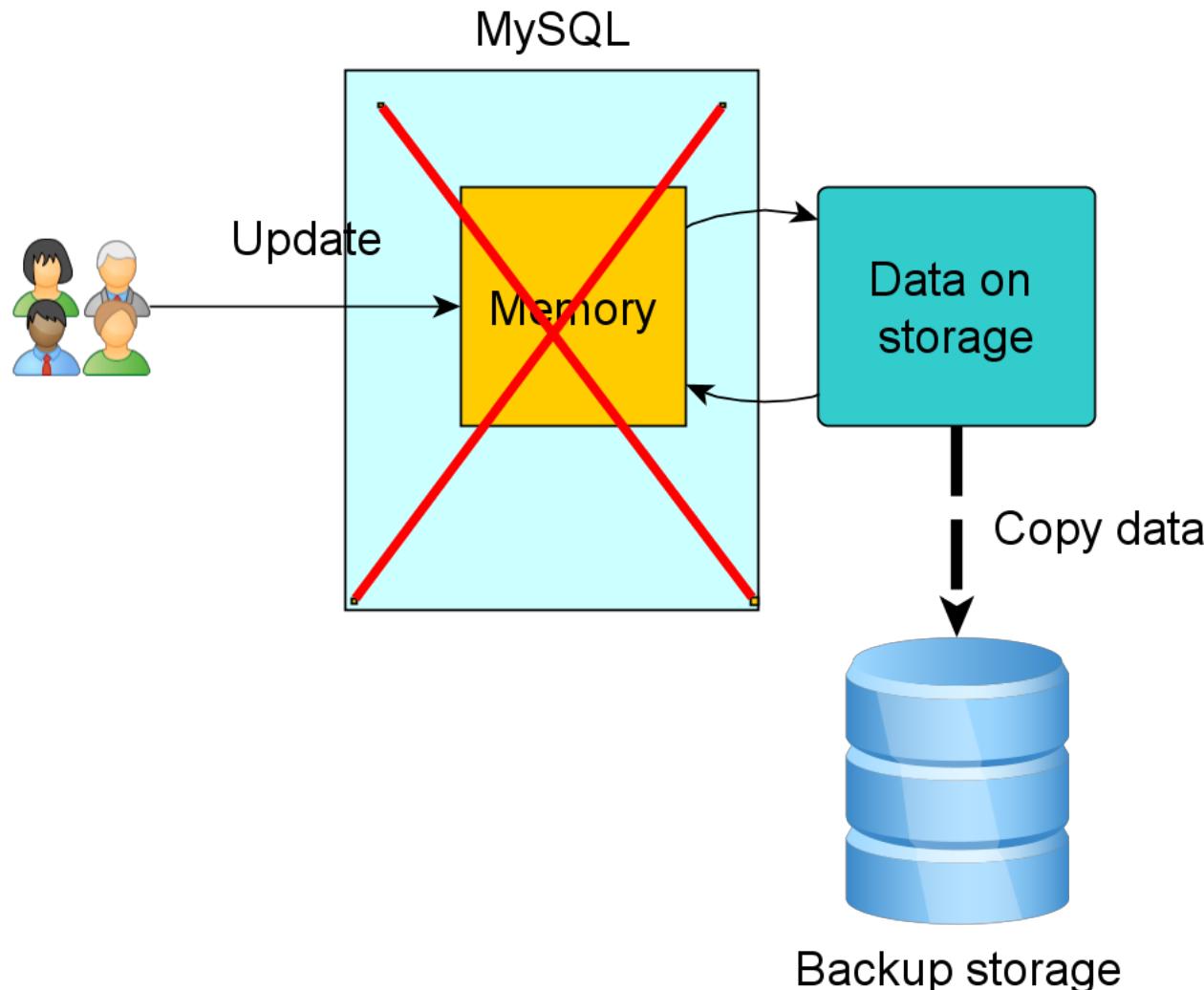
Database



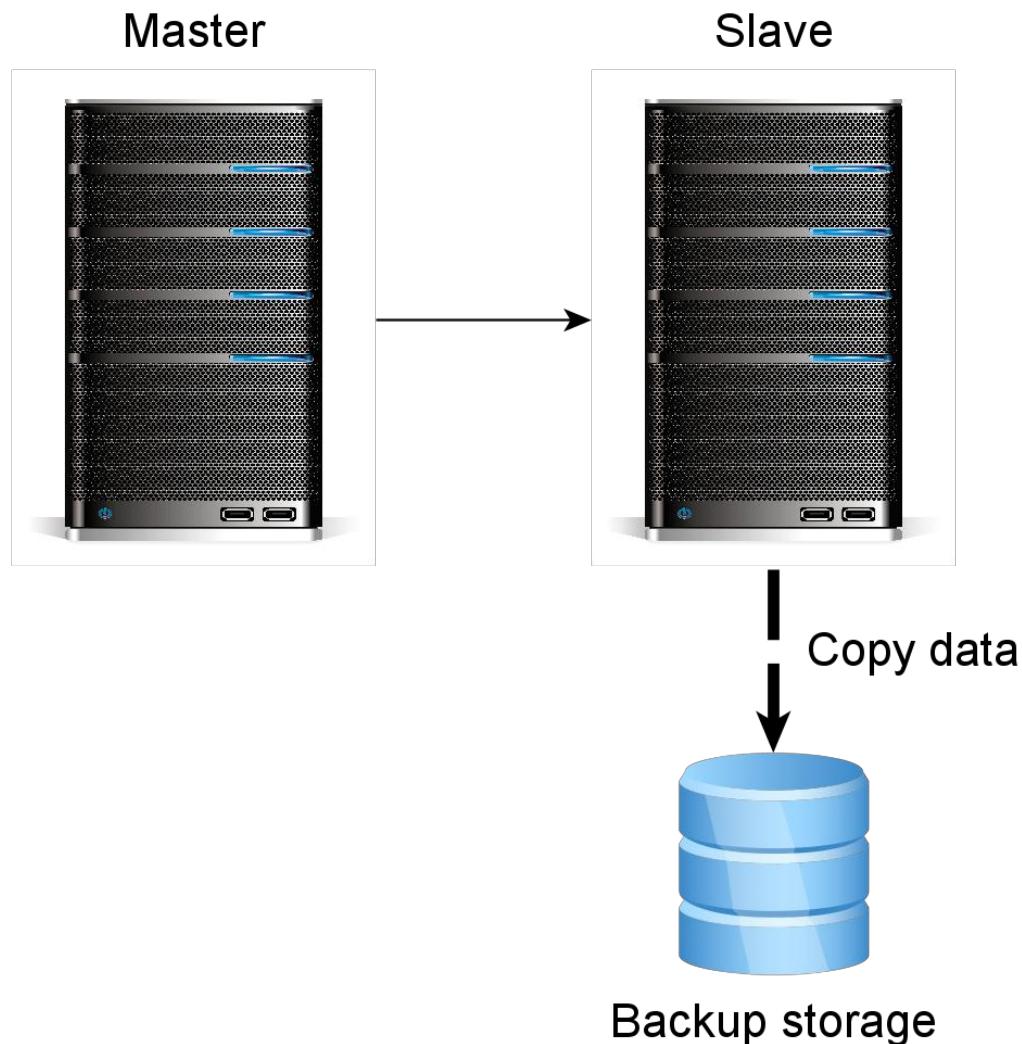
InnoDB updates



Cold copy



Slave backup



Cold slave backup

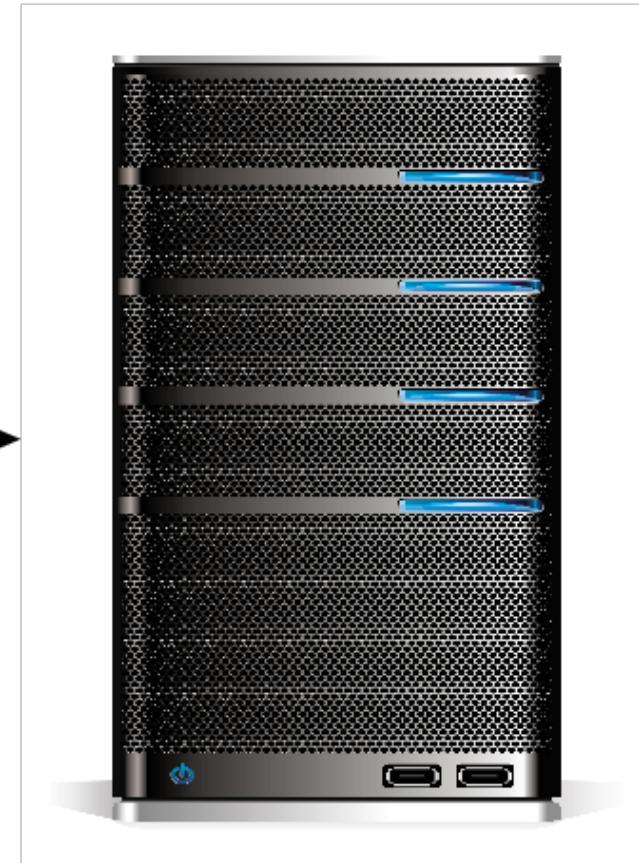


Slave as “backup”

Master



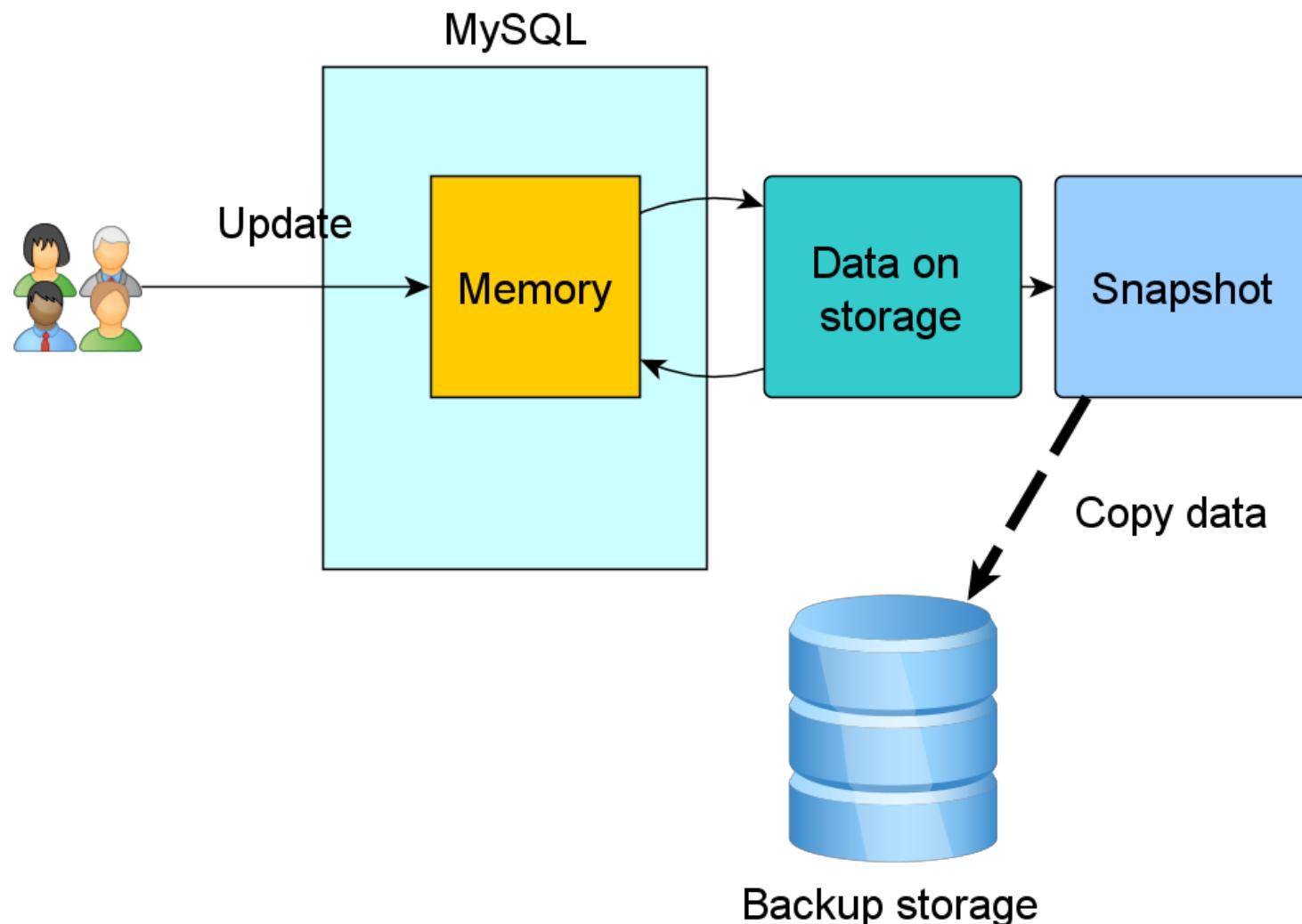
Slave as backup storage



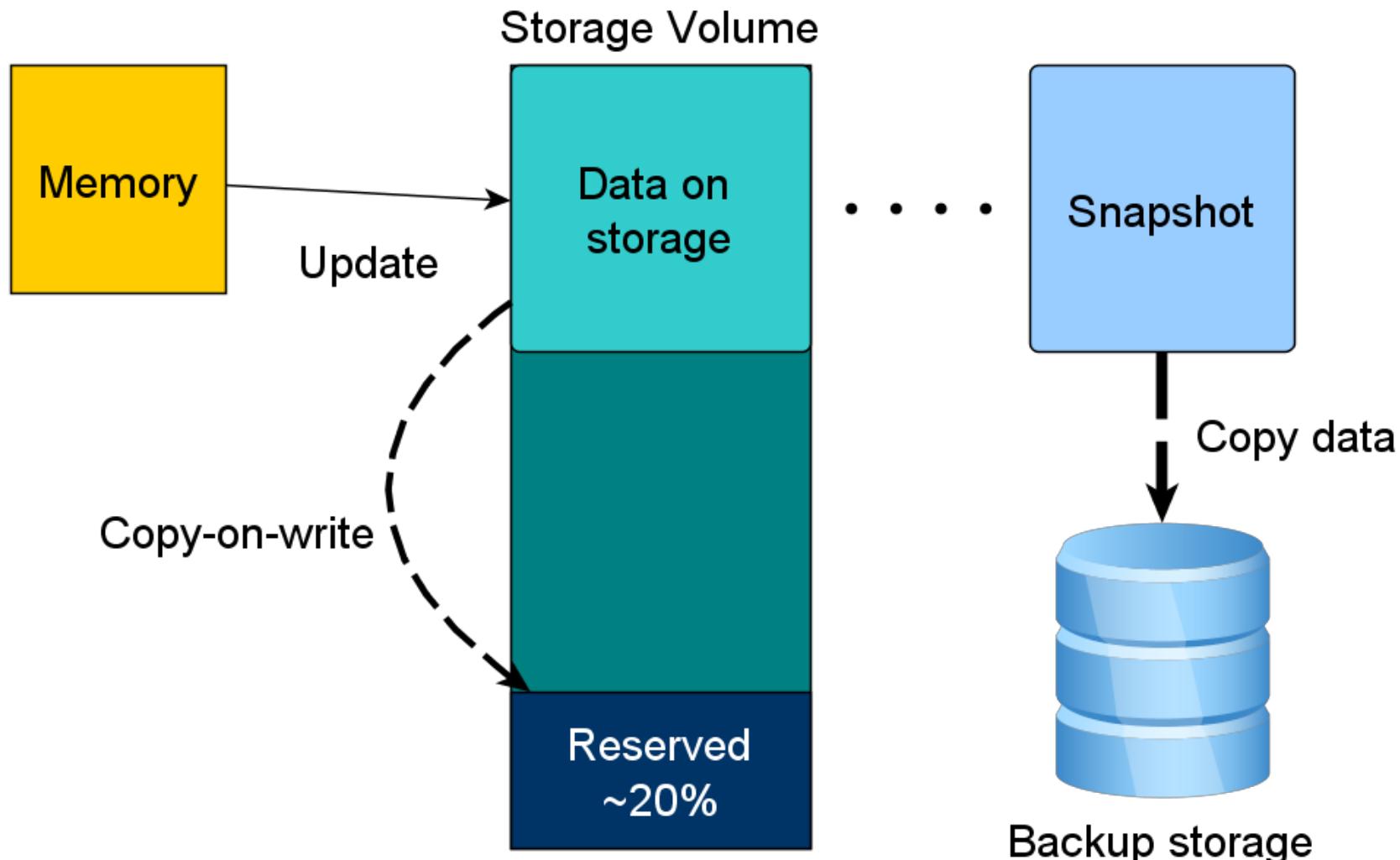
It works till first DROP TABLE or bad UPDATE

Shhhh.... It happens

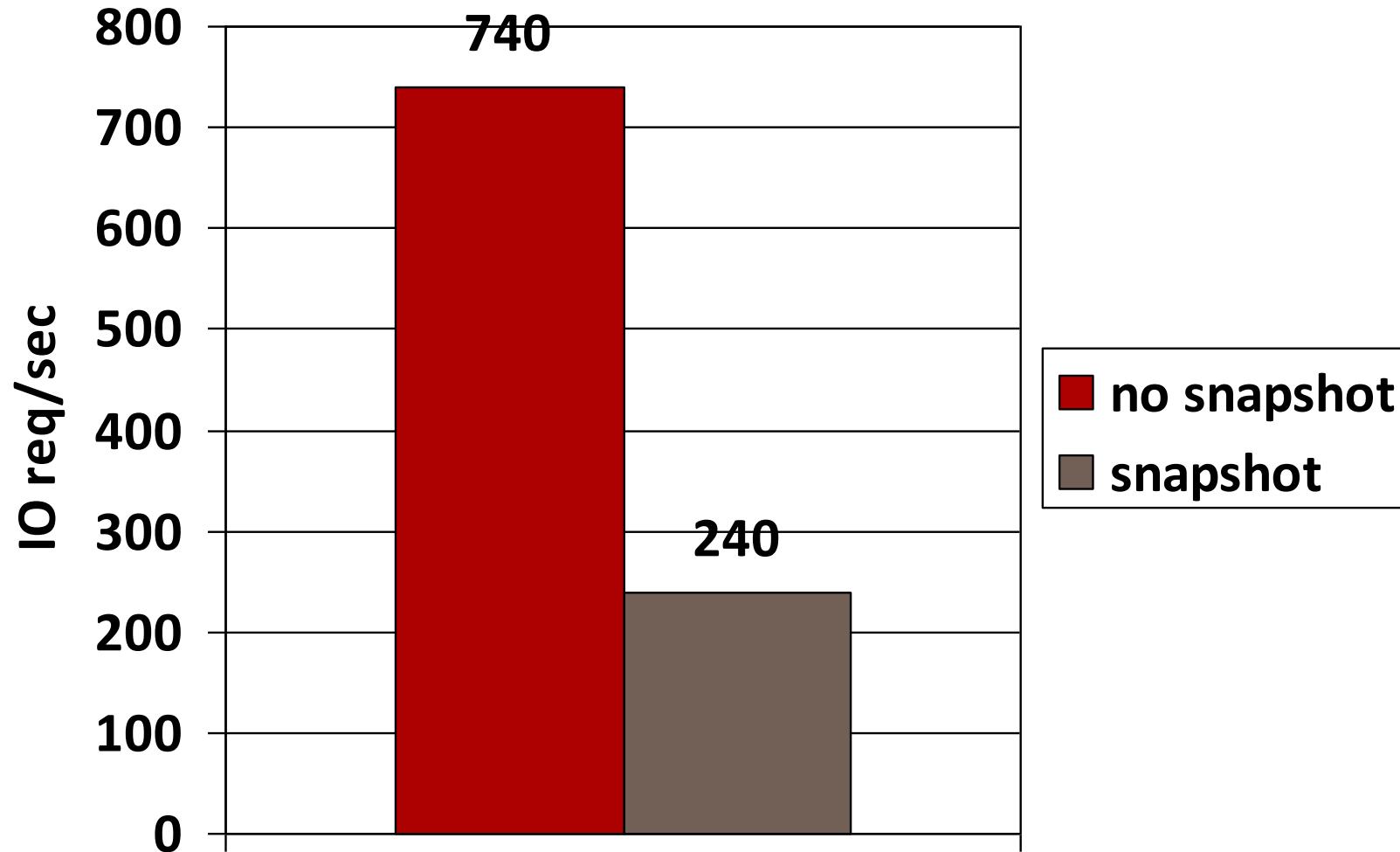
Snapshots



LVM snapshot



LVM snapshot performance



ZFS snapshots

I am told it works fine

ZFS available only

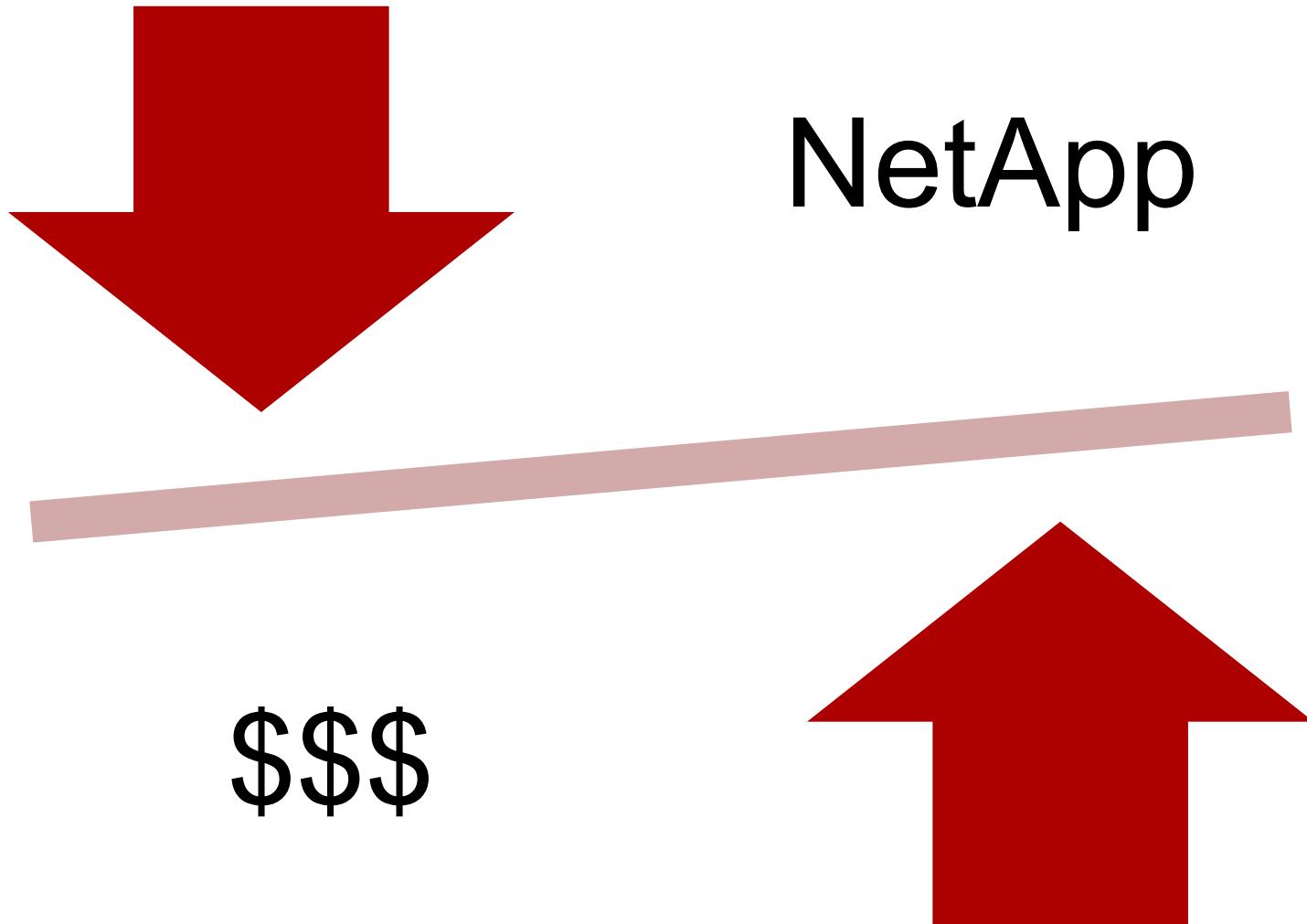
Solaris

FreeBSD

SAN snapshots

\$\$\$

NetApp

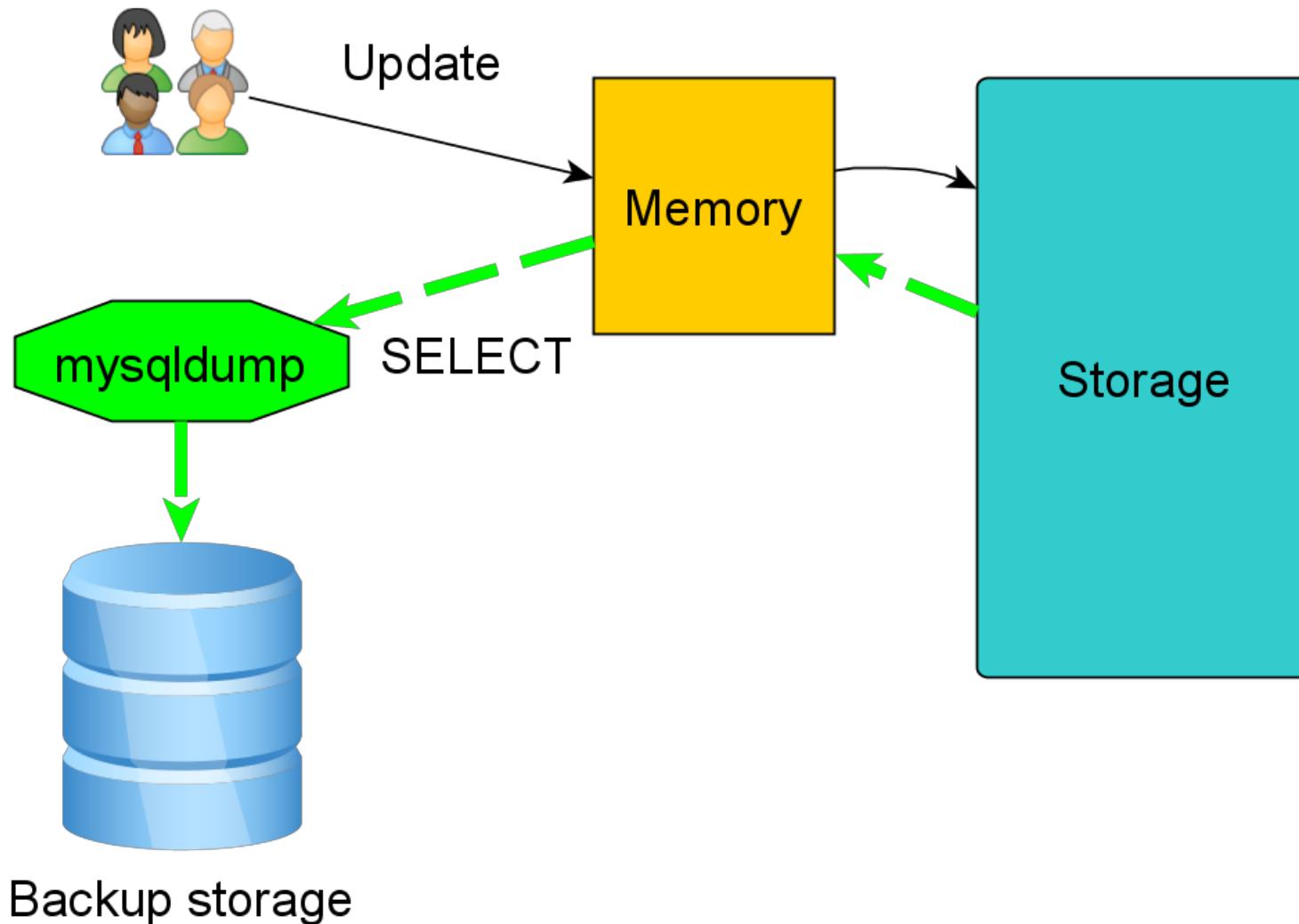


R1Soft

I do not know much about it

Logical backup

mysqldump



“Backup that takes 30 hours
can’t be a daily backup”

Mysqldump benefits

Easy to
use

Per-table

Per-
database

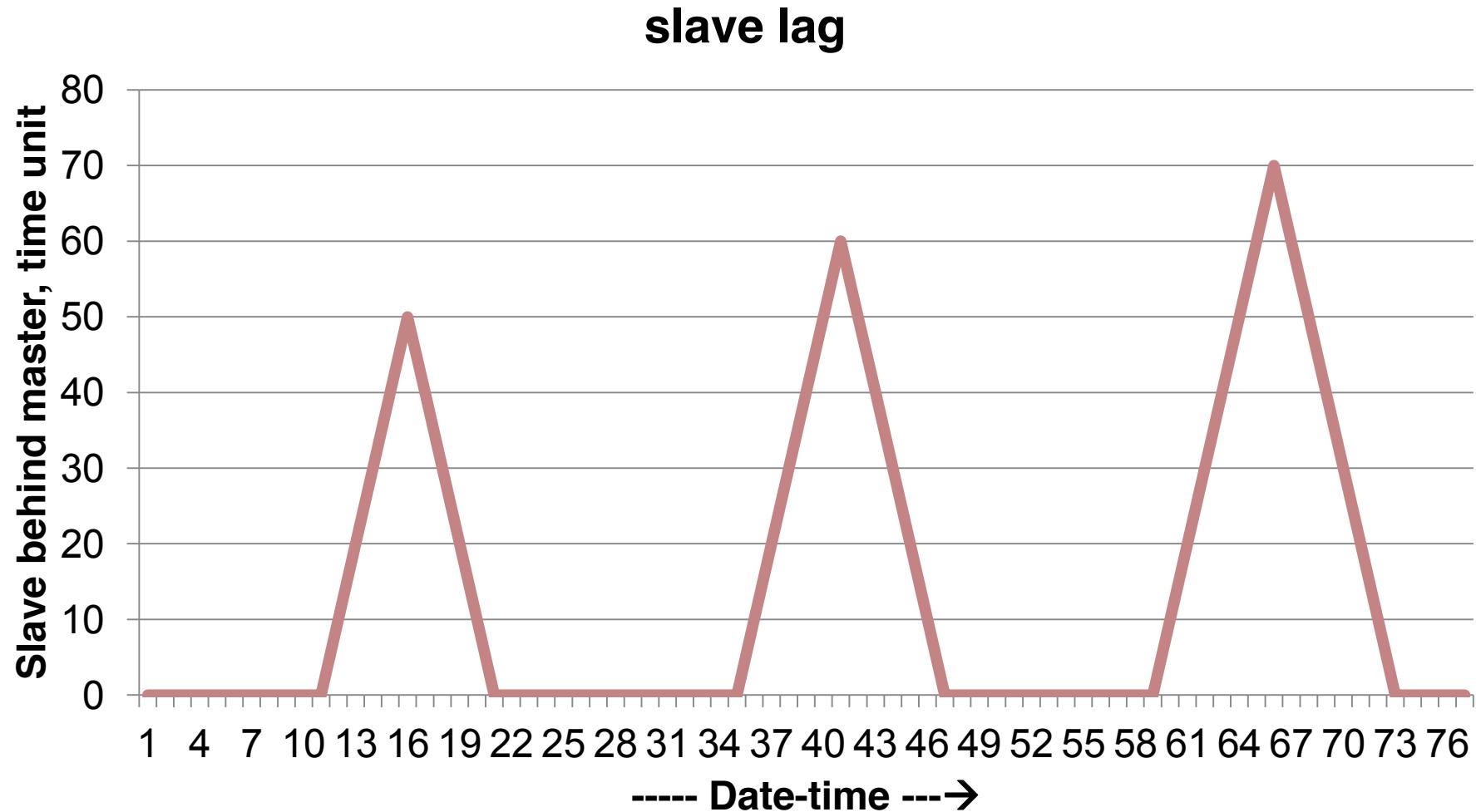
Mysqldump problems

Intrusive

Slow

Single
thread

Slave lag under mysqldump



mydumper

Parallel mysqldump

MySQL Enterprise Backup

MEB benefits

Similar to XtraBackup

MEB drawbacks

Closed
source

Annual
subscription

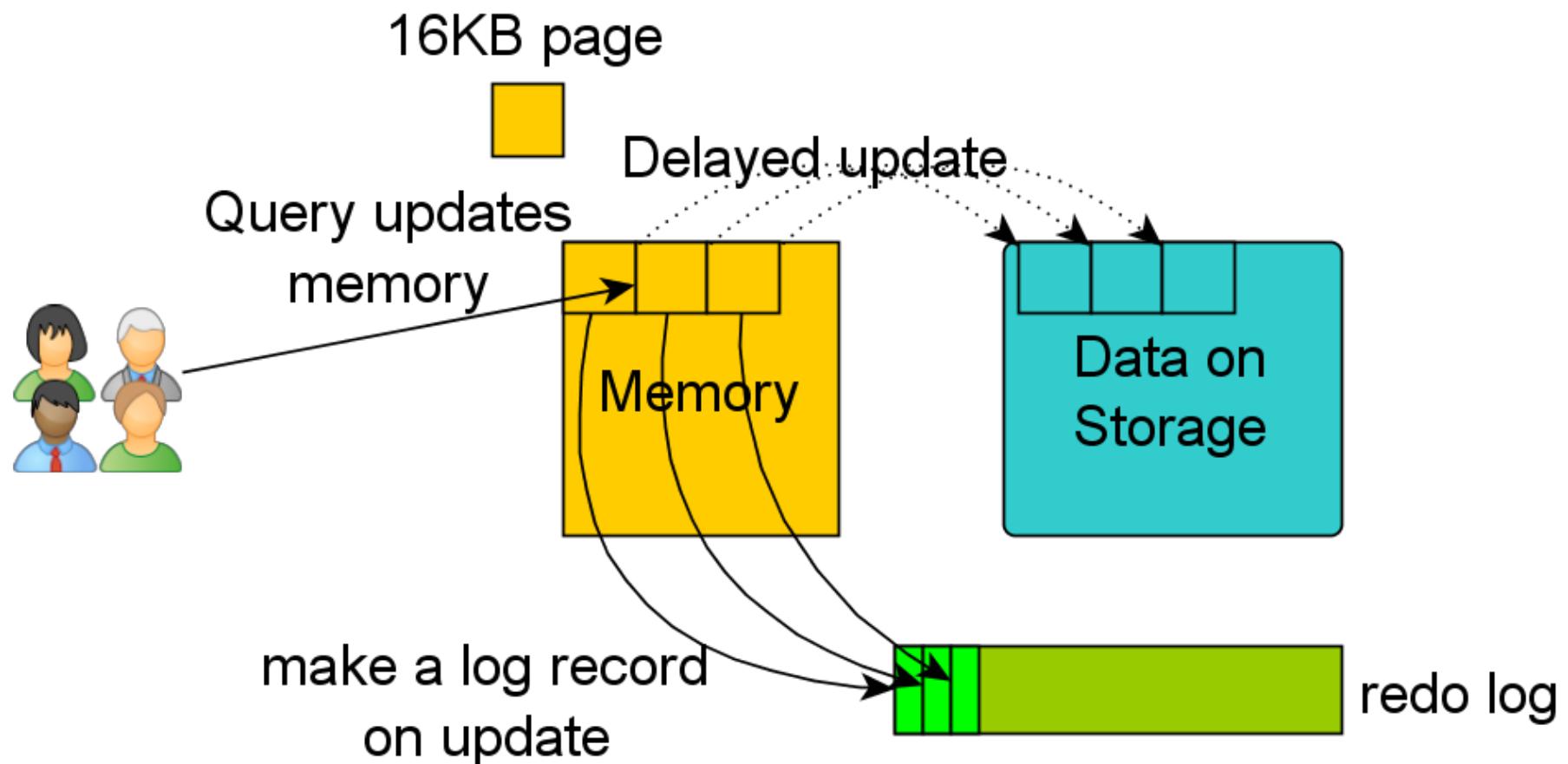
Not true

XtraBackup is “reverse engineering” of MEB

What is Percona XtraBackup

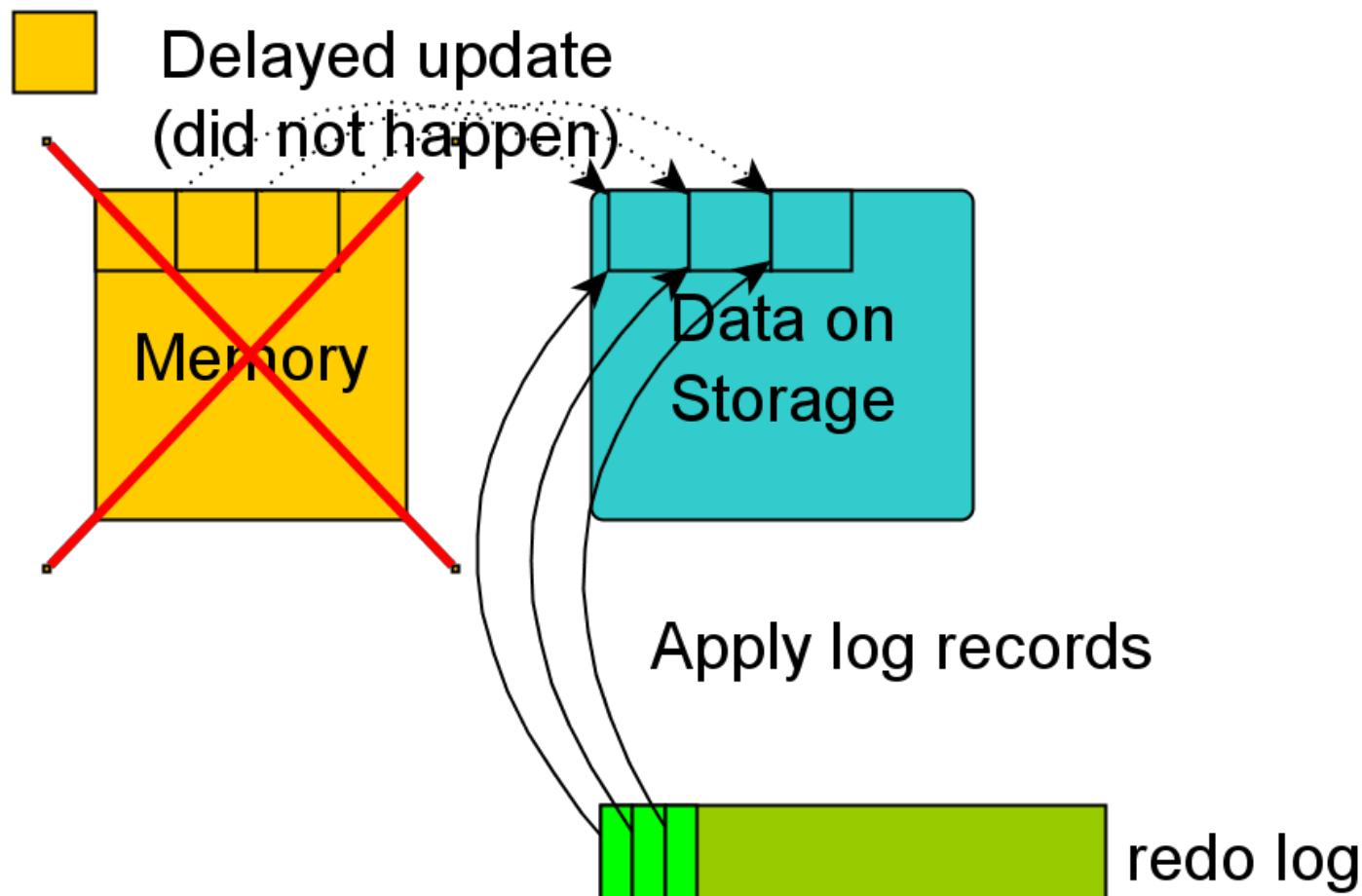
Open-source hot
backup tool for
MySQL-based
servers

InnoDB internals

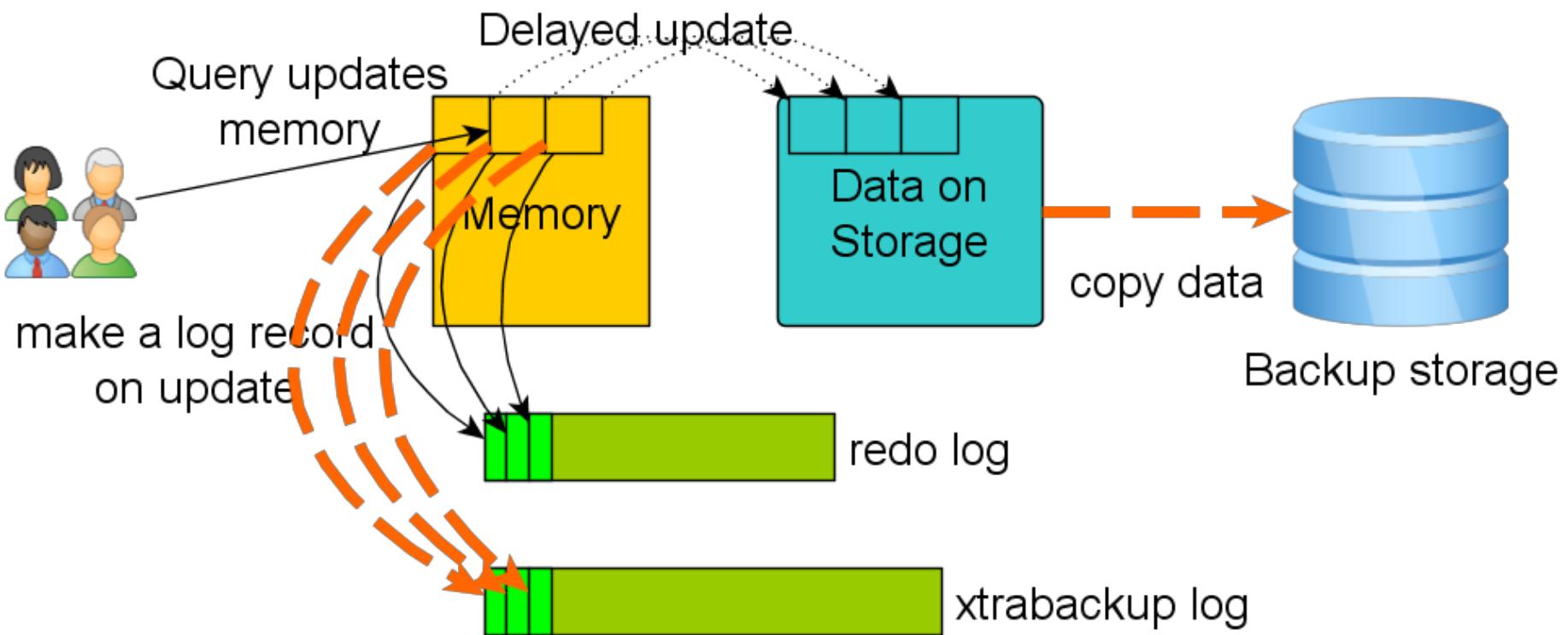


InnoDB recovery

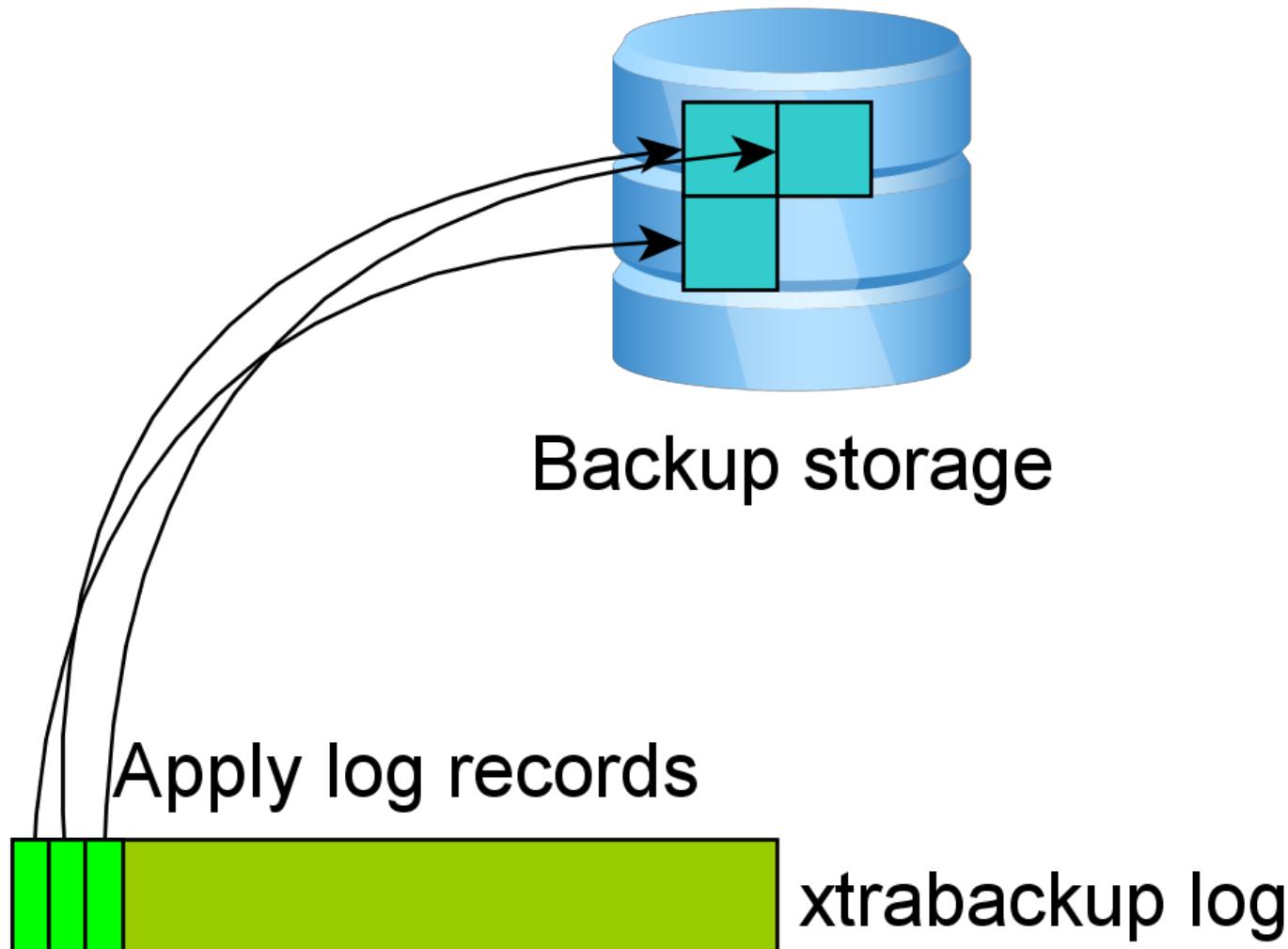
16KB page



Backup idea

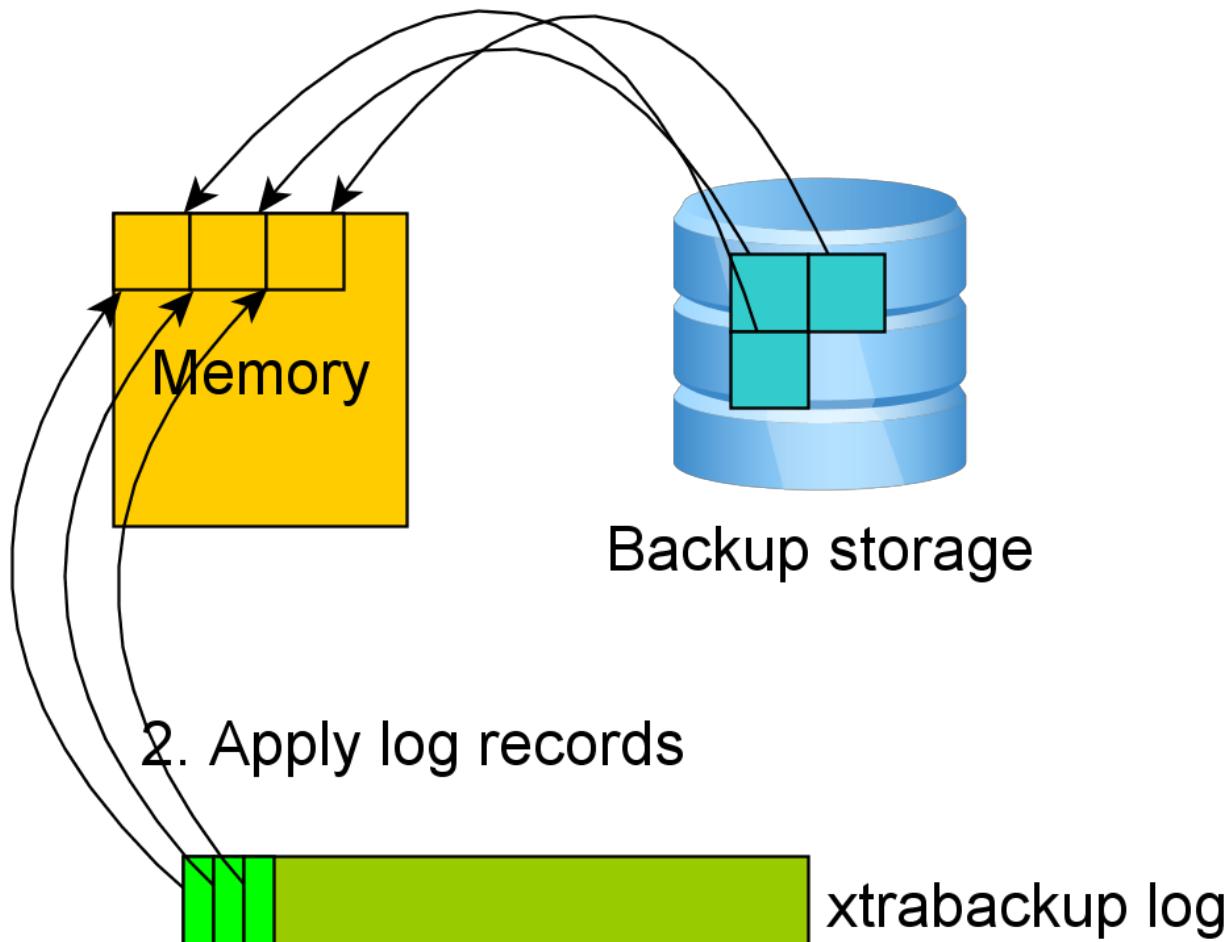


Backup 2nd stage



Recovery uses memory

1. Read data into memory



Backup logic is identical to InnoDB recovery logic

Backup logic is a part of **InnoDB** source code

No needs for “reverse engineering”

XtraBackup source code structure

Patches to InnoDB

C code

Innodbupex script

Streaming utilities

bzr branch lp:percona-xtrabackup

<https://launchpad.net/percona-xtrabackup>

Build tools

- build/ directory
- build.sh
 - build.sh xtradb55
- build-rpm.sh
 - xtrabackup.spec
- build-dpkg.sh

Binary packages

Yum repository

- `rpm -Uhv http://www.percona.com/downloads/percona-release/percona-release-0.0-1.x86_64.rpm`
- `yum install percona-xtrabackup`

Apt repository

- deb `http://repo.percona.com/apt squeeze main`
 - Replace “squeeze” by your favorite name
- deb-src `http://repo.percona.com/apt squeeze main`
- apt-get update

Binary packages

- <http://www.percona.com/downloads/XtraBackup/XtraBackup-2.0.0/>
 - Binary tar.gz
 - Centos based, may not work on SUSE
 - deb
 - RPM
 - RedHat based
 - source

Supported MySQL versions

- MySQL 5.0
- MySQL 5.1 + built-in InnoDB
- MySQL 5.1 + InnoDB-plugin
- MySQL 5.5
- Percona Server 5.0
- Percona Server 5.1
- Percona Server 5.5
- Percona XtraDB Cluster 5.5

“Should work” products

- MariaDB
 - May need small fix
- Drizzle
 - Shipped with its own xtrabackup

Supported engines

- InnoDB/XtraDB
 - Hot backup
- MyISAM
 - With read-lock
- Archive, CSV
 - With read-lock
- Your favorite exotic engine
 - May work if supports FLUSH TABLES WITH READ LOCK

Supported platforms

- Linux
 - RedHat 5; RedHat 6
 - CentOS, Oracle Linux
 - Debian 6
 - Ubuntu LTS
 - Any other via compiling source code

Solaris

- Binaries will be soon

Mac OS X

- Binaries will be soon

Windows

- Experimental releases
- Active Perl

FreeBSD

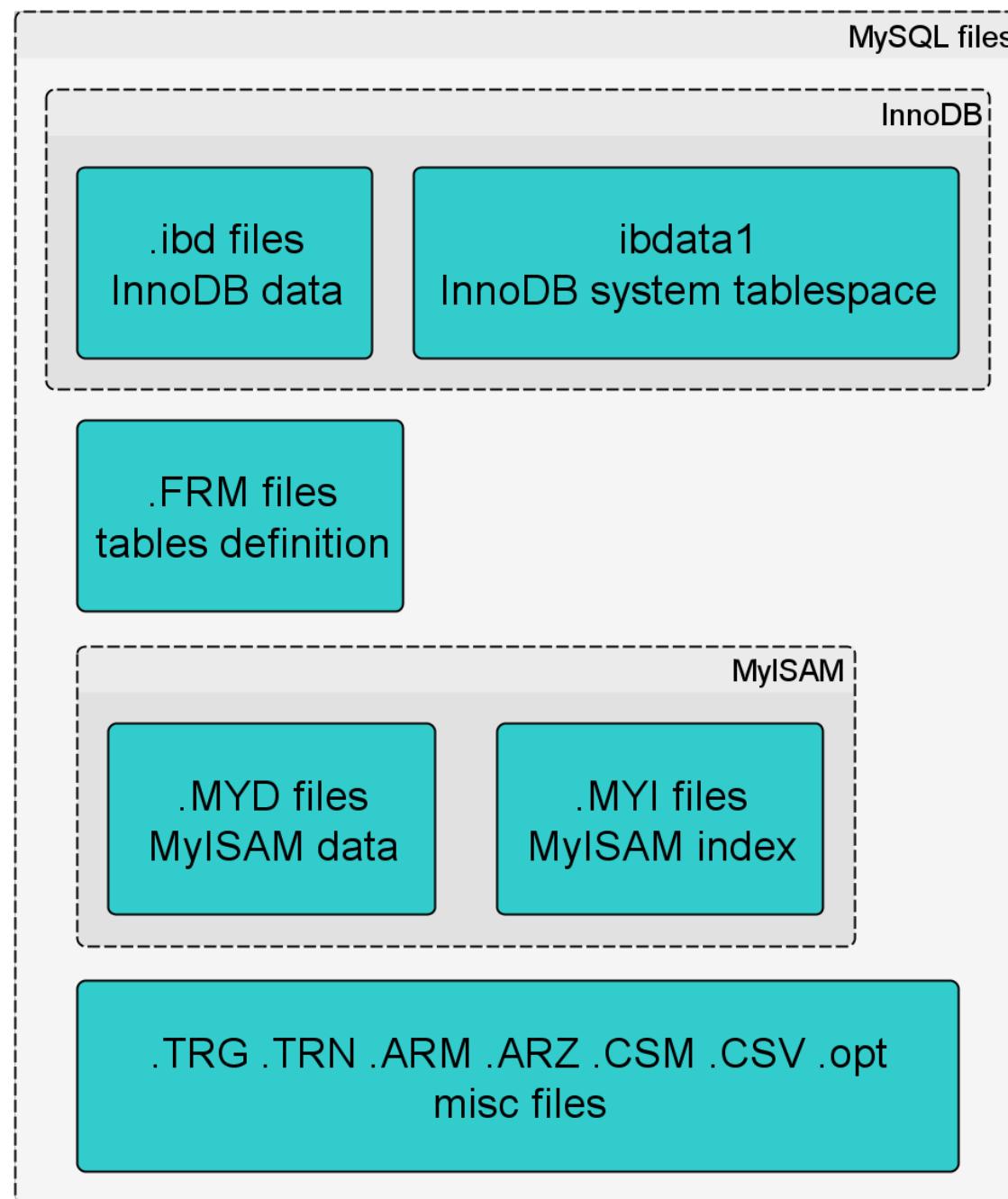
- Compile source code

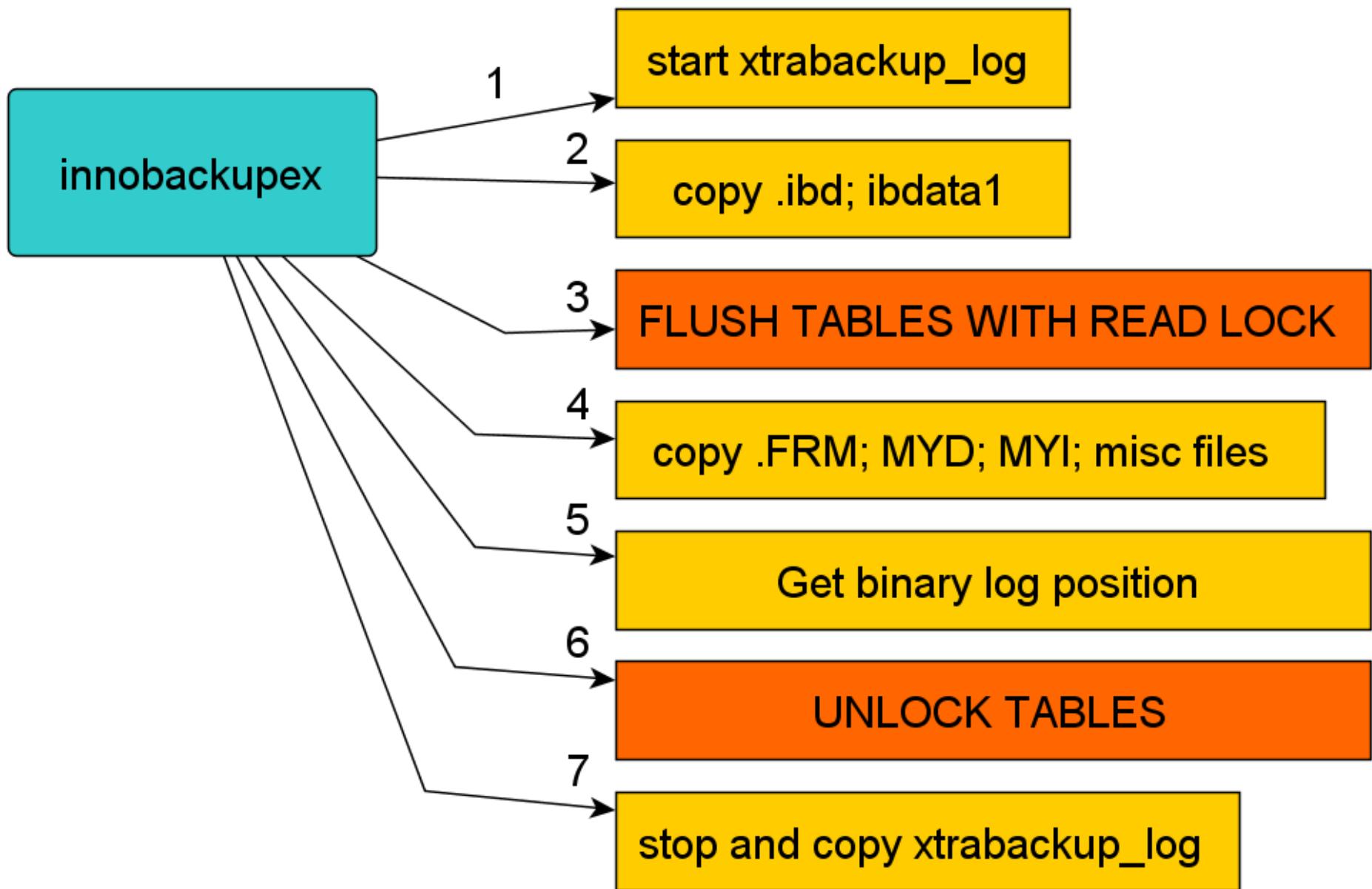
Binaries structure

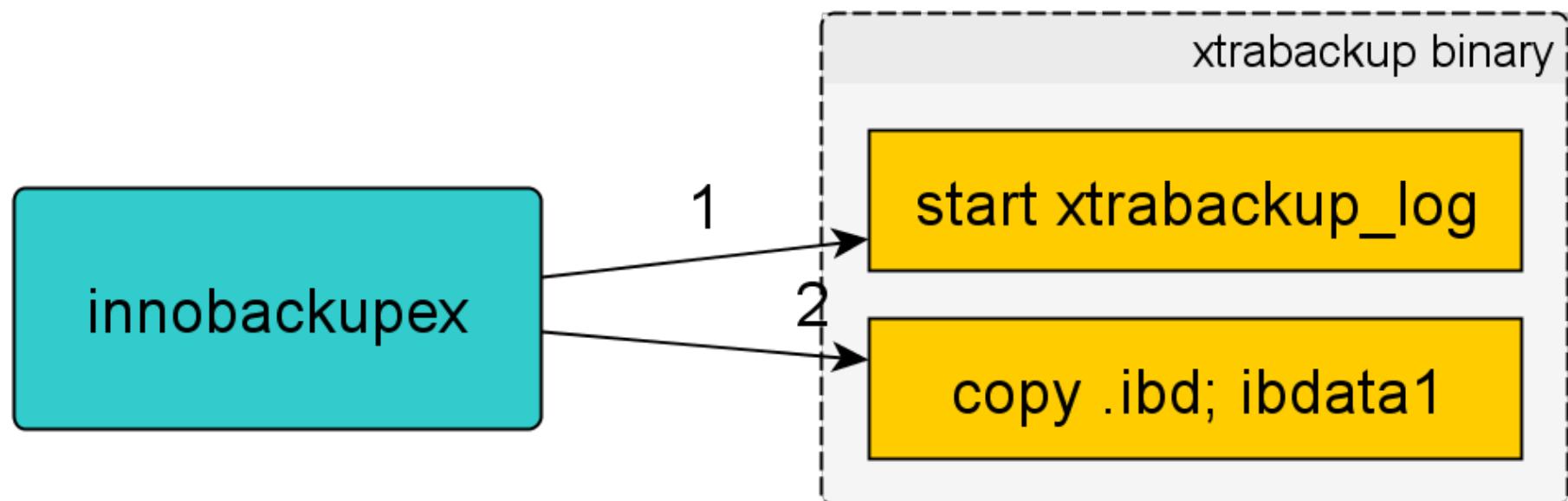
- **innobackupex**
 - Perl script
- **xtrabackup**
 - Percona Server 5.1; MySQL 5.1 + InnoDB-plugin
- **xtrabackup_51**
 - MySQL 5.0; Percona Server 5.0; MySQL 5.1+builtin InnoDB
- **xtrabackup_55**
 - MySQL 5.5; Percona Server 5.5
- **tar4ibd**
 - Only in XtraBackup-1.6; built-in in 2.0
- **xbstream**
 - Only in 2.0

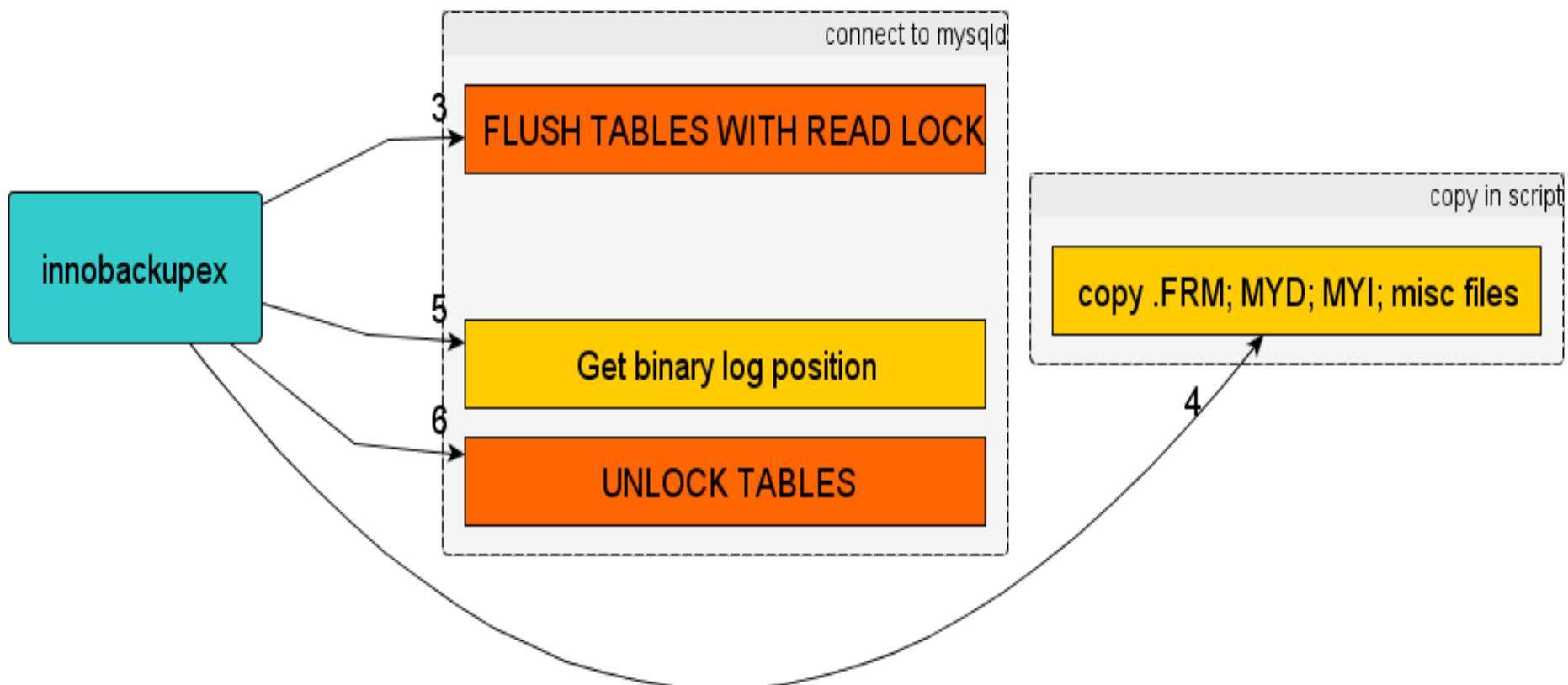


Data on Storage









Basic command

- `innobackupex /data/backup/mysql-data`

Innodbupex connects to mysqld

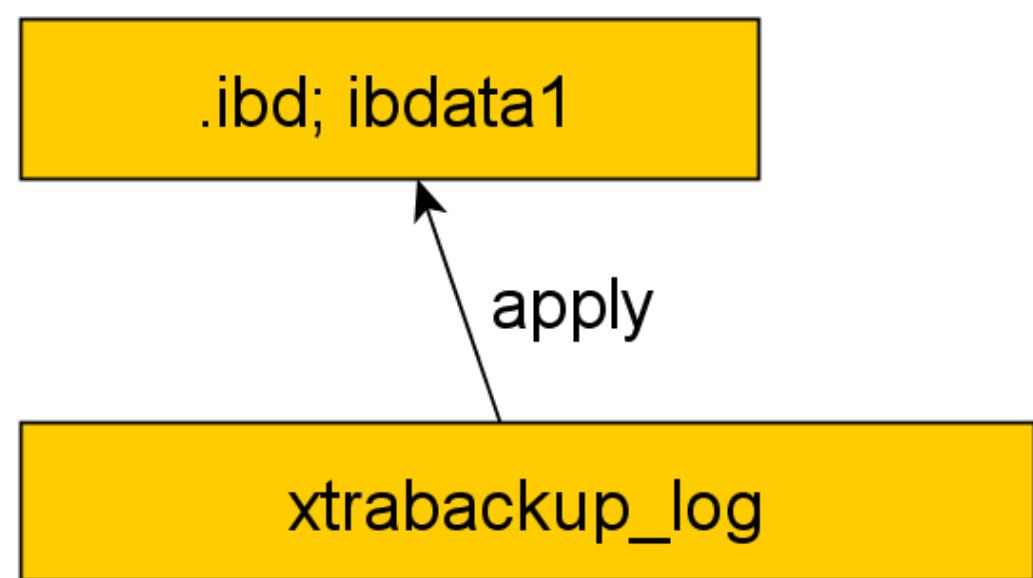
- `--defaults-file`
 - Datadir (default assumption /var/lib/mysql) !
- `--user`
- `--password`
- `--host`
 - If you need to connected via specific IP
- `--socket`

Innobackupex basic output

File xtrabackup-out1

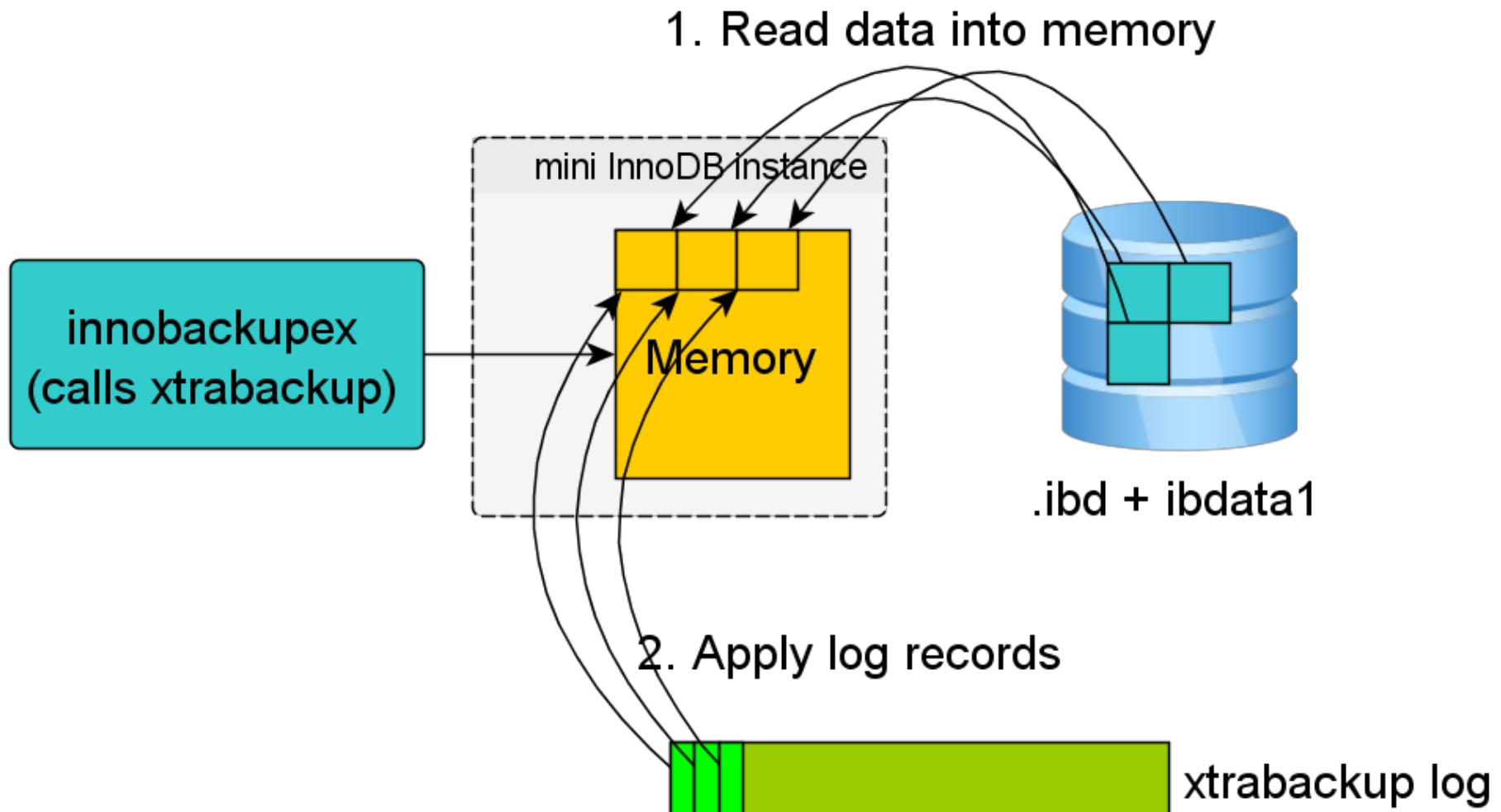
Stage 2

innobackupex
(calls xtrabackup)



apply command

- `innobackupex --apply-log /data/backup/mysql-data`



Use memory

- `innobackupex --apply-log --use-memory=10G /data/backup/mysql-data`
- `use-memory == innodb_buffer_pool_size` for `mysqld`

Innodbupex apply-log output

File xtrabackup-apply-out

Apply length: 8 minutes 25 seconds

Innodbupex apply-log –use-memory output

File xtrabackup-apply-use-memory-out

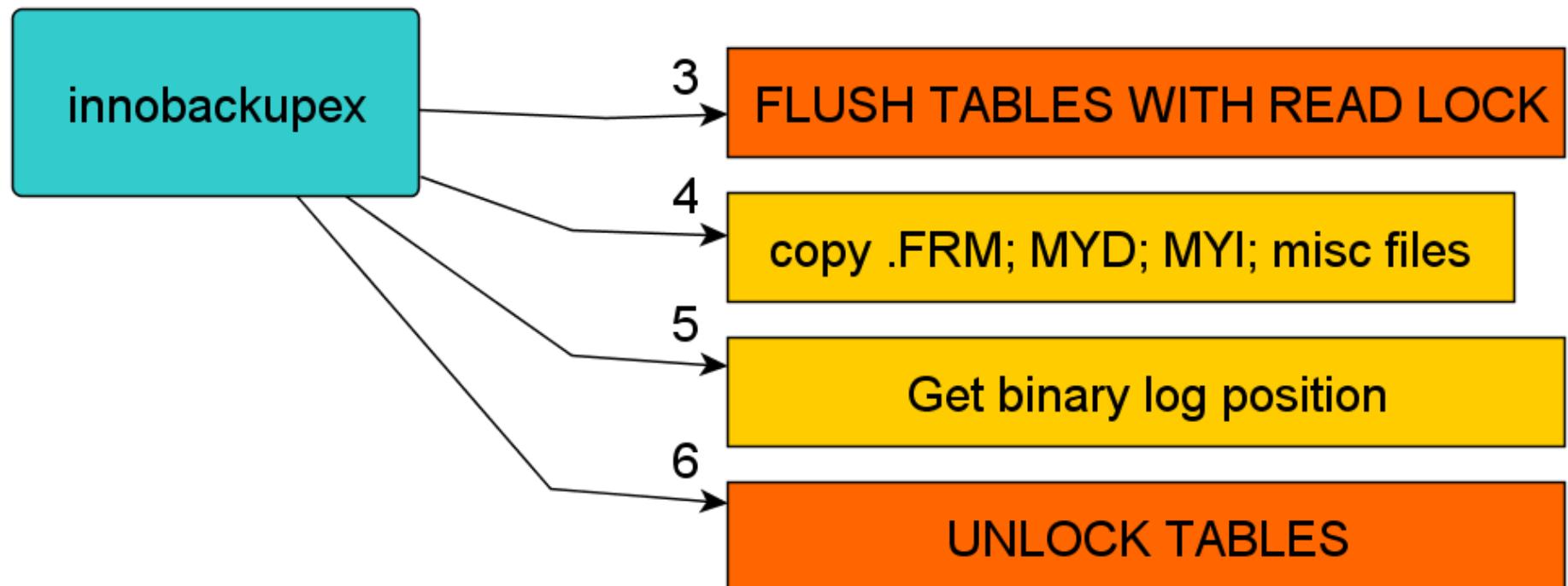
Apply length: 3 minutes 36 seconds

Why xtrabackup –prepare 2 times?

- Second time to create innodb log files
- Just for convenience

/data/backup/mysql-
data is ready for usage

FLUSH TABLES WITH READ LOCK



FTWRL

- set the **global read lock** - after this step, insert/update/delete/replace/alter statements cannot run
- **close open tables** - this step will block until all statements started previously have stopped
- set a flag to **block commits**

Why? Consider:

- Copy table1.frm
- Copy table2.frm
- Copy table3.frm
- Copy table4.frm
 - Meantine ALTER TABLE table1 started
- Copy table5.frm
- Copy table6.frm
- Copy table7.frm

With FTWRL

- FLUSH TABLES WITH READ LOCK
- Copy table1.frm
- Copy table2.frm
- Copy table3.frm
- Copy table4.frm
 - ALTER TABLE table1 ---→ **LOCKED**
- Copy table5.frm
- Copy table6.frm
- Copy table7.frm
- UNLOCK TABLES

The same with MyISAM

Not transactional engine. No redo logs

FTWRL problems

Problem 1

- **close open tables** - this step will block until all statements started previously have stopped
- **Long running SELECT** will block FTWRL, which will block other statements – **Queries piles up and lock down server**

Pileup

- SELECT 1mIn rows FROM BIG_TABLE
 - Blocks: FLUSH TABLES WITH READ LOCK
 - Blocks all following queries:
 - SELECT ...
 - UPDATE ...
 - SELECT ...

Solution: shoot long queries

Problem 2

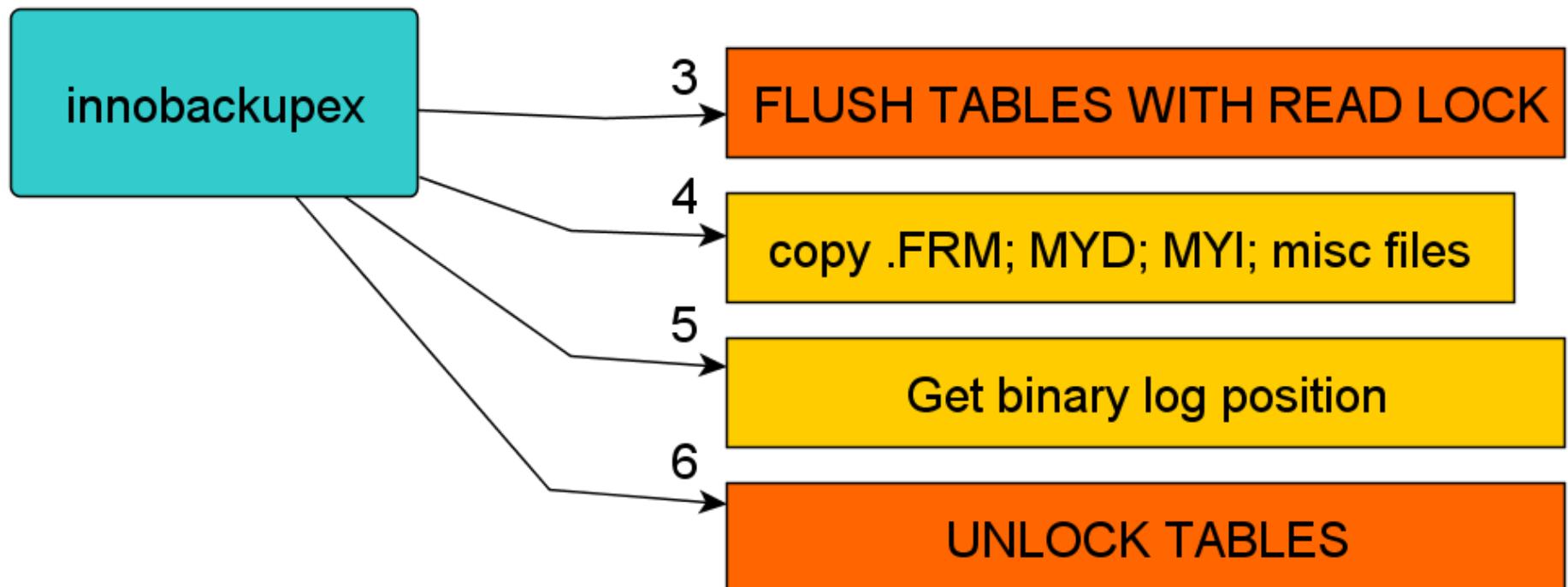
- **Big MyISAM tables**
- **Block write queries** on time copying MyISAM
- Example: you created 100GB MyISAM table and forgot about

--no-lock

--no-lock

Do not execute FLUSH TABLES WITH READ LOCK

To remind:



You take responsibility

- There is NO ALTER TABLE statements
- There is NO INSERT/UPDATE/DELETE to MyISAM tables
- What about Binary log position?
 - We will cover it

Directory after backup

- MySQL files +
- `xtrabackup_binlog_info`
- `xtrabackup_binlog_pos_innodb`
 - Only after `-apply-log`
- `xtrabackup_slave_info`
 - When `-slave-info` is used
- `xtrabackup_checkpoints`
- `xtrabackup_logfile`
- `xtrabackup_binary`
- `backup-my.cnf`

xtrabackup_binlog_info

- Master binary log position
- Result of SHOW MASTER STATUS

```
binlog.000001    68212201
```

xtrabackup_checkpoints

- Information about InnoDB checkpoints LSN, useful for incremental backup

xtrabackup_logfile

- Can be very big
- The bigger this file – the longer –apply-log process

xtrabackup_binary

- Exact xtrabackup binary used for backup
- This is used for –apply-log

backup-my.cnf

- NOT BACKUP of my.cnf
- Information to start mini-InnoDB instance during –apply-info
- E.g.

```
[mysqld]
datadir=/data/bench/back/test/2012-04-06_13-09-42
innodb_data_home_dir=/data/bench/back/test/2012-04-06_13-09-42
innodb_log_files_in_group=2
innodb_log_file_size=1992294400
innodb_fast_checksum=0
innodb_page_size=16384
innodb_log_block_size=4096
```

Summary

Backups are important

Recovery time is important

Backup methods for MySQL

Percona XtraBackup overview

XtraBackup basic usage

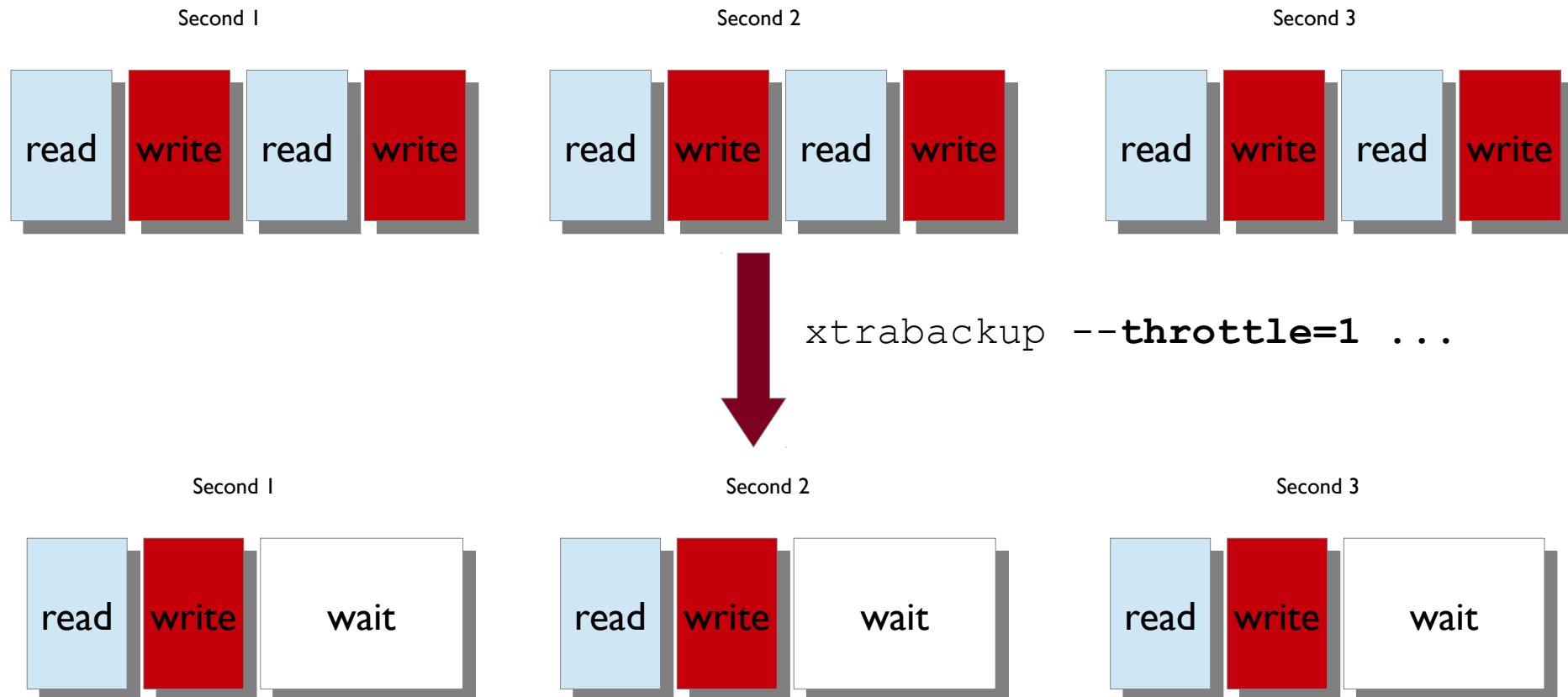
Agenda

- Minimizing footprint
 - I/O throttling
 - FS cache optimization
- Parallel file copying
- Restoring individual tables
- Partial backups
 - individual partitions backup

Minimizing footprint: I/O throttling

--throttle=N

Limit the number of I/O operations per second in 1 MB units



Minimizing footprint: I/O throttling

Limitation:

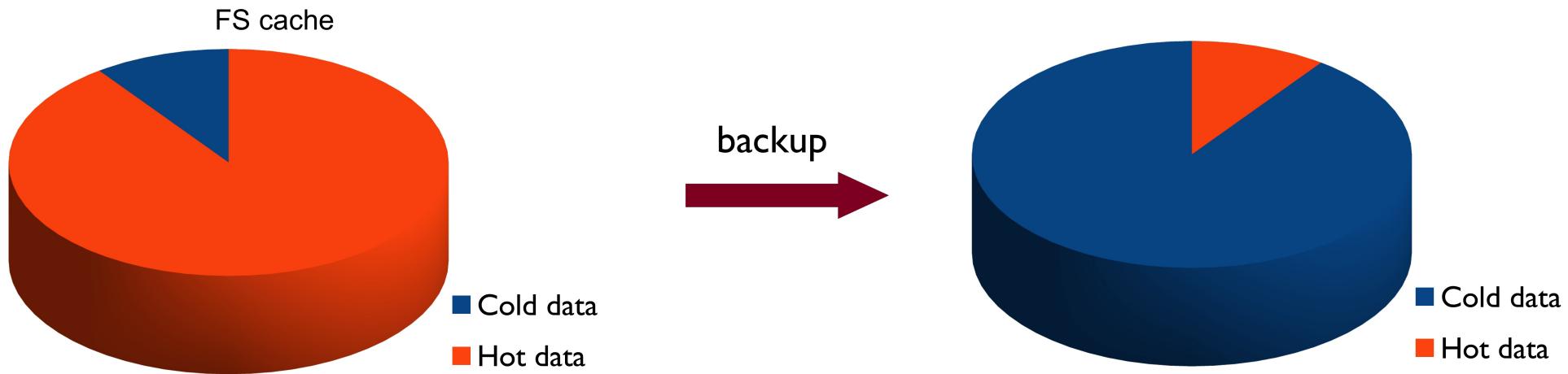
--throttle only has effect for InnoDB tables, other files are copied by the **innobackupex** script with `cp` or `rsync`.

Use streaming backups and the **pv** utility as a workaround:

```
$ innobackupex --stream=tar /tmp | pv -q -L1m | tar -xf - -C /data/backup
```

- **pv** utility
 - monitor the progress of data through a pipe
 - available in most Linux distributions
 - `-L` to limit the transfer to a specified rate

Minimizing footprint: FS cache optimizations

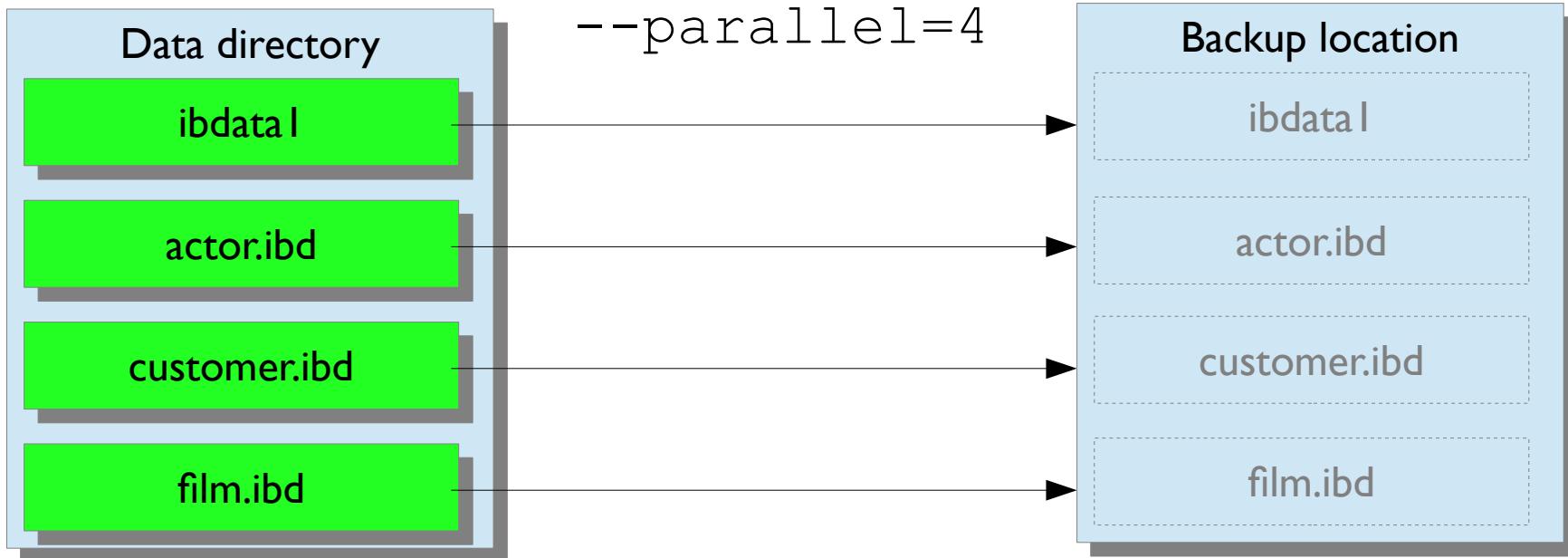


- FS cache may get filled with cold data
- especially bad for MyISAM which relies on OS to cache data files
- causes VM pressure and swapping

Minimizing footprint: FS cache optimizations

- **XtraBackup on Linux:**
 - `posix_fadvise(POSIX_FADV_DONTNEED)`
 - hints the kernel the application will not need the specified bytes again
 - works automatically, no option to enable
- Didn't really work in XtraBackup 1.6, fixed in 2.0

Parallel file copying



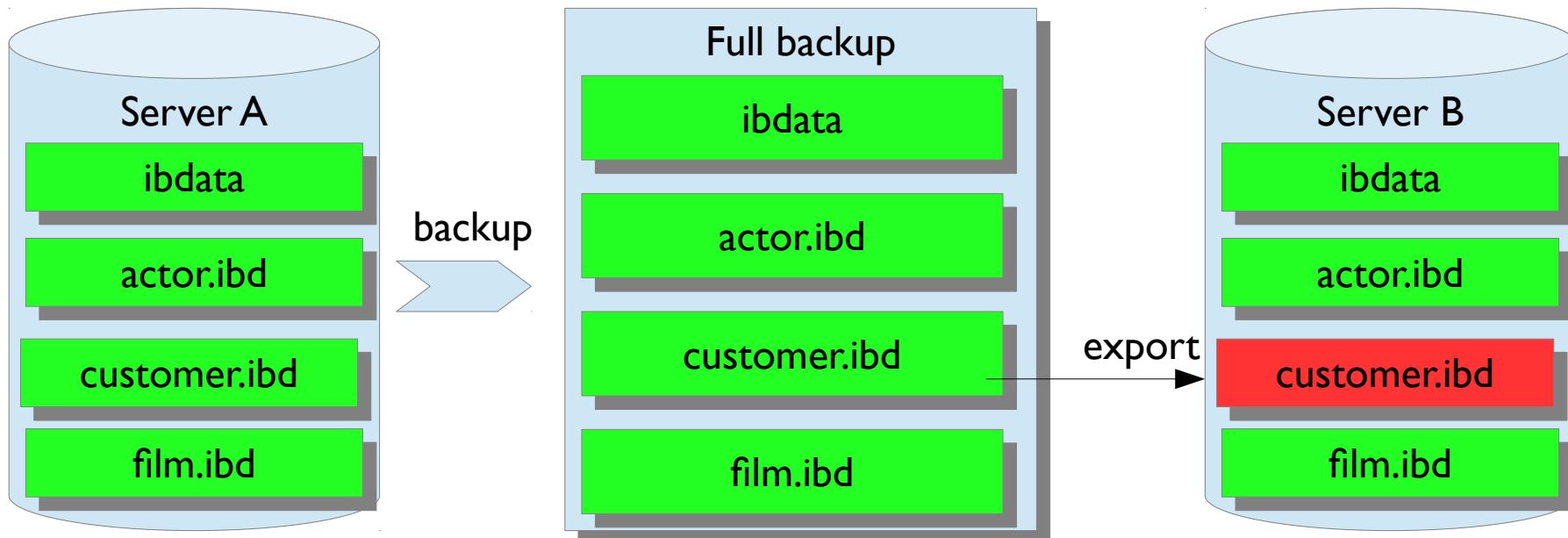
- creates N threads, each thread copying one file at a time
- utilizes disk hardware by copying multiple files in parallel
 - best for SSDs
 - less seeks on HDDs due to more merged reqs by I/O scheduler
 - YMMV, benchmarking before using is recommended

Parallel file copying

```
$ innobackupex --parallel=4 --no-timestamp /data/backup
...
[01] Copying ./ibdata1 to /data/backup/ibdata1
[02] Copying ./sakila/actor.ibd to /data/backup./sakila/actor.ibd
[03] Copying ./sakila/customer.ibd to /data/backup./sakila/customer.ibd
[04] Copying ./sakila/film.ibd to /data/backup./sakila/film.ibd
```

- **works only with multiple InnoDB tablespaces**
(innodb_file_per_table=1)
- in XtraBackup 1.6 could not be used with streaming backups
- **works with any backup types in XtraBackup 2.0**

Restoring individual tables: export



- problem: restore individual InnoDB table(s) from a full backup to another server
- use `--export` to prepare
- use improved table import feature in Percona Server to restore
- `innodb_file_per_table=1`

Restoring individual tables: export

Why not just copy the .ibd file?

- metadata:
 - InnoDB data dictionary (space ID, index IDs, pointers to root index pages)
 - .ibd page fields (space ID, LSNs, transaction IDs, index ID, etc.)
- xtrabackup --export dumps index metadata to .exp files on prepare
- Percona Server uses .exp files to update both data dictionary and .ibd on import

Restoring individual tables: export

```
$ xtrabackup --prepare --export --innodb-file-per-table=1  
--target-dir=/data/backup
```

...

```
xtrabackup: export metadata of table 'sakila/customer' to  
file `./sakila/customer.exp` (4 indexes)
```

```
xtrabackup: name=PRIMARY, id.low=23, page=3
```

```
xtrabackup: name=idx_fk_store_id, id.low=24, page=4
```

```
xtrabackup: name=idx_fk_address_id, id.low=25, page=5
```

```
xtrabackup: name=idx_last_name, id.low=26, page=6
```

...

Restoring individual tables: import

- improved import only available in Percona Server
- can be either the same or a different server instance:
 - (on different server to create .frm)
`CREATE TABLE customer(...);`
 - `SET FOREIGN_KEY_CHECKS=0;`
 - `ALTER TABLE customer DISCARD TABLESPACE;`
 - <copy `customer.ibd` to the database directory>
 - `SET GLOBAL innodb_import_table_from_xtrabackup=1;`
(Percona Server 5.5)
or
`SET GLOBAL innodb_expand_import=1;`
(Percona Server 5.1)
 - `ALTER TABLE customer IMPORT TABLESPACE;`
 - `SET FOREIGN_KEY_CHECKS=1;`

Restoring individual tables: import

- Improved table import is only available in Percona Server
- tables can only be imported to the same server with MySQL (with limitations):
 - there must be no DROP/CREATE/TRUNCATE/ALTER between taking backup and importing the table

```
mysql> ALTER TABLE customer DISCARD TABLESPACE;
```

<copy customer.ibd to the database directory>

```
mysql> ALTER TABLE customer IMPORT TABLESPACE;
```

Partial backups

- backup individual tables/schemas rather than the entire dataset
- InnoDB tables:
 - require `innodb_file_per_table=1`
 - restored in the same way as individual tables from a full backup
 - same limitations with the standard MySQL server (same server, no DDL)
 - no limitations with Percona Server when `innodb_import_table_from_xtrabackup` is enabled

Partial backups: selecting what to backup

innobackupex:

- streaming backups:
 - `--databases="database1[.table1] ..."`,
e.g.: `--databases="employees sales.orders"`
- local backups:
 - `--tables-file=filename`, file contains database.table, one per line
 - `--include=regexp`,
e.g.: `--include='^database(1|2)\.reports.*'`

xtrabackup:

- `--tables-file=filename` (same syntax as with innobackupex)
- `--tables=regexp` (equivalent to `--include` in innobackupex)

Partial backups: selecting what to backup

innobackupex:

- **streaming backups:**

- `--databases="database1[.table1] ..."`,
e.g.:`--databases="employees sales.orders"`

- **local backups:**

- `--tables-file=filename`, file contains database.table, one per line
- `--include=regexp`,
e.g.:`--include='^database(1|2)\.reports.*'`

xtrabackup:

- `--tables-file=filename` (same syntax as with innobackupex)
- `--tables=regexp` (equivalent to `--include` in innobackupex)

Partial backups: selecting what to backup

innobackupex:

- streaming backups:

- --databases="database1[.table1] ...",
e.g.: --databases="employees sales.orders"

- local backups:

- --tables-file=*filename*, file contains database.table, one per line
- --include=*regexp*,
e.g.: --include='^database(1|2)\.reports.*'

xtrabackup:

- --tables-file=*filename* (same syntax as with innobackupex)
- --tables=*regexp* (equivalent to --include in innobackupex)

Partial backups: selecting what to backup

innobackupex:

- streaming backups:
 - `--databases="database1[.table1] ..."`,
e.g.: `--databases="employees sales.orders"`
- local backups:
 - `--tables-file=filename`, file contains database.table, one per line
 - `--include=regexp`,
e.g.: `--include='^database(1|2)\.reports.*'`

xtrabackup:

- `--tables-file=filename` (same syntax as with innobackupex)
- `--tables=regexp` (equivalent to `--include` in innobackupex)

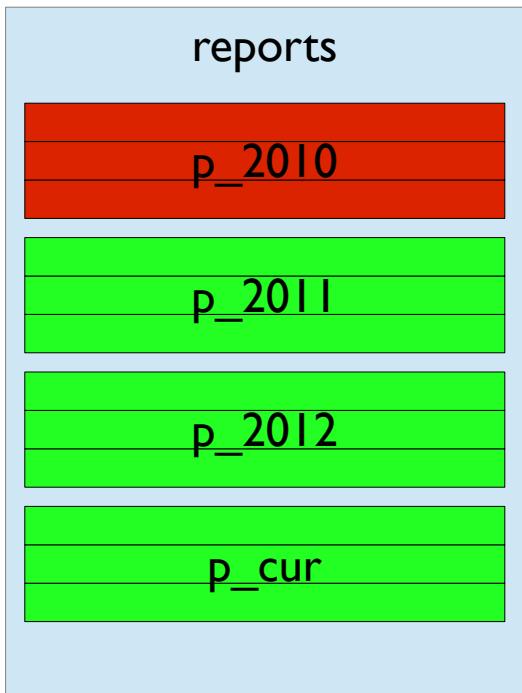
Partial backups: preparing

```
$ xtrabackup --prepare --export --target-dir=./  
...  
120407 18:04:57  InnoDB: Error: table 'sakila/store'  
InnoDB: in InnoDB data dictionary has tablespace id 24,  
InnoDB: but tablespace with that id or name does not exist. It will be  
removed from data dictionary.  
...  
xtrabackup: export option is specified.  
xtrabackup: export metadata of table 'sakila/customer' to file  
`./sakila/customer.exp` (4 indexes)  
xtrabackup:      name=PRIMARY, id.low=62, page=3  
xtrabackup:      name=idx_fk_store_id, id.low=63, page=4  
xtrabackup:      name=idx_fk_address_id, id.low=64, page=5  
xtrabackup:      name=idx_last_name, id.low=65, page=6  
...
```

Partial backups: restoring

- Non-InnoDB tables
 - just copy files to the database directory
- InnoDB (MySQL):
 - ALTER TABLE ... DISCARD/IMPORT TABLESPACE
 - same limitations on import (must be same server, no ALTER/DROP/TRUNCATE after backup)
- XtraDB (Percona Server):
 - xtrabackup --export **on prepare**
 - innodb_import_table_from_xtrabackup=1;
 - ALTER TABLE ... DISCARD/IMPORT TABLESPACE
 - no limitations

Partial backups: individual partitions



```
CREATE TABLE reports (d DATE, t TEXT)
PARTITION BY RANGE (YEAR(d))
( PARTITION p_2010 VALUES LESS THAN (2010),
  PARTITION p_2011 VALUES LESS THAN (2011),
  PARTITION p_2012 VALUES LESS THAN (2012),
  PARTITION p_cur VALUES LESS THAN (MAXVALUE)
) ;
```

Problem:

- archive p_2010 first
- purge p_2010

Partial backups: individual partitions

Can I backup individual partitions with XtraBackup?

Yes!

MyISAM or InnoDB (with multiple tablespaces) partitions are like normal tables with specially formatted names:

```
$ ls -l /usr/local/mysql/data/test/reports*  
  
/usr/local/mysql/data/test/reports#P#p_2010.ibd  
/usr/local/mysql/data/test/reports#P#p_2011.ibd  
/usr/local/mysql/data/test/reports#P#p_2012.ibd  
/usr/local/mysql/data/test/reports#P#p_cur.ibd  
/usr/local/mysql/data/test/reports.frm  
/usr/local/mysql/data/test/reports.par
```

Partial backups: individual partitions

```
$ xtrabackup --backup  
--datadir=/usr/local/mysql/data/  
--tables='reports#P#p_2010'  
  
...  
  
[01] Copying ./ibdata1 to /data/backup/ibdata1  
[01]     ...done  
[01] Copying ./test/reports#P#p_2010.ibd to  
/data/backup./test/reports#P#p_2010.ibd  
[01]     ...done  
[01] Skipping ./test/reports#P#p_2011.ibd  
[01] Skipping ./test/reports#P#p_2012.ibd  
[01] Skipping ./test/reports#P#p_cur.ibd
```

Partial backups: individual partitions

To prepare:

```
$ xtrabackup --prepare --export
--innodb_file_per_table --target-dir=./
```

```
$ ls -1 test/
```

```
reports#P#p_2010.exp
reports#P#p_2010.ibd
```

Partial backups: individual partitions

To restore:

- no support for attaching partitions in MySQL

```
mysql> CREATE TABLE reports_2010 (d DATE, t TEXT);
```

```
mysql> ALTER TABLE reports_2010 DISCARD TABLESPACE;
```

Partial backups: individual partitions

To restore (continued):

```
$ cp /data/backup/test/reports#P#p_2010.exp  
/usr/local/mysql/data/test/reports_2010.exp  
  
$ cp /data/backup/test/reports#P#p_2010.ibd  
/usr/local/mysql/data/test/reports_2010.ibd  
  
mysql> SET GLOBAL  
innodb_import_table_from_xtrabackup=1;  
  
mysql> ALTER TABLE report_2010 IMPORT TABLESPACE;
```

- same procedure to backup/restore MyISAM partitions (except DISCARD/IMPORT TABLESPACE)

Questions ?

End of part 1

Q&A

Part 2

Streaming

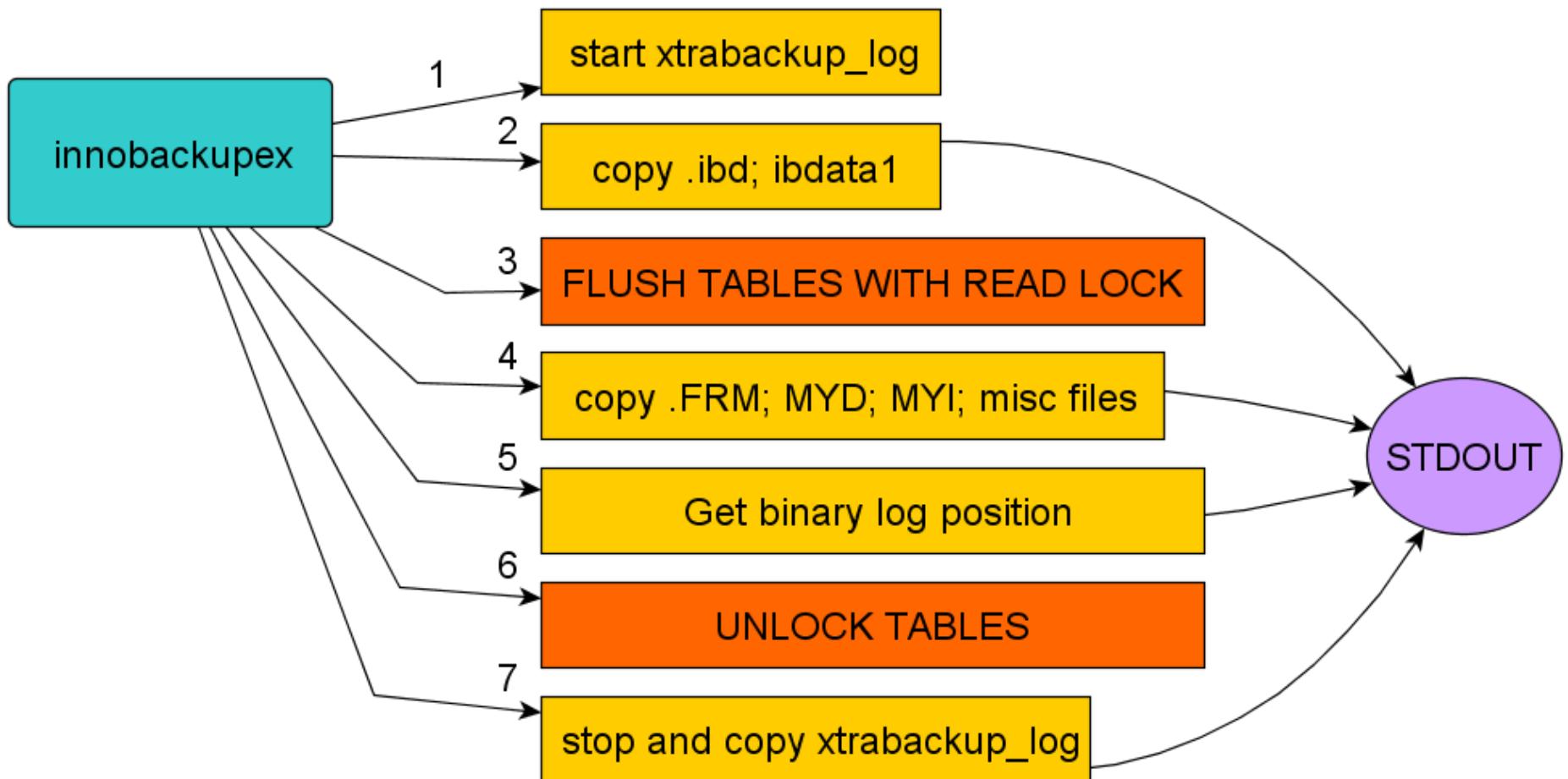
Unique for XtraBackup, not available in MEB

Local copy to

- Local partitions
- NFS partitions

Unix pipes

- `tar zcvf - /wwwdata | ssh root@192.168.1.201 "cat > /backup/wwwdata.tar.gz"`
- We want:
 - `innobackupex | script1 | scrip2 | ...`



Basic command

- `innobackupex --stream=tar /tmpdir`

Stream output

- ... File xtrabackup-stream-out ...

Tar -i

- -i is IMPORTANT

```
> tar --list -f backup.tar  
backup-my.cnf
```

```
> tar --list -i -f backup.tar  
backup-my.cnf  
ibdata1  
.sbtest/sbtest3.ibd  
.sbtest/sbtest2.ibd  
.sbtest/sbtest1.ibd  
.sbtest/sbtest4.ibd  
xtrabackup_binlog_info  
mysql/time_zone_leap_second.MYI  
mysql/ndb_binlog_index.MYD  
mysql/plugin frm  
mysql/time_zone_leap_second.MYD  
mysql/time_zone_name.frm  
mysql/tables_priv.frm  
mysql/event.MYD  
mysql/time_zone_name.MYI  
mysql/ndb_binlog_index.MYI  
mysqlproxies_priv.MYD  
mysqlproxies_priv.MYI  
mysqlfunc.MYI
```

Usage: compression

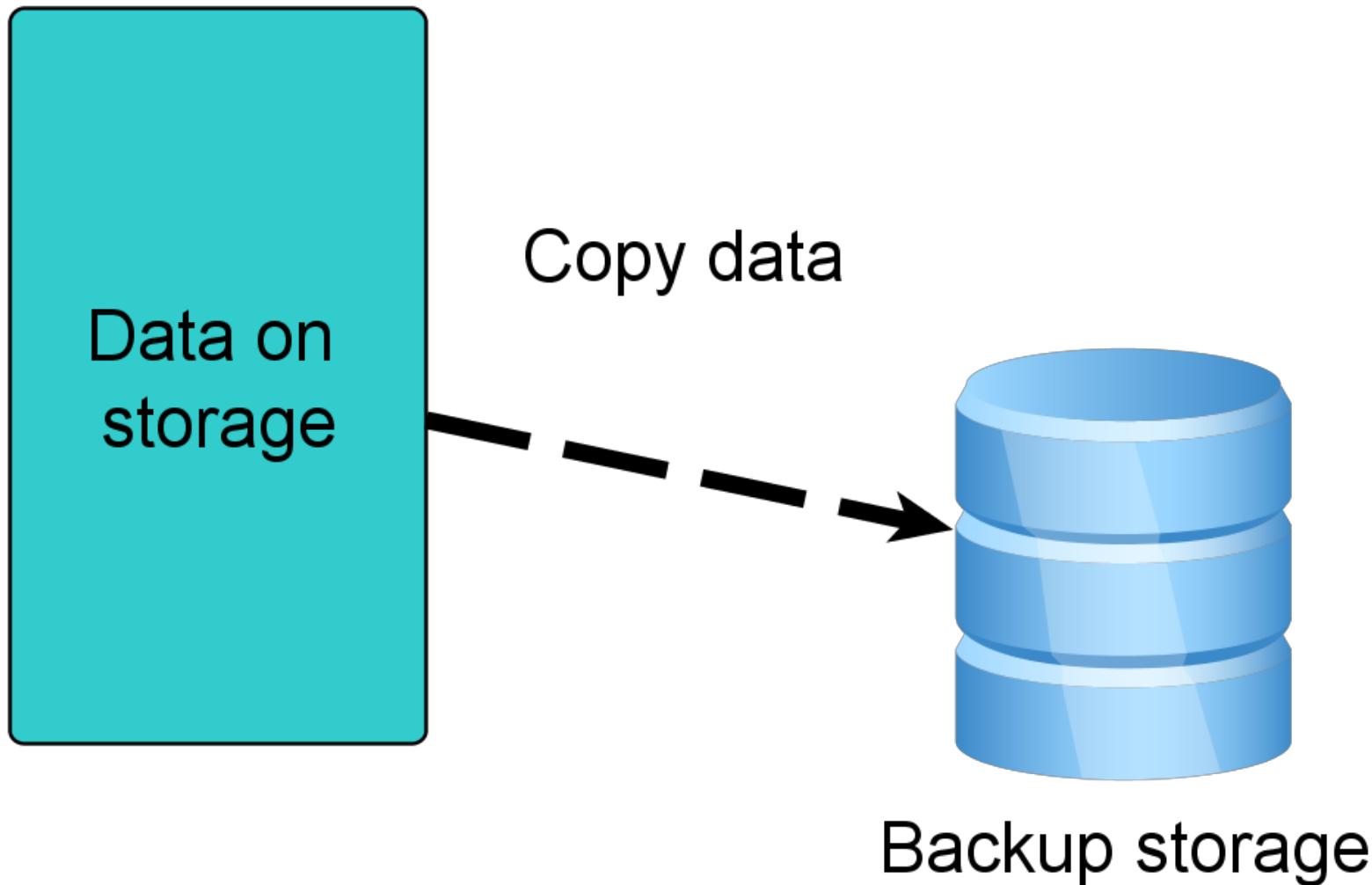
- `innobackupex --stream=tar ./ | gzip - > backup.tar.gz`
- `innobackupex --stream=tar ./ | qpress -io xtrabackup.tar > backup.tar.qpress`

Usage: copy to remote host

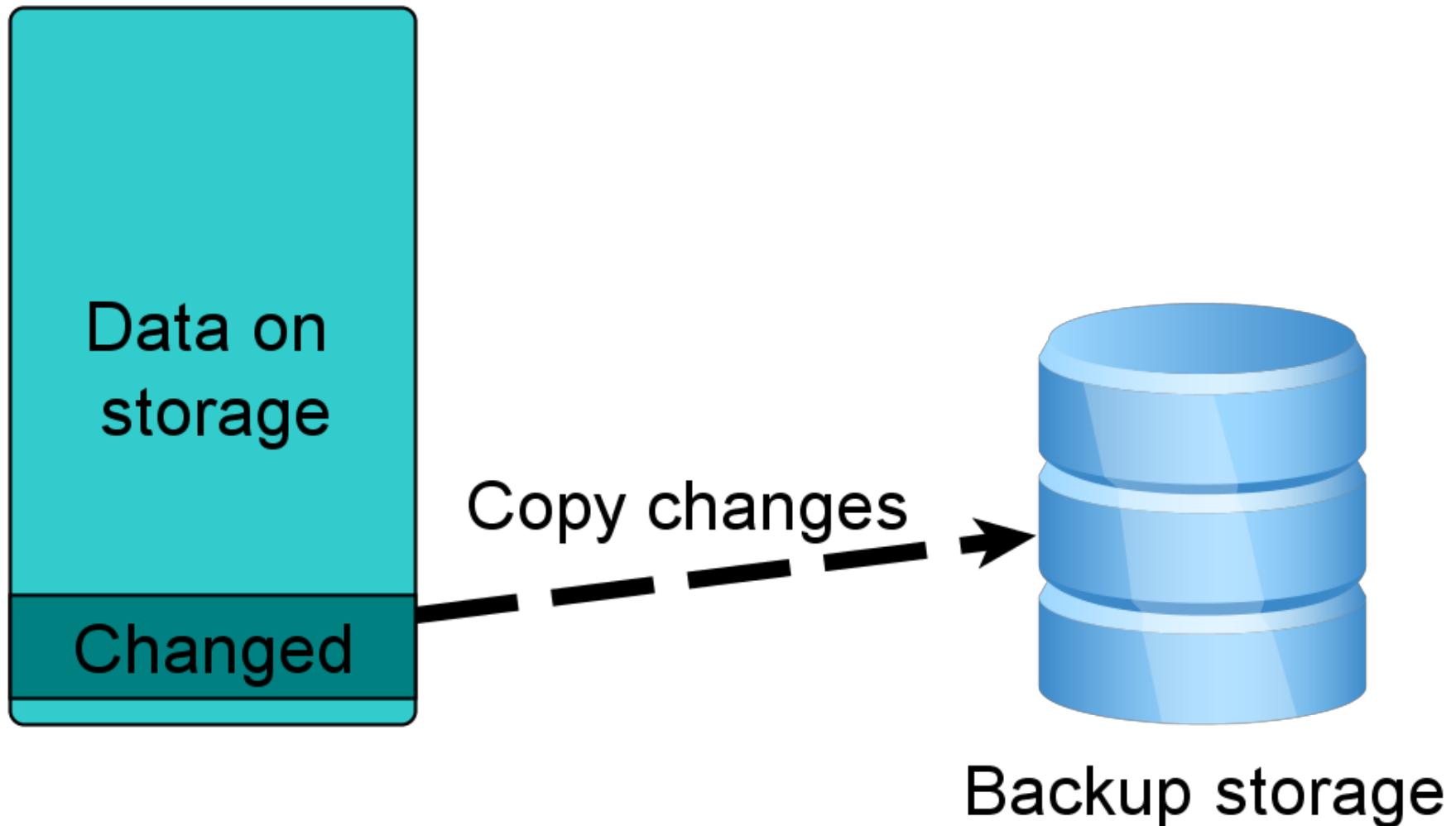
- `innobackupex --stream=tar ./ | ssh vadim@desthost "cat - > /data/vol1/mysqluc/backup.tar"`
- `ssh user@desthost "(nc -l 9999 > /data/backups/backup.tar.qpress &)" && innobackupex --stream=tar ./ | qpress -io backup.tar | nc desthost 9999`
- `ssh user@desthost "(nc -l 9999 | qpress -io backup.tar > /data/backups/backup.tar.qpress &)" && innobackupex --stream=tar ./ | nc desthost 9999`

Incremental backup

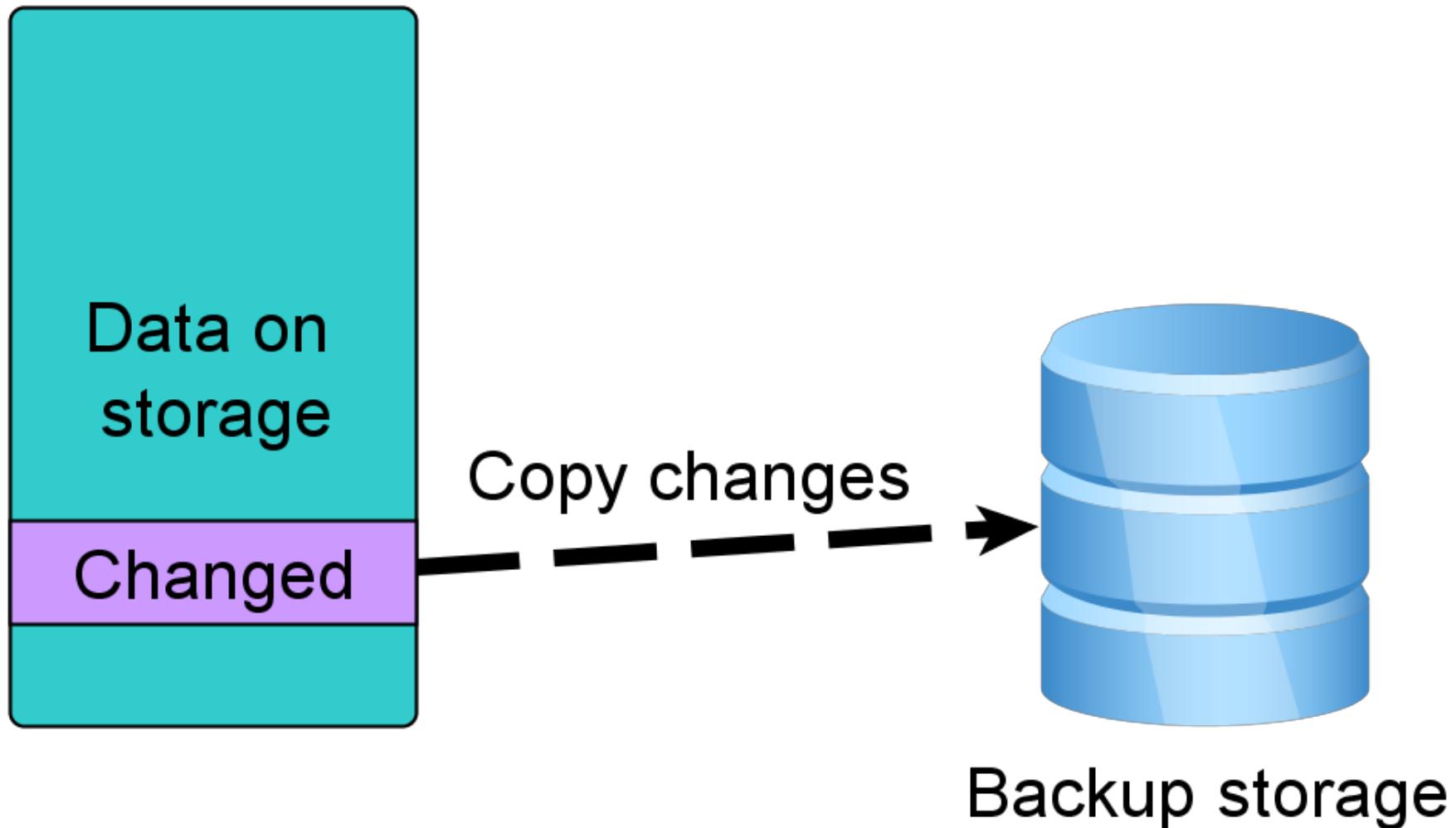
Full copy



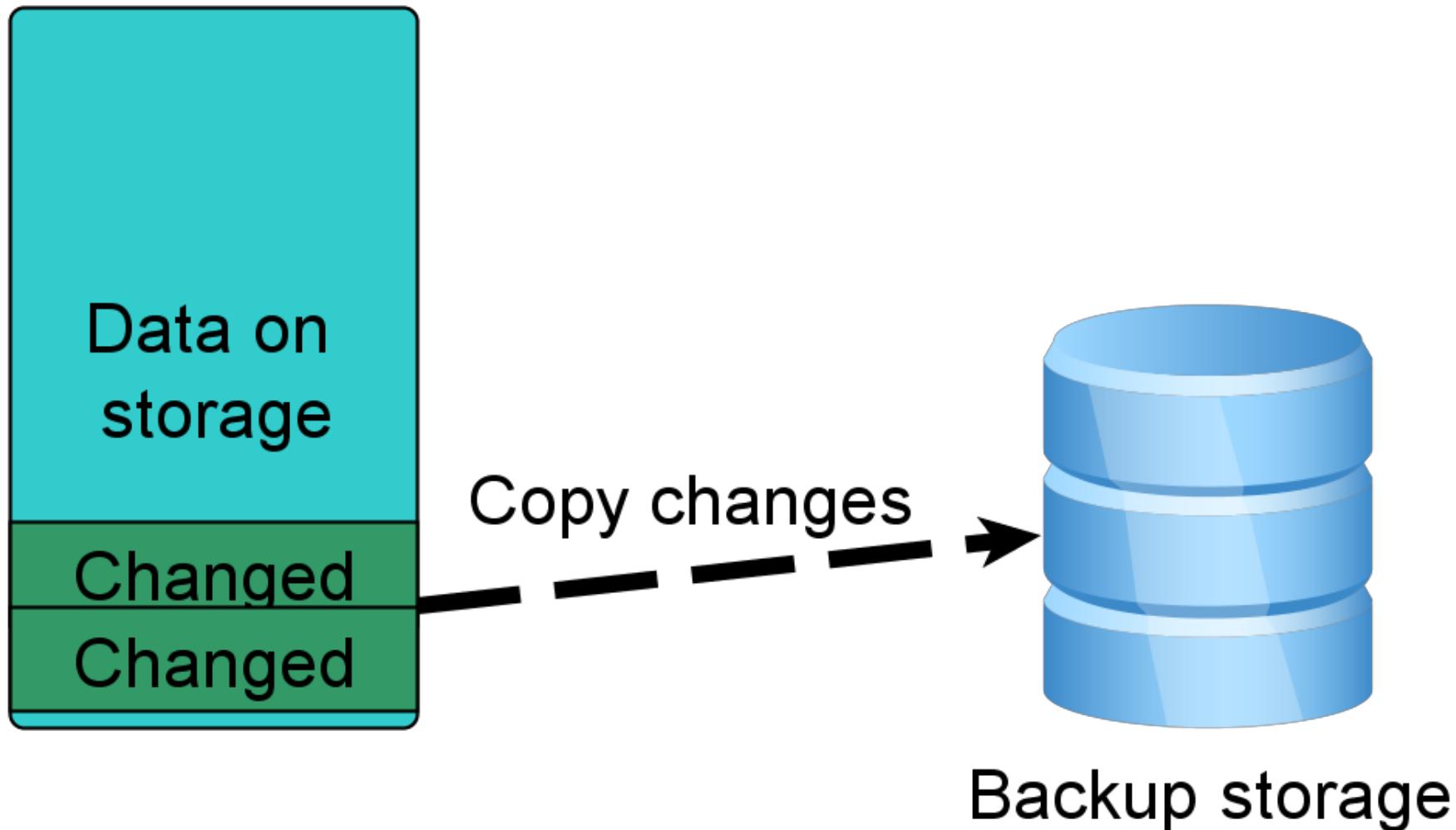
Incremental. Day 1 – changes since full



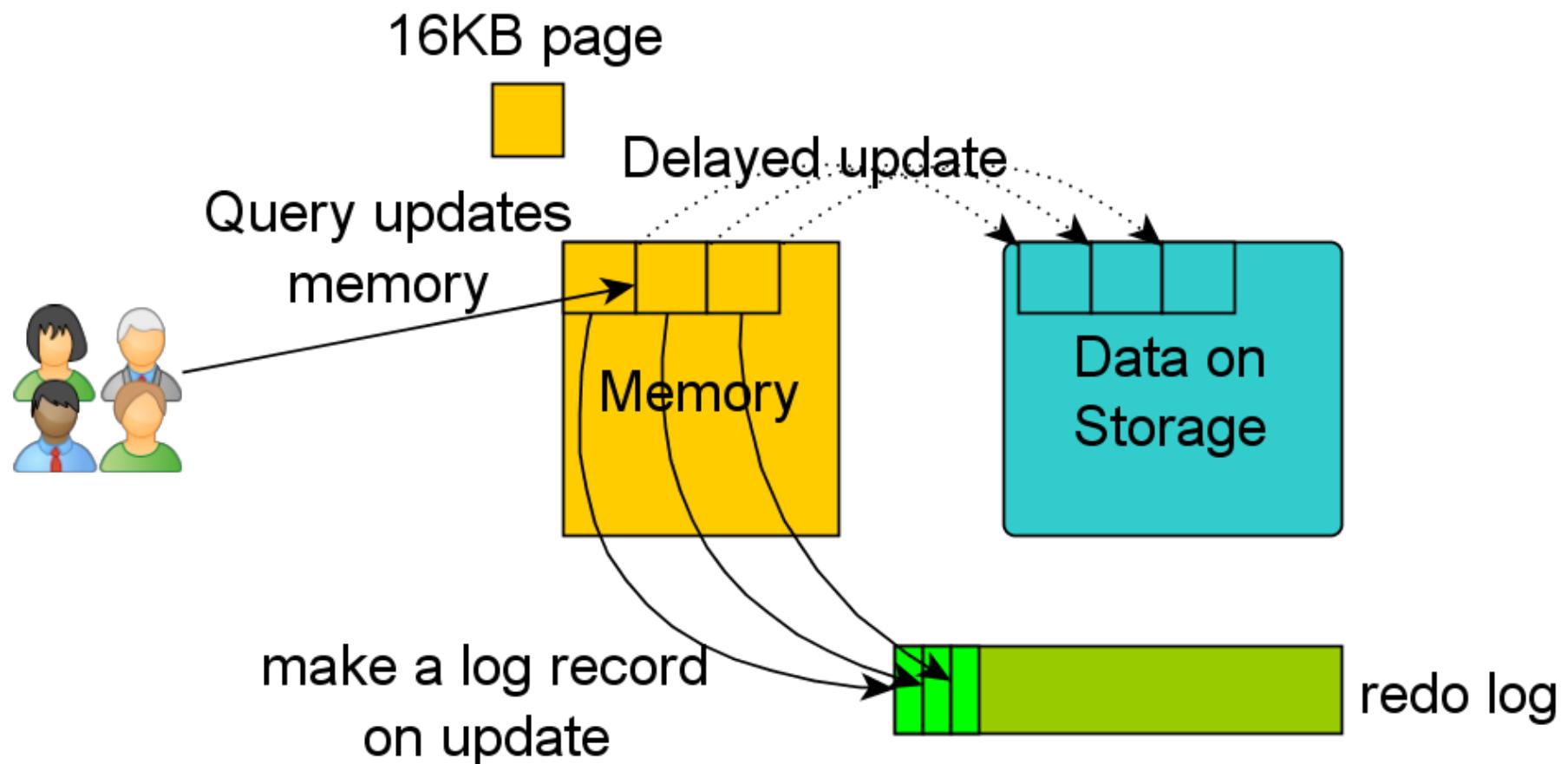
Incremental. Day 2 – changes since Incremental

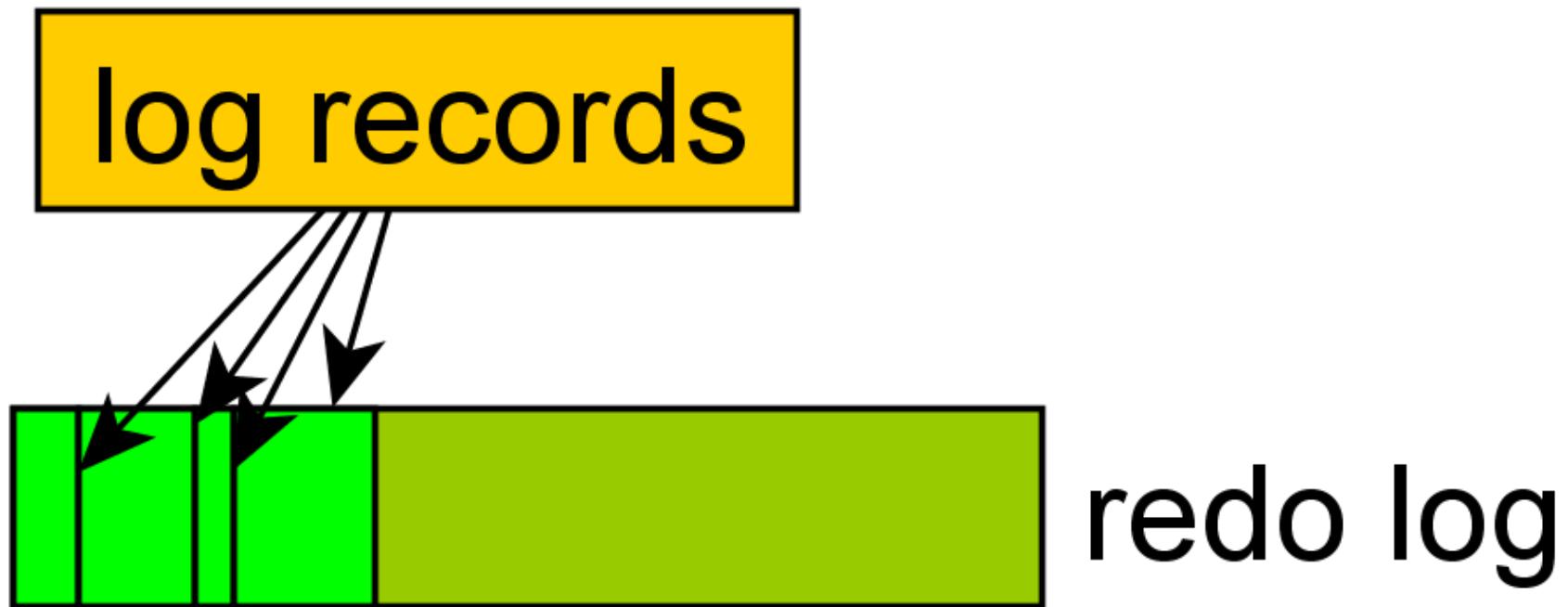


Differential. Changes since Full



To remind: InnoDB logs





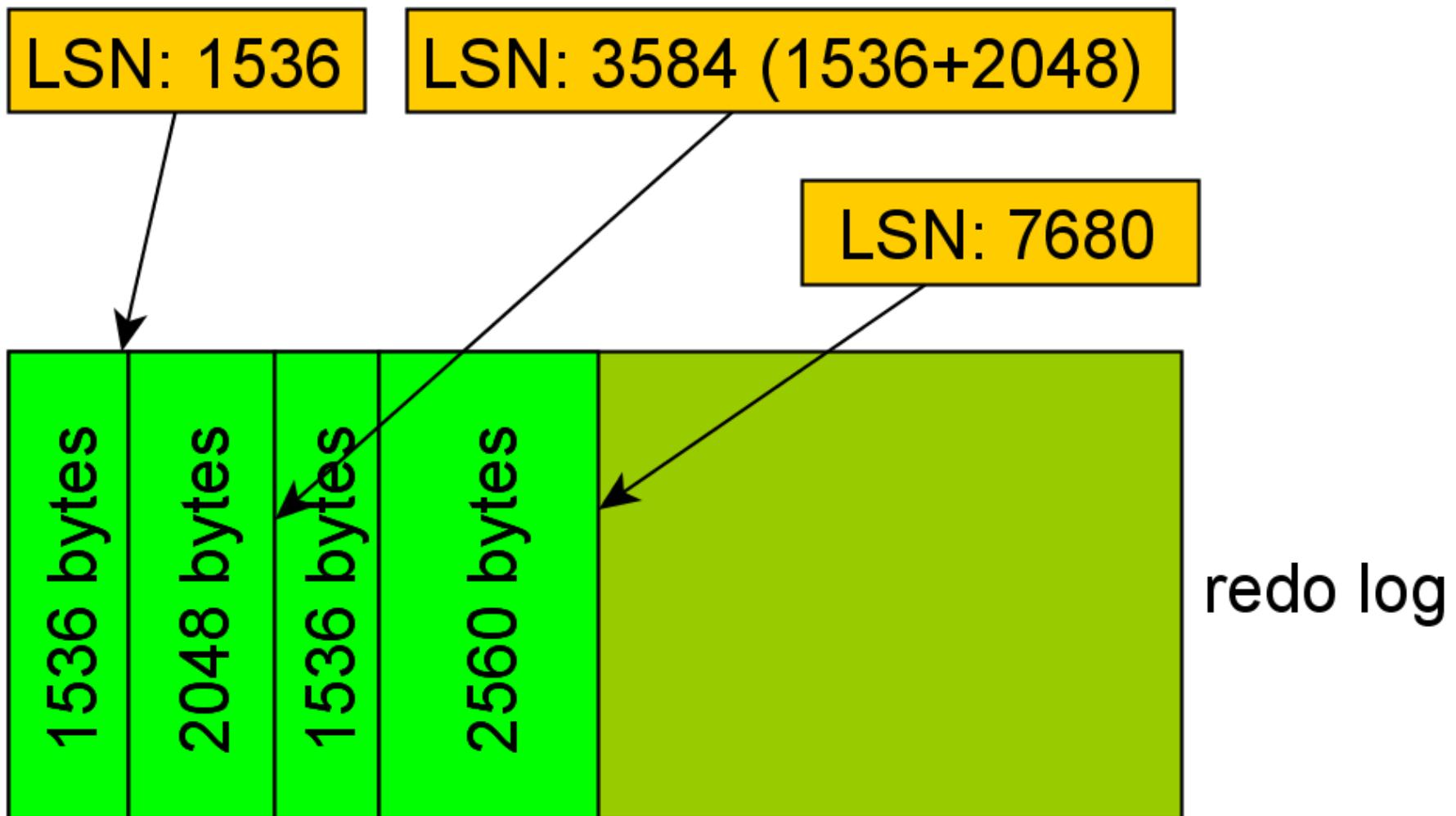
Log Sequential Number

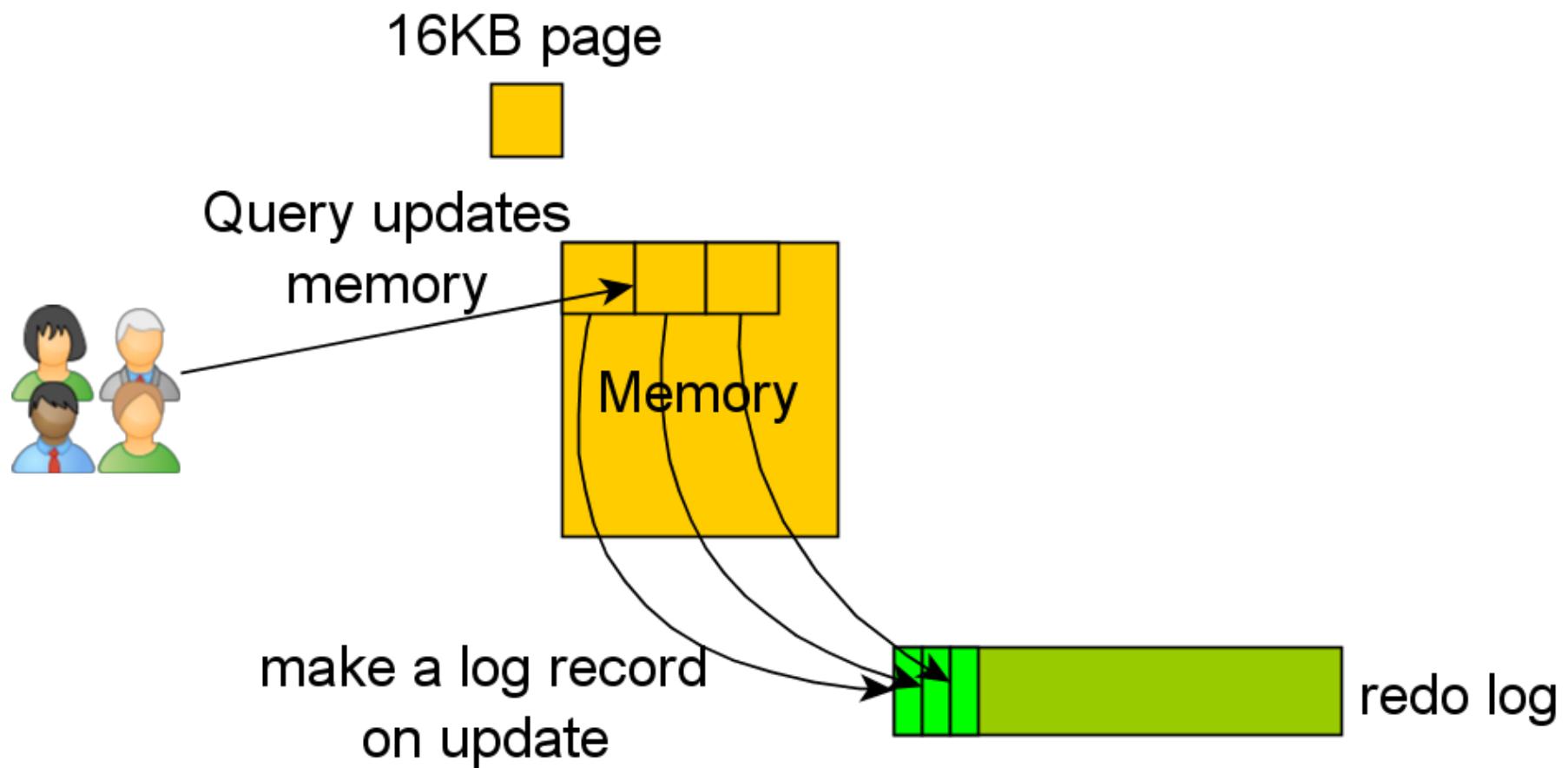
LSN

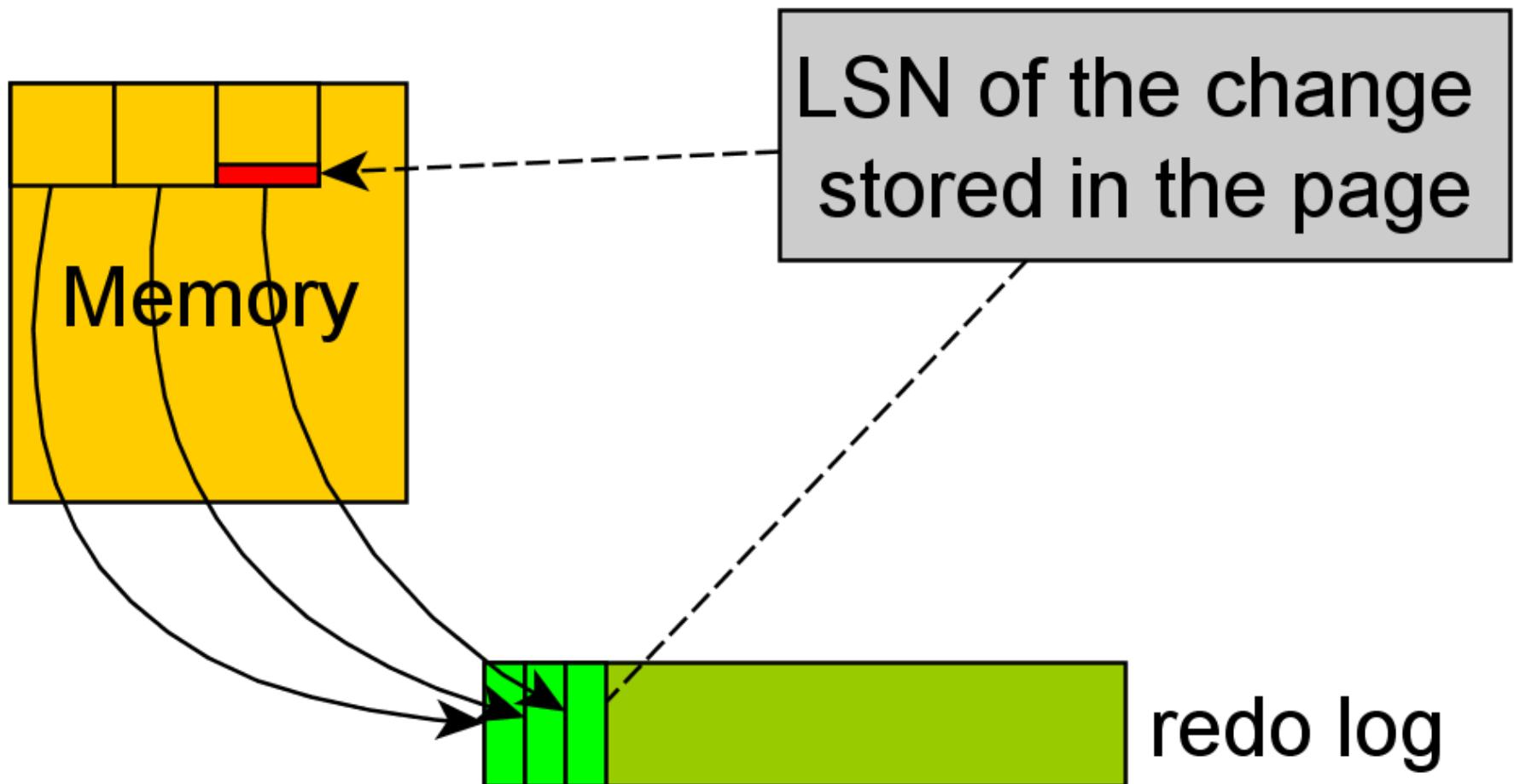
LSN

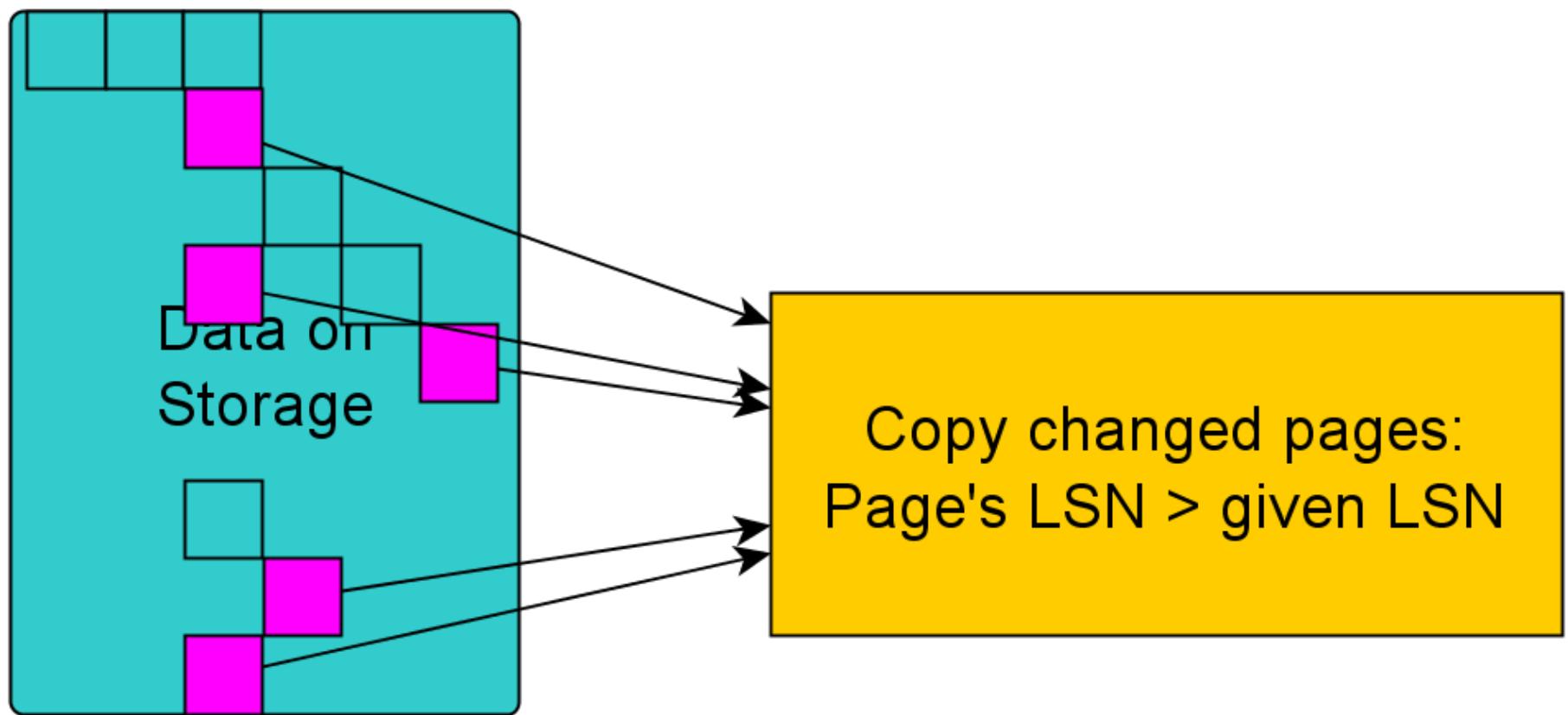
```
SHOW ENGINE INNODB STATUS\G

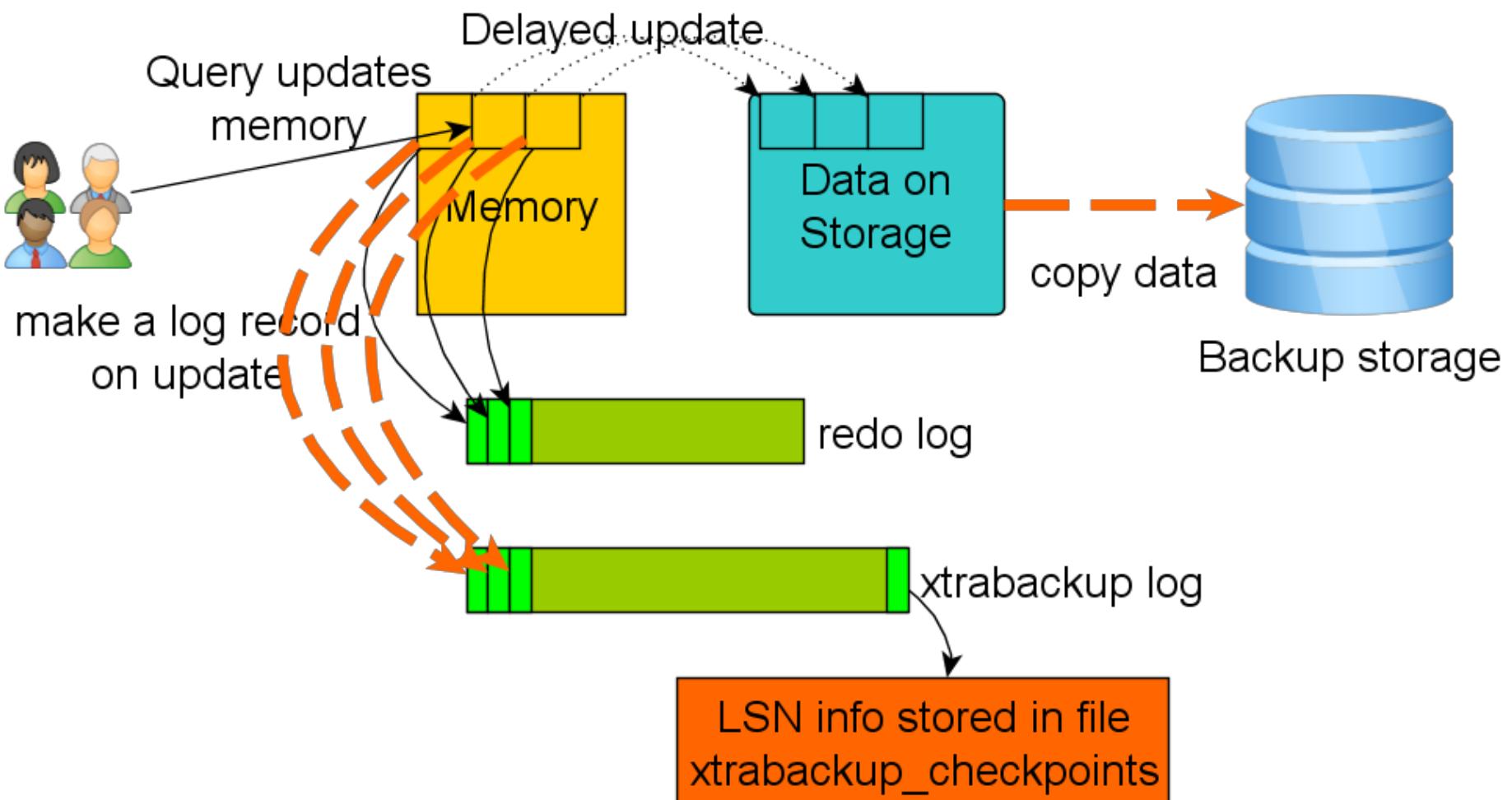
-----
LOG
-----
Log sequence number 180661841734
```





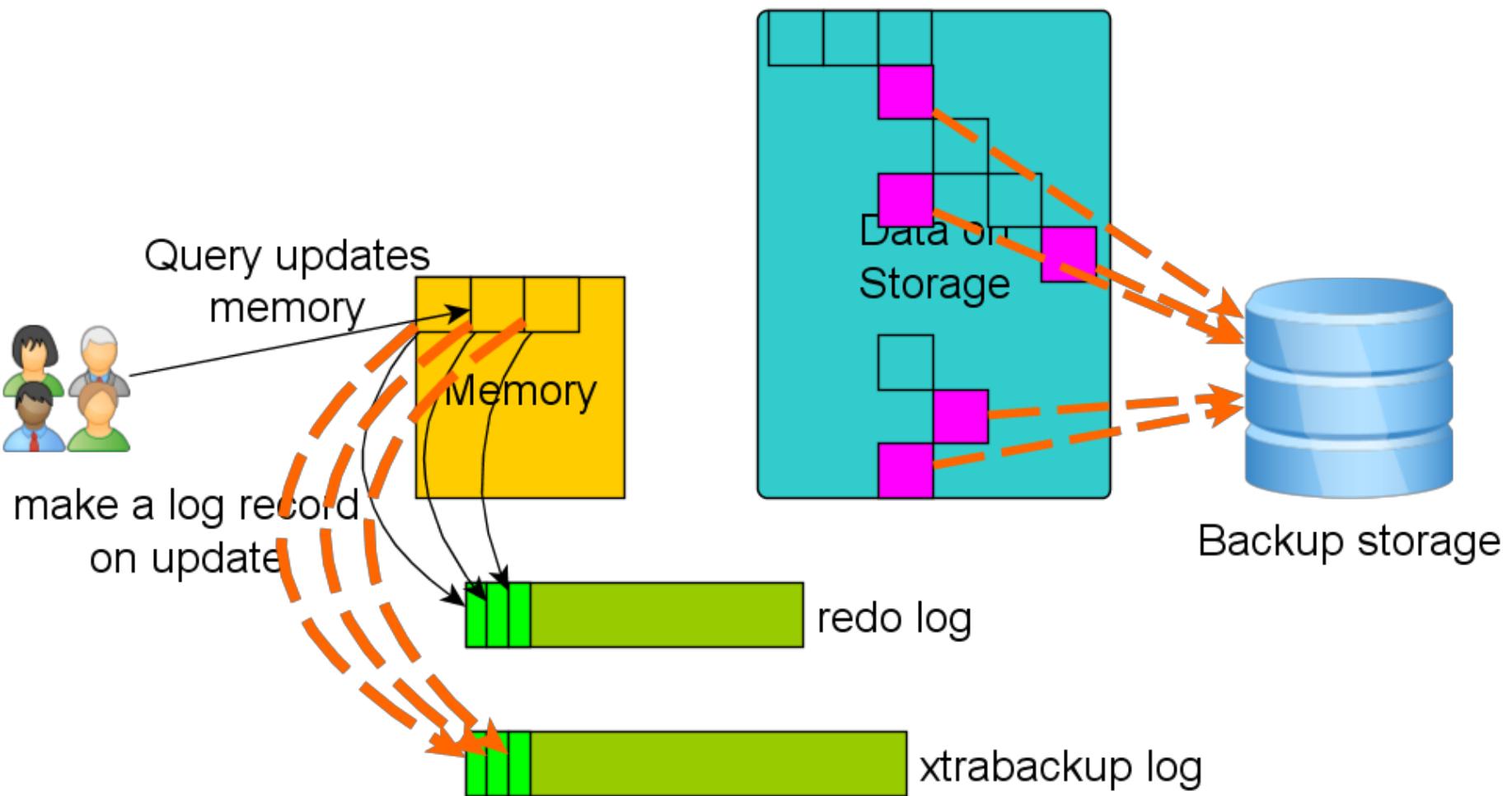




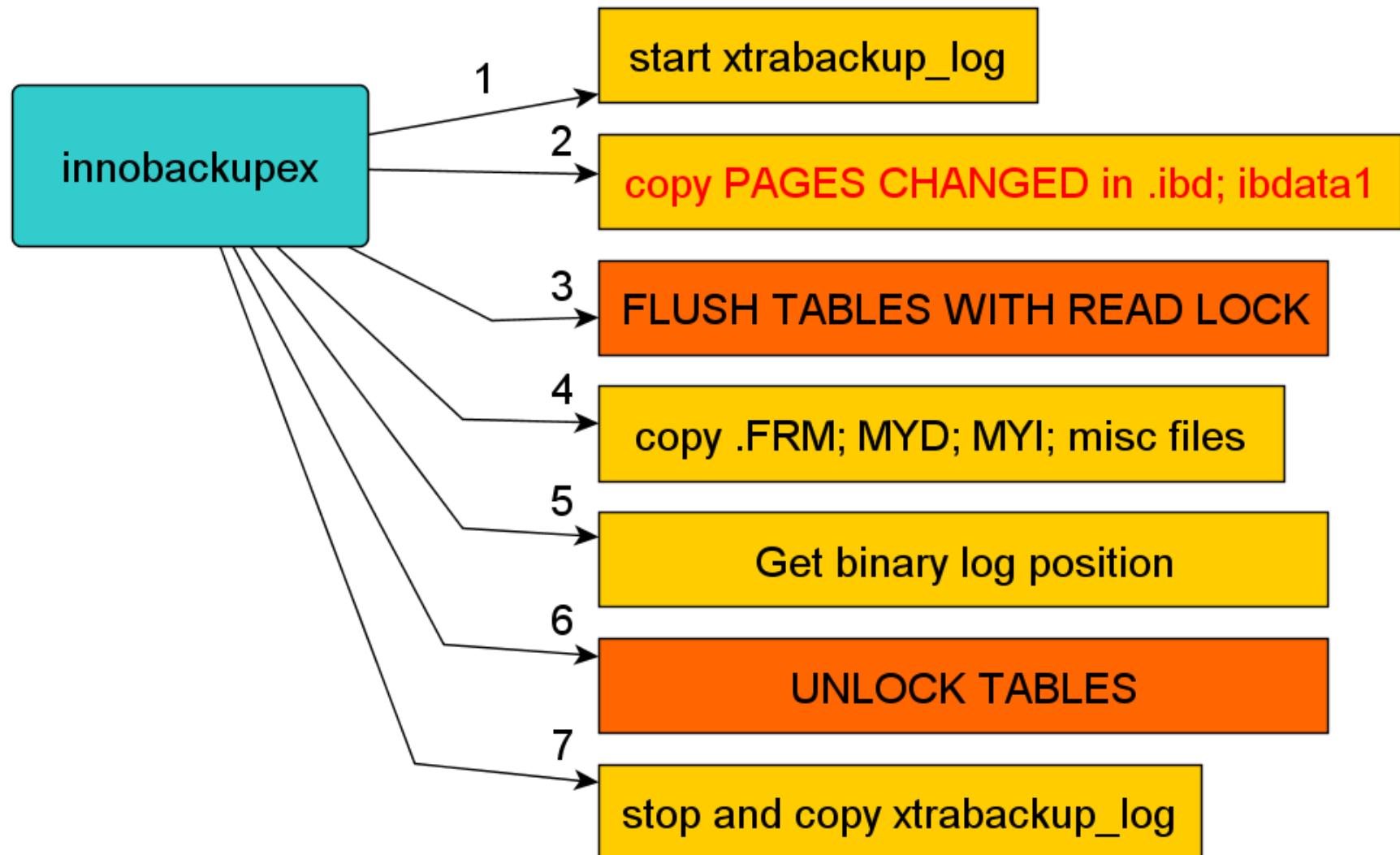


Xtrabackup_checkpoints

```
cat xtrabackup_checkpoints
backup_type = full-backuped
from_lsn = 0
to_lsn = 172573554953
last_lsn = 172728858570
```



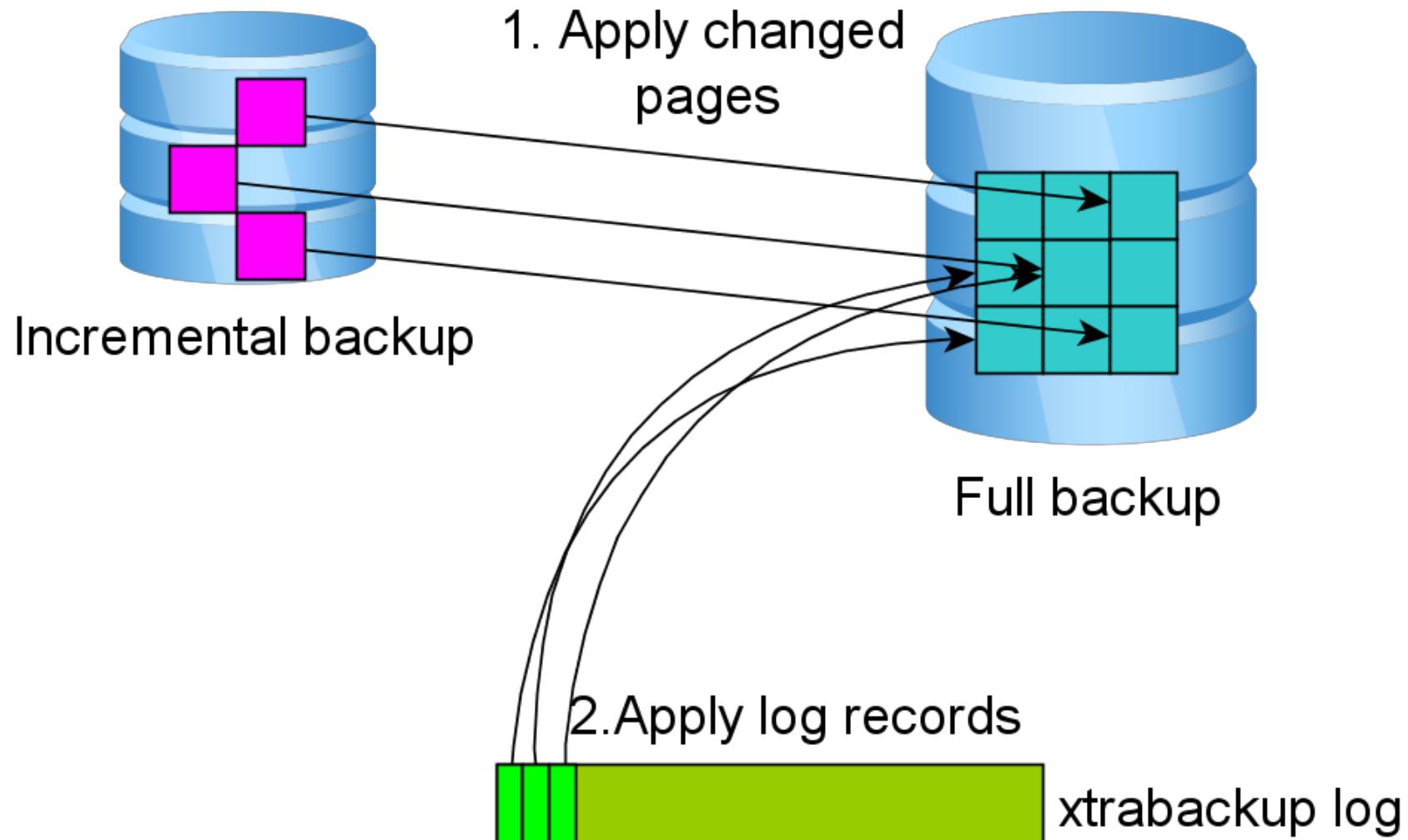
Innodb backup incremental process



Innobackupex incremental:

- Handles incremental changes to InnoDB
- DOES not handle MyISAM and other engines
 - Full copy instead

Incremental apply



Innodbupex commands

- innobackupex --incremental \
 --incremental-basedir=
 /previous/full/or/incremental \
 /data/backup/inc

Innodbupex from LSN

- `innobackupex --incremental \
--incremental-lsn=172573554953 \
/data/backup/inc`

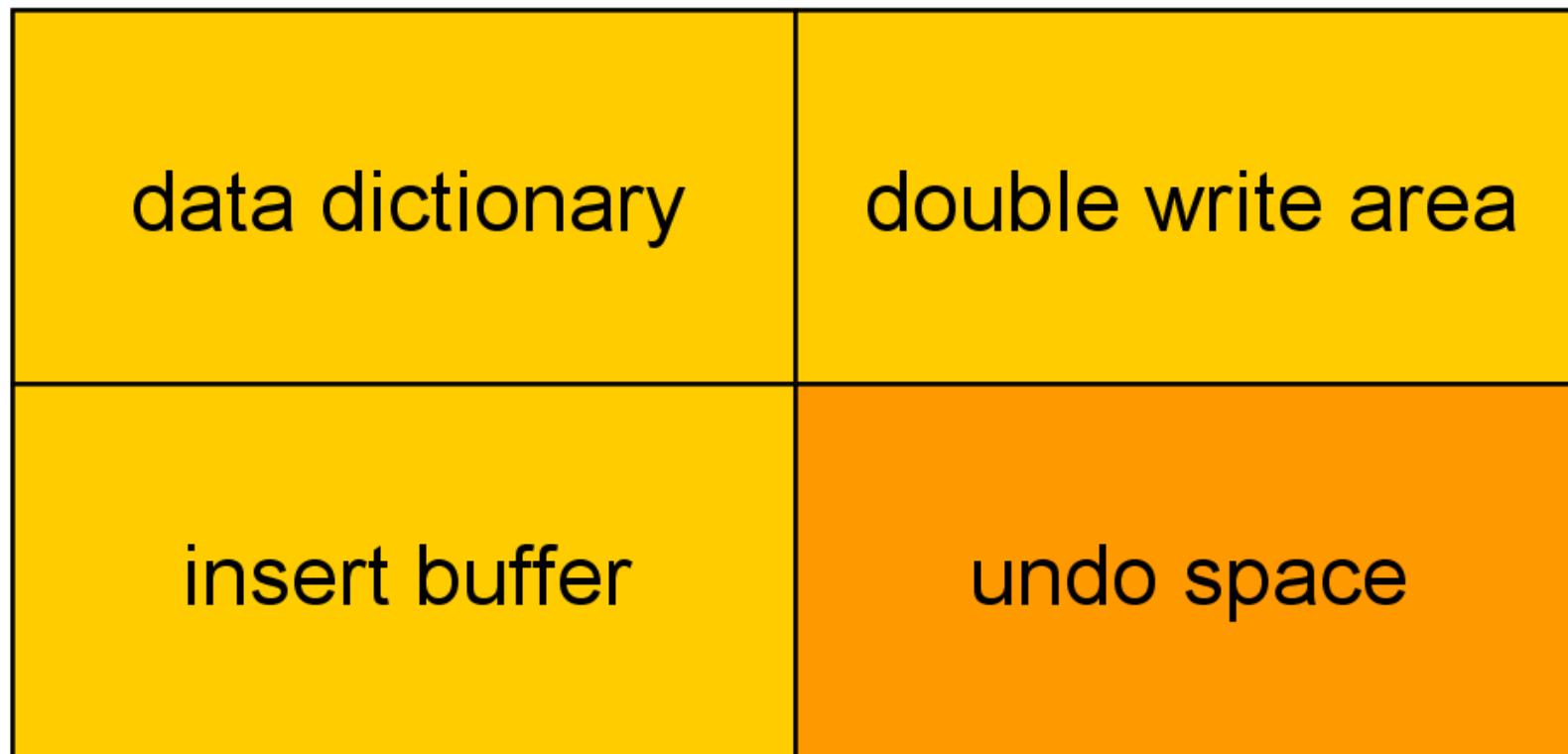
Innobackup incremental output

- ... File xtrabackup-incremental-out ...

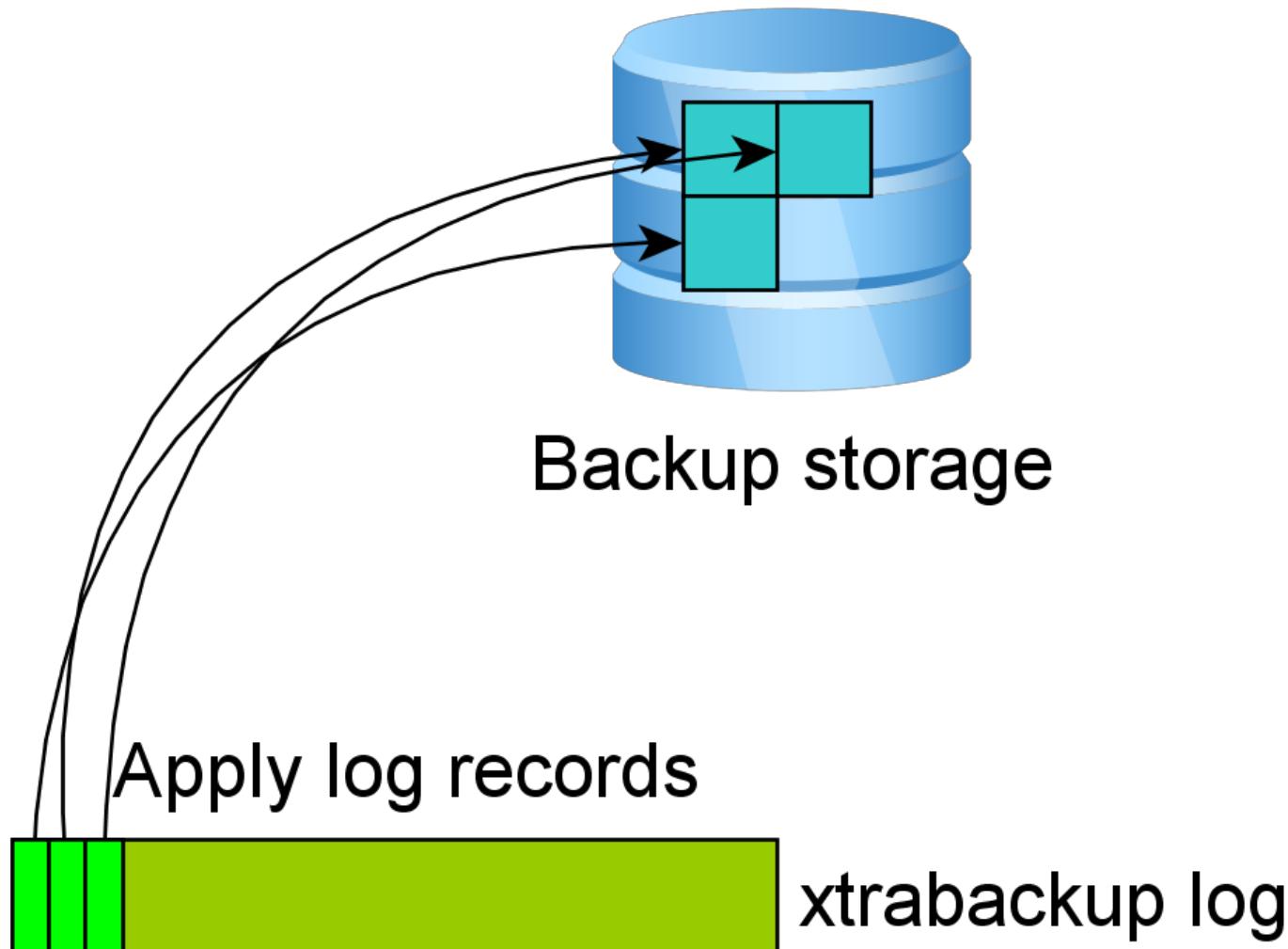
Incremental apply complications

Ibdata1 structure

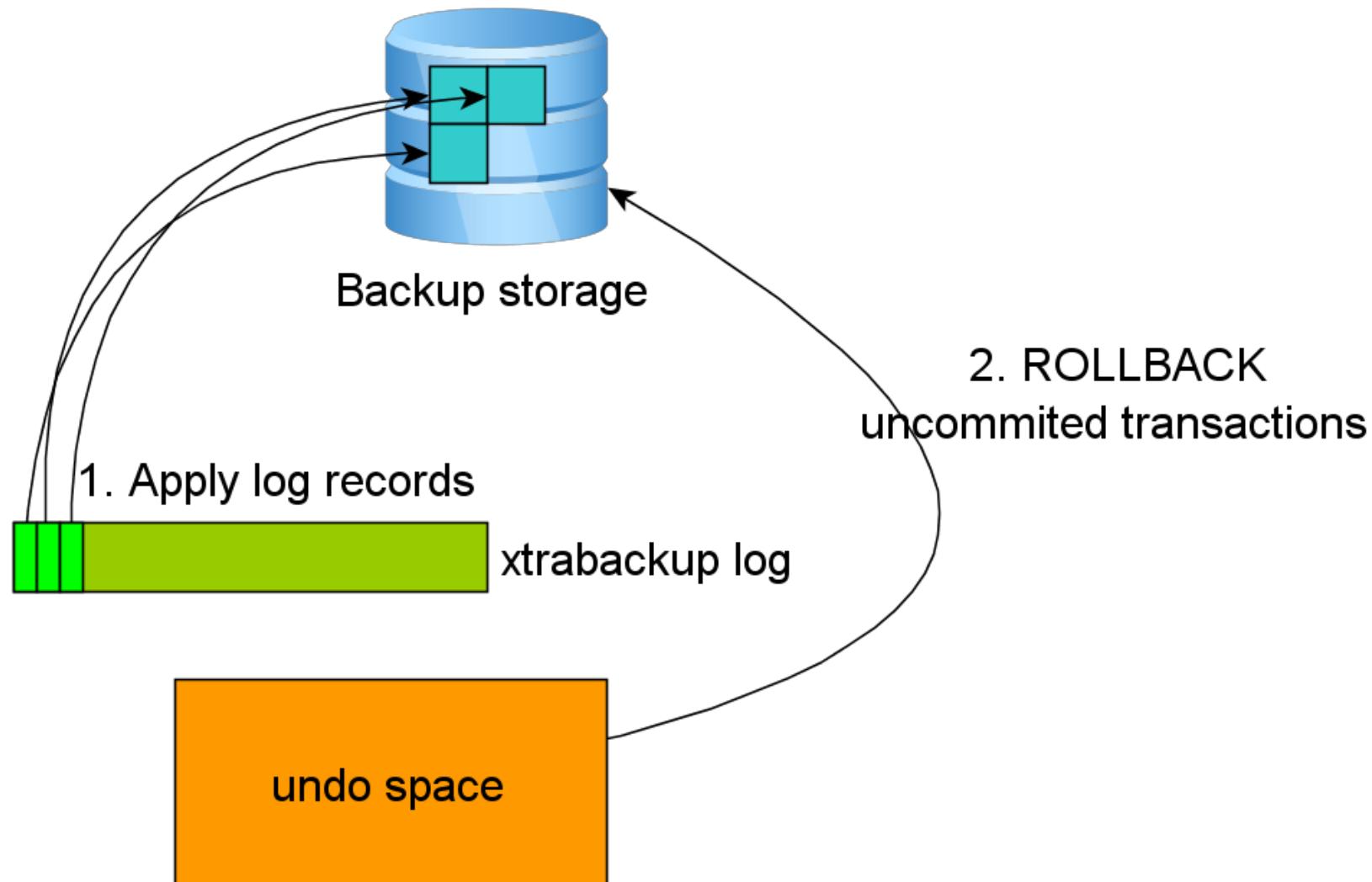
ibdata1
[system tablespace]



To remind: recovery process



This is what really happens



No need for ROLLBACK if
we apply incremental later

Innodbupex –redo-only

- Full backup: apply with –redo-only
- `innobackupex -apply-log -redo-only -use-memory=10G /data/backup/mysql-data`

Innodbupex apply incremental

- innobackupex --apply-log --use-memory=10G
/data/**full**/backup --incremental-
dir=/data/**incremental**/backup

Innodbupex apply output

- ... File xtrabackup-incremental-apply-out ...

Incremental streaming

New in Percona-XtraBackup 2.0

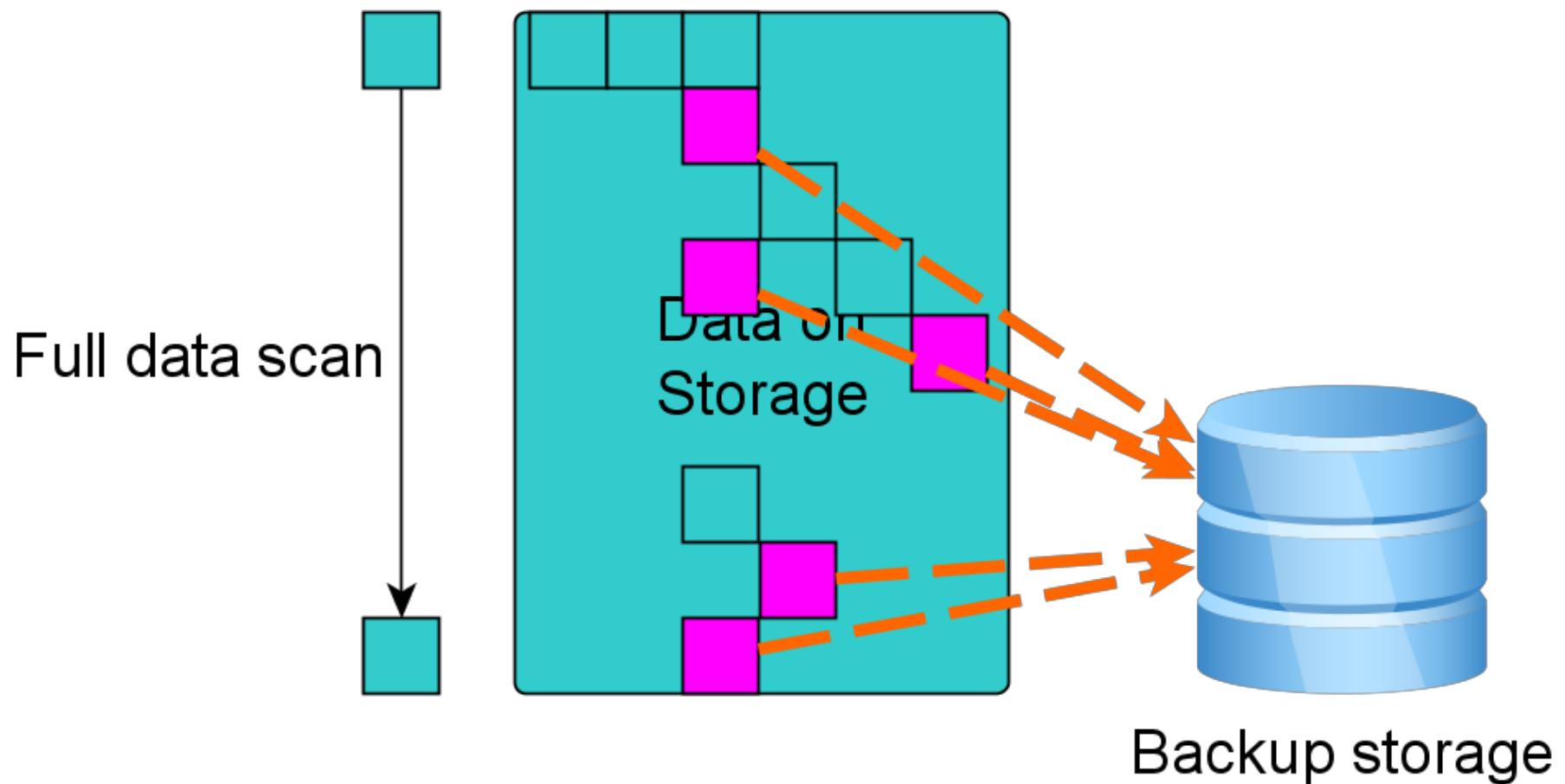
Why not in 1.6?

- Tar format has limitations
 - We need to know filesize before sending to tar
- New streaming format xbstream
- Xbstream supports compression and incremental backups
 - Custom format

Usage:

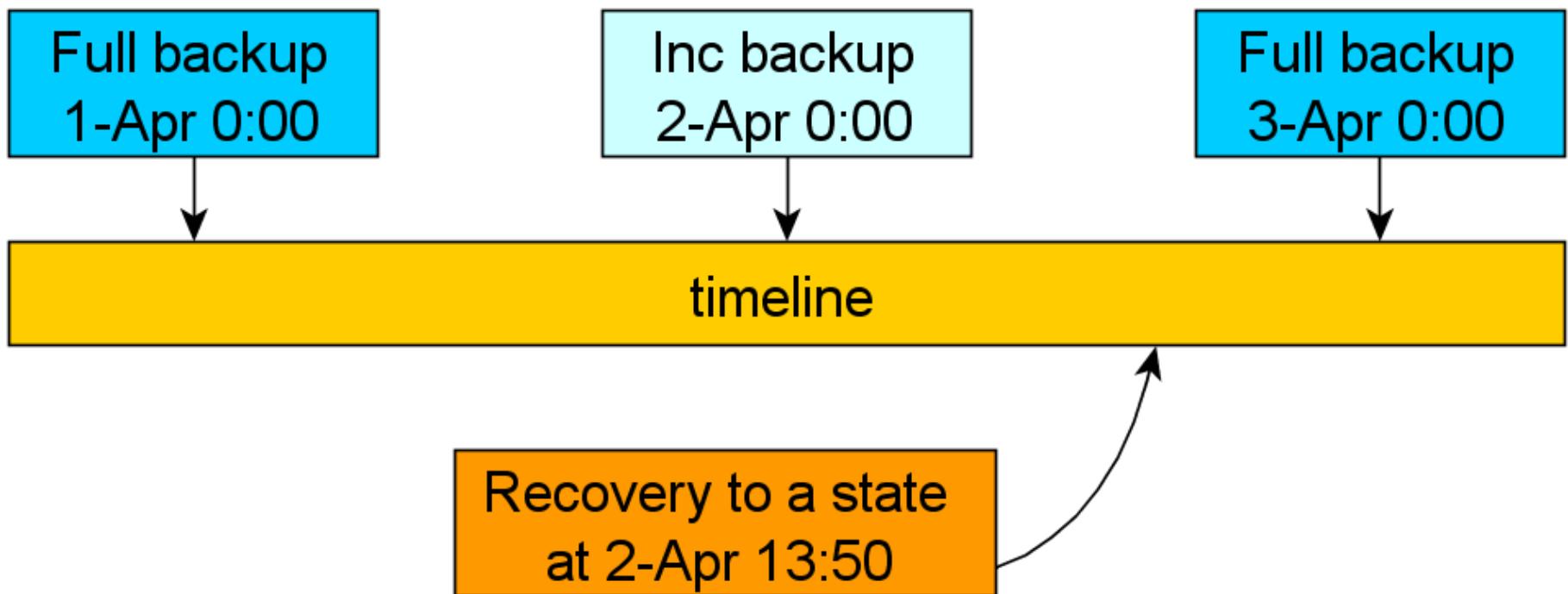
- `innobackupex --incremental --incremental-lsn=166678320660 --stream=xstream ./ > incremental.bak`
- Unpack:
- `xstream -x < incremental.bak`
- Incremental directly to remote server
- `innobackupex --incremental --incremental-lsn=166678320660 --stream=xstream ./ | ssh vadim@desthost "cat - | xstream -x -C /data/vol1/mysqluc/"`

Incremental now:

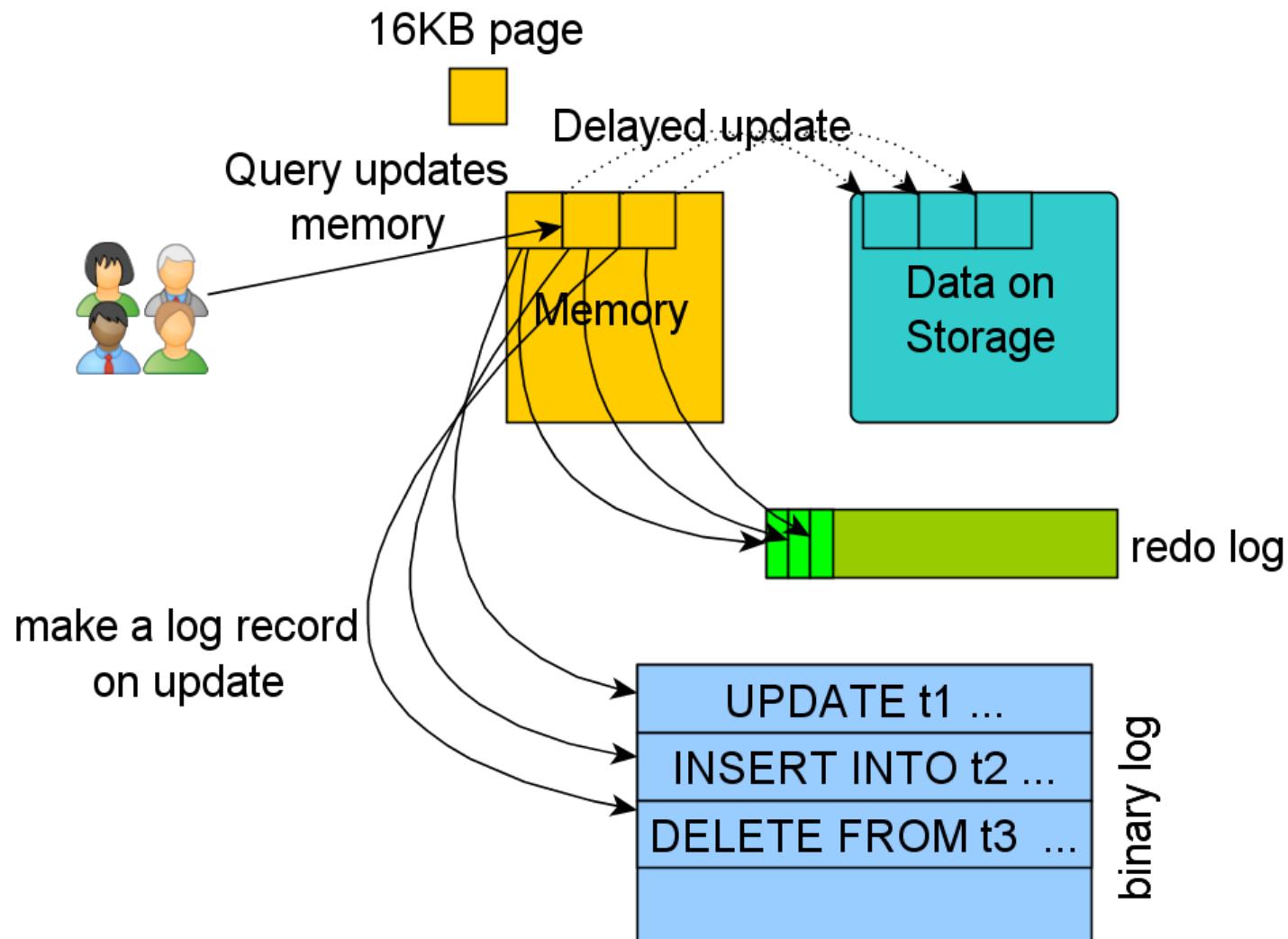


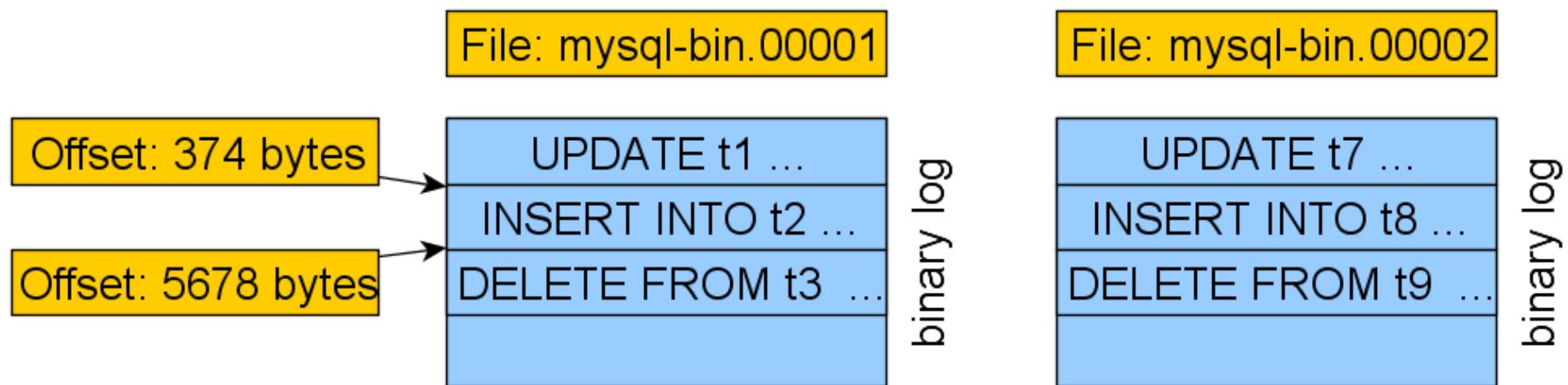
Point-In-Time Recovery

PITR



Binary logs





Position: mysql-bin.00001:5678

event header:

- timestamp

- ...

DELETE FROM t3 ...

binary log

Xtrabackup has binlog coordinates:

xtrabackup_binlog_info

- Master binary log position
- Result of SHOW MASTER STATUS

```
binlog.000001    68212201
```

xtrabackup_binlog_pos_innodb

- Master binary log position **STORED IN InnoDB** (ibdata1)

```
./binlog.000001 68212201
```

- Available only after –apply-log
- Can be used with –no-lock

In innobackupex output

- Master binary log position **STORED IN InnoDB** (ibdata1)

```
[notice (again)]
```

If you use binary log and don't use any hack
of group commit,

the binary log position seems to be:

```
InnoDB: Last MySQL binlog file position 0  
68212201, file name ./binlog.000001
```

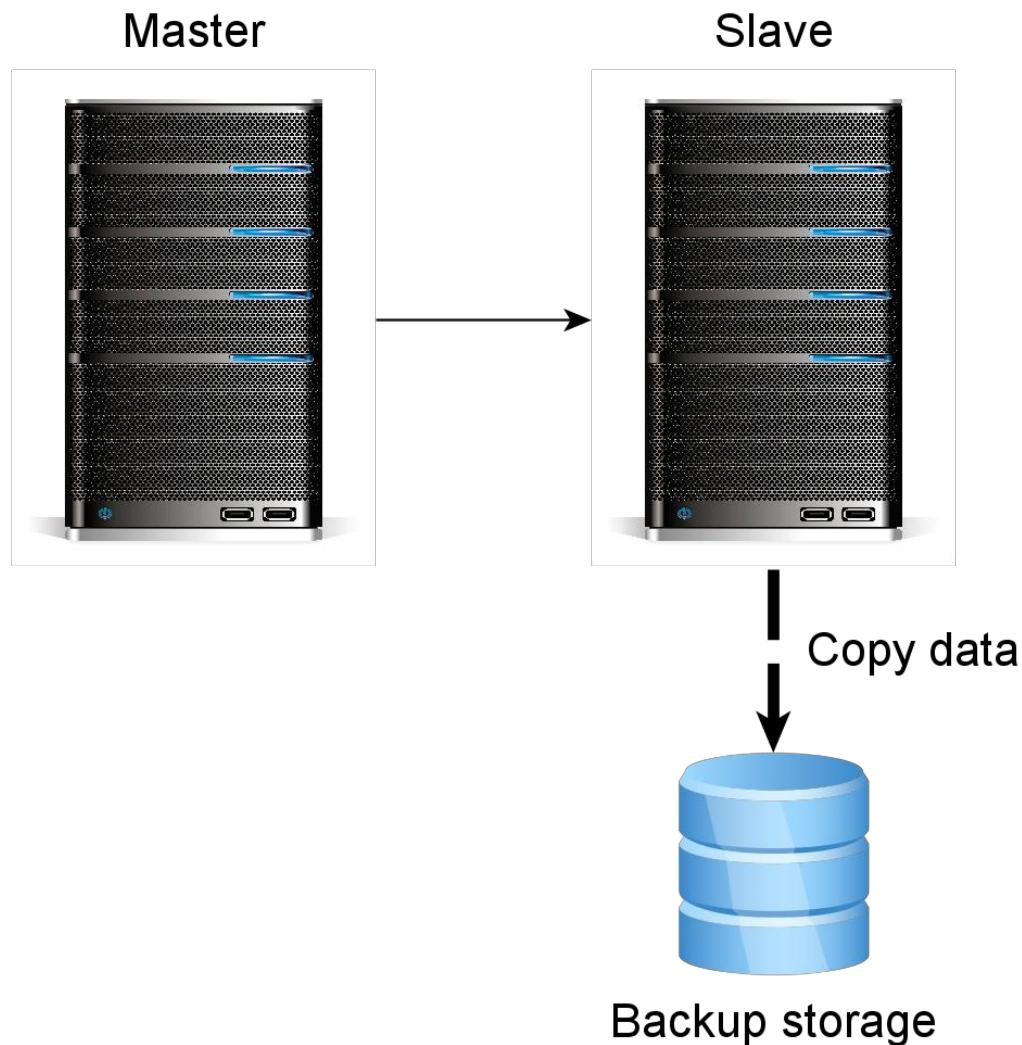
- Available only after –apply-log
- Can be used with –no-lock

mysqlbinlog

- mysqlbinlog to apply binary logs from position
- `mysqlbinlog -start-position=68212201 binlog.000001`
- `mysqlbinlog --stop-datetime=`
 - To stop at exact time

Slave topics

Slave backup



xtrabackup_slave_info

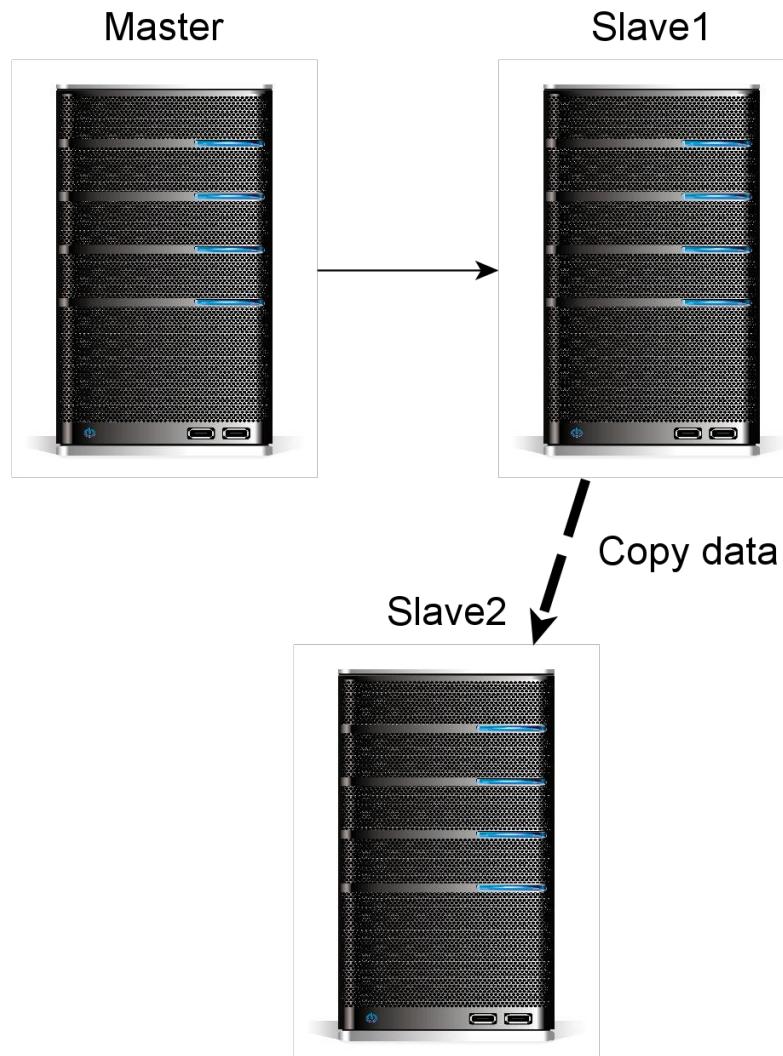
- innobackupex -slave-info
- Master binary log position for slave
- Result of SHOW SLAVE STATUS

```
CHANGE MASTER TO  
MASTER_LOG_FILE='binlog.000001',  
MASTER_LOG_POS=68212201
```

xtrabackup_slave_info

- Used for PITR when we backup from slave
- Slave cloning

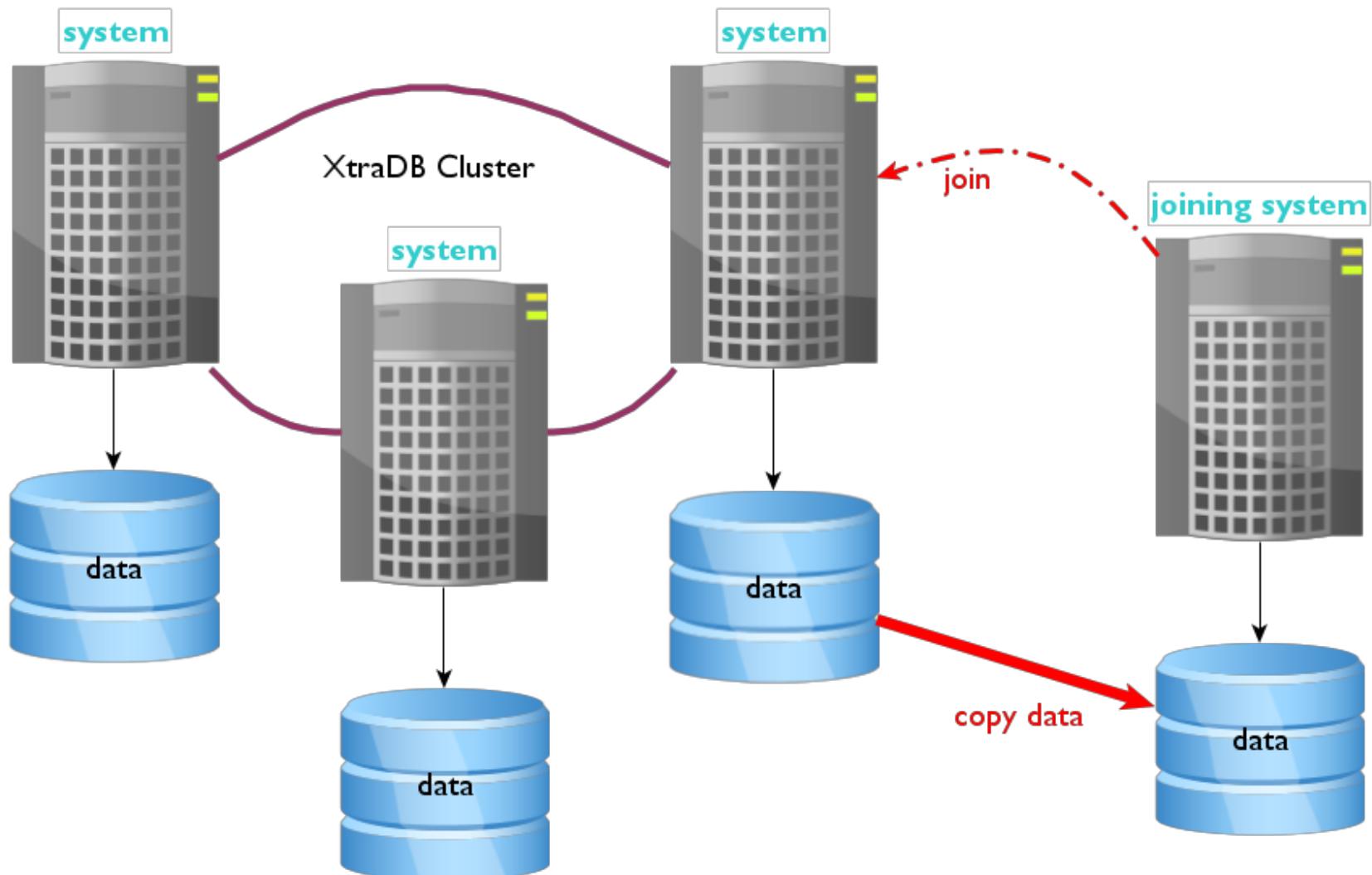
Slave cloning



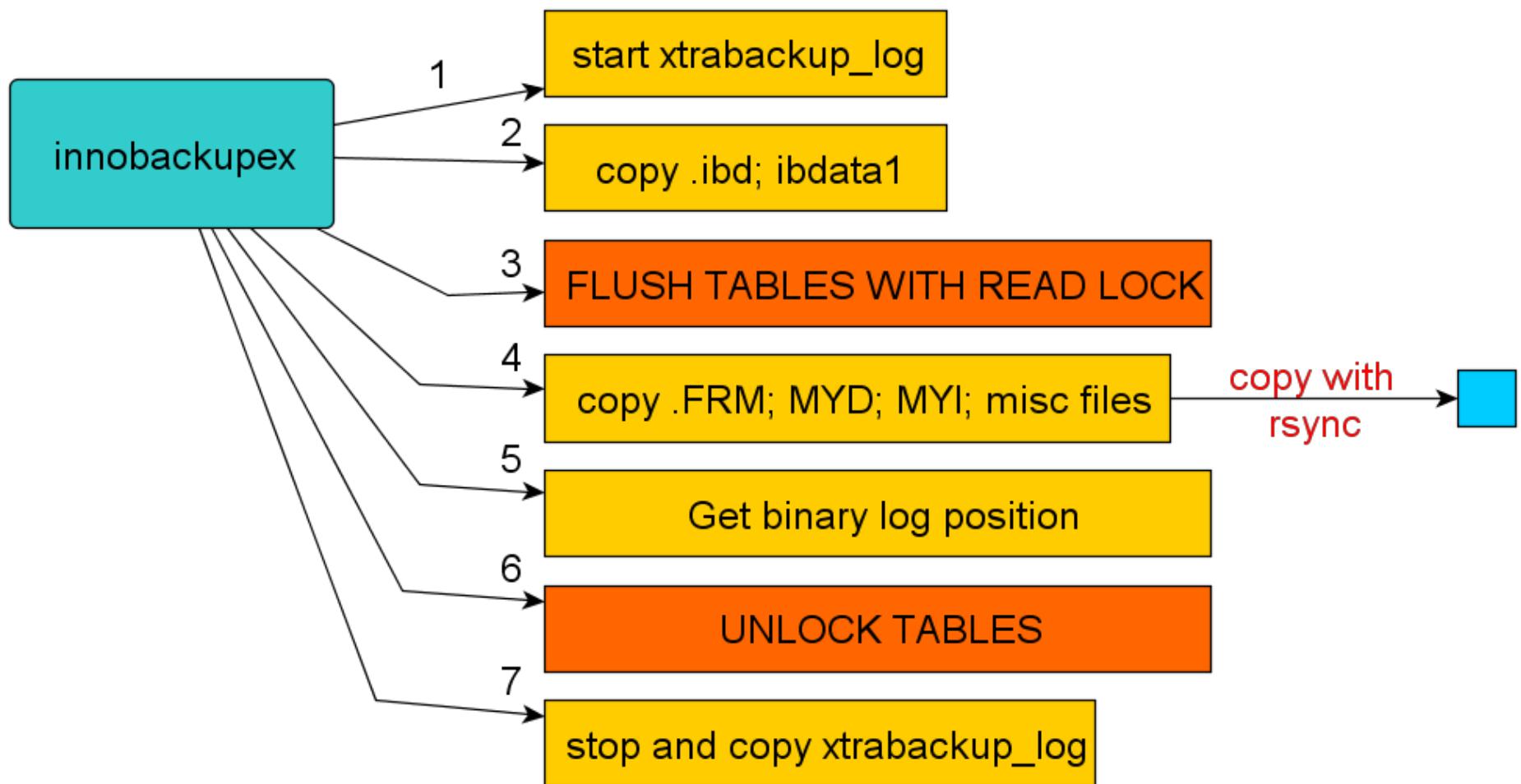
Using streaming

- `innobackupex --stream=tar /tmp --slave-info | ssh user@NEWSLAVE "tar xfi - -C /DESTDIR"`
- `innobackupex --apply-log /DESTDIR`
- Execute `CHANGE MASTER` from `xtrabackup_slave_info`

XtraDB Cluster join process.



Rsync for copying



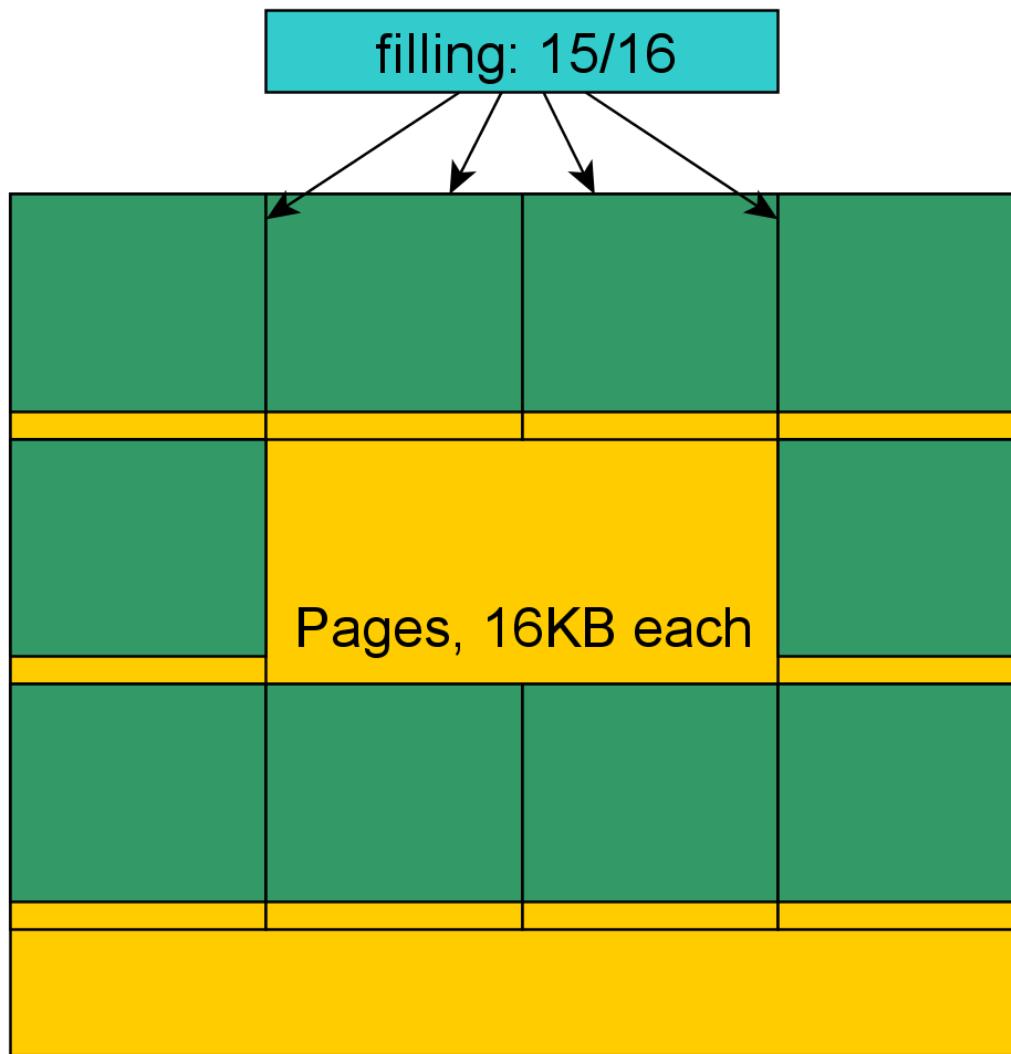
Rsync

- Innobackupex –rsync
- Two stage
 - 1. Rsync before FLUSH TABLES WITH READ LOCK
 - 2. Rsync changes inside FLUSH TABLES WITH READ LOCK
- Can't be used with –stream

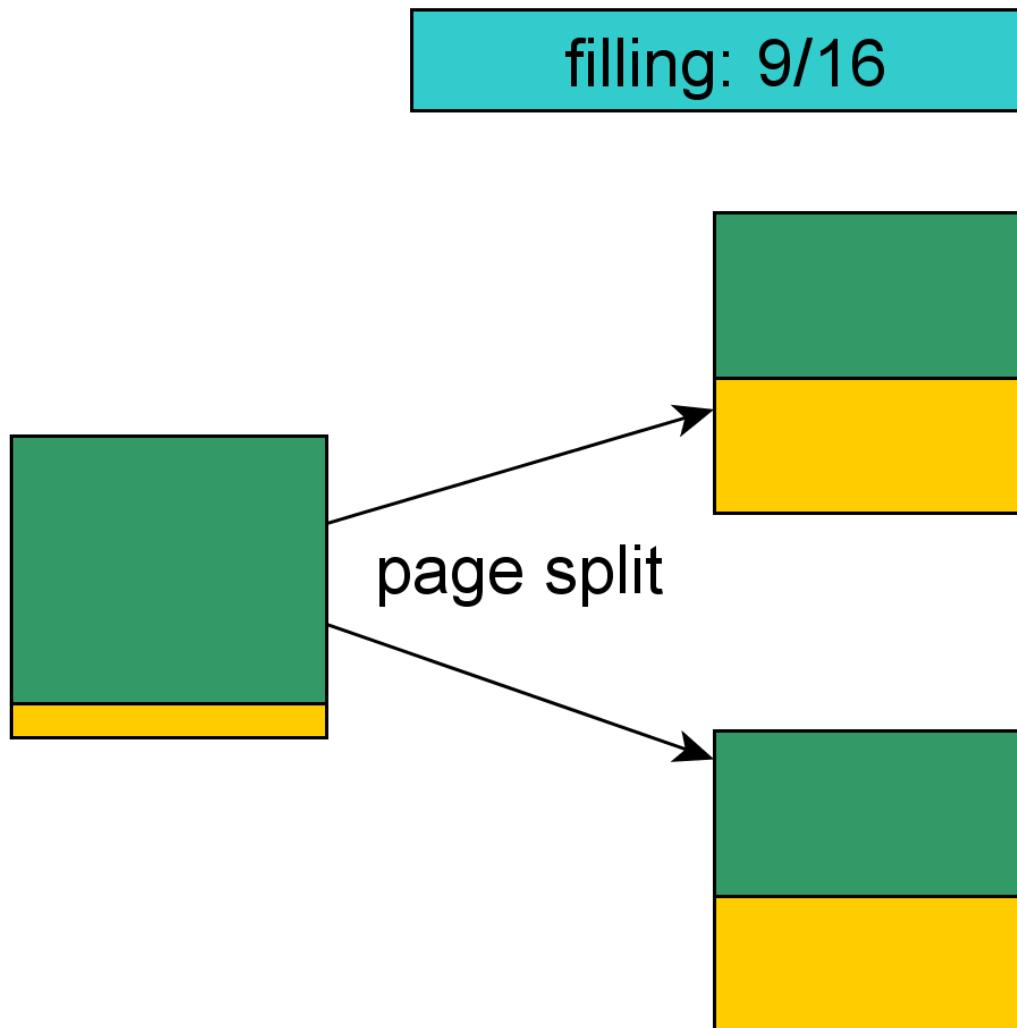
Rsync output

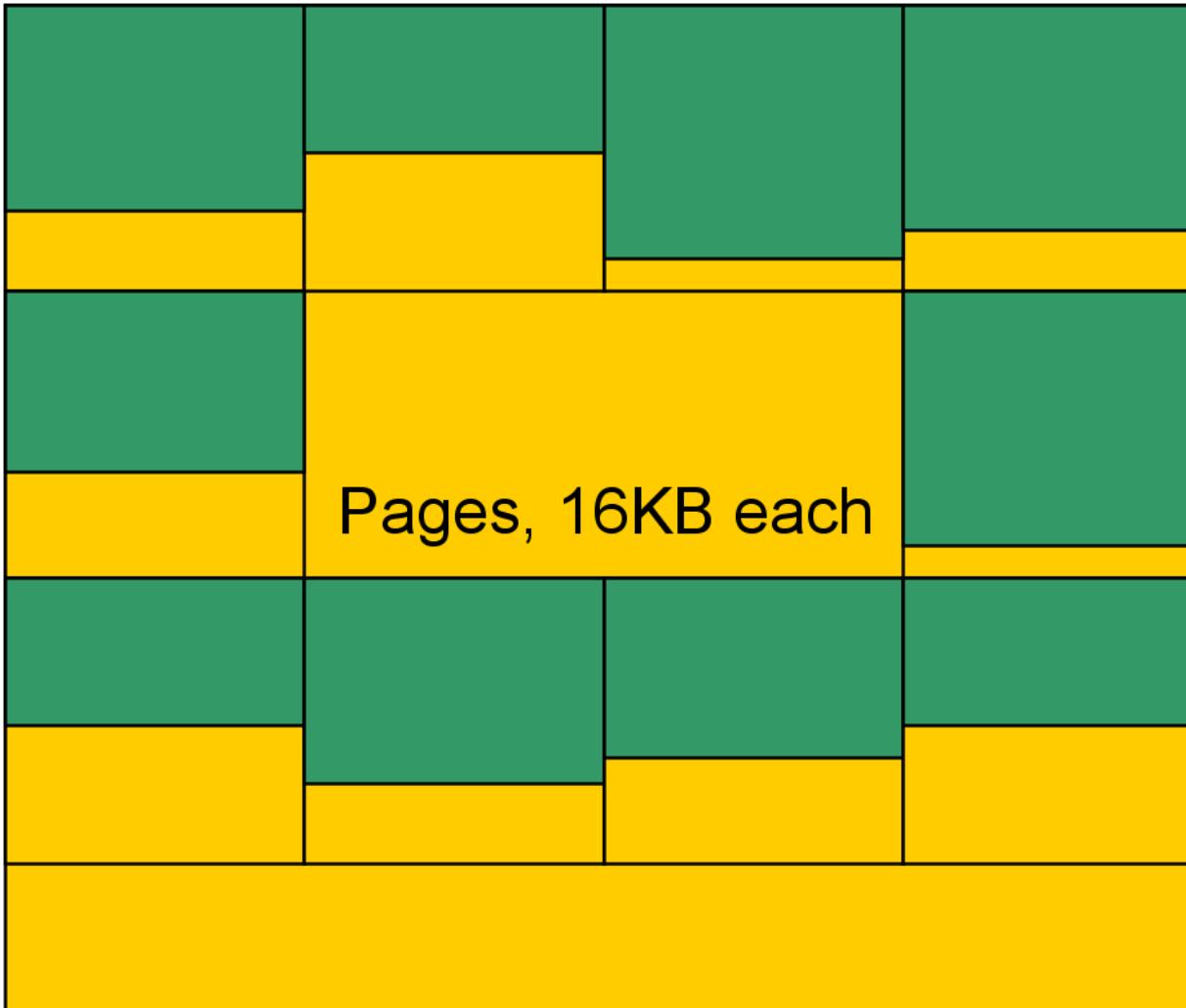
- ... File xtrabackup-rsync output

Statistics



Random inserts – page splits





Xtrabackup --stats

```
table: tpcc/order_line, index: PRIMARY, space id: 25, root
page: 3, zip size: 0
estimated statistics in dictionary:
    key vals: 32471816, leaf pages: 264267, size pages: 302592
real statistics:
    level 2 pages: pages=1, data=8406 bytes, data/pages=51%
    level 1 pages: pages=467, data=4756806 bytes,
data/pages=62%
    leaf pages: recs=35116662, pages=264267,
data=2410713870 bytes, data/pages=55%
```

Optimize?

- OPTIMIZE TABLE for Primary Key (data)
- DROP KEY / CREATE KEY (via Fast Index Creation) for Secondary Keys

Backup schemas

Daily full backups

- /data/backup/2012-04-01
- /data/backup/2012-04-02
- /data/backup/2012-04-03
- /data/backup/2012-04-04
- /data/backup/2012-04-05
- ...
- As many copies as you need

Daily full backups + binary logs

- /data/backup/2012-04-01
- /data/backup/2012-04-02
- /data/backup/2012-04-03
- /data/backup/2012-04-04
- /data/backup/2012-04-05
- ...
- As many copies as you need
- **+ backup of binary logs**

Daily backup with local copy

- /data/backup/2012-04-01
- /data/backup/2012-04-02
- /data/backup/2012-04-03
- /data/backup/2012-04-04
- /data/backup/2012-04-05
- ...
- As many copies as you need
- **The latest copy is stored also locally**
 - Compressed

Incremental into to the mix

- For low rate of changes. ~10%

Schema

- Full /data/backup/2012-04-01
- Incremental /data/backup/2012-04-02
- Incremental /data/backup/2012-04-03
- Incremental /data/backup/2012-04-04
- Full /data/backup/2012-04-05
- ...

We do not provide scripts

DIY or use third-party tools

Third-party tools

- Zmanda
- XtraBackup manager

Resources

- Facebook hybrid incremental backup
 - http://www.facebook.com/note.php?note_id=10150098033318920

Thank you!

Questions ?

vadim@percona.com

New features

- XtraBackup 2.0
 - streaming incremental backups
 - parallel compression
 - xbstream
 - LRU dump backups
- Future releases
 - true incremental backups
 - log archiving
 - compact backups
 - merged innobackupex + xtrabackup

Streaming incremental backups

Problem: send an incremental backup to a remote host:

```
innobackupex --stream=tar  
              --incremental ... | ssh ...
```

- didn't work in XtraBackup 1.6
 - innobackupex used external utilities to generate TAR streams, didn't invoke xtrabackup binary
 - xtrabackup binary must be used for incremental backups to scan data files and generate deltas
- in XtraBackup 2.0:
 - xtrabackup binary can produce TAR or XBSTREAM streams on its own

Compression

- disk space is often a problem
- compression with external utilities has some serious limitations
- built-in parallel compression in XtraBackup 2.0

Compression: external utilities

- local backups:

- create a local uncompressed backup first, then gzip files
 - must have sufficient disk space on the same machine
 - data is read and written twice

- streaming backups:

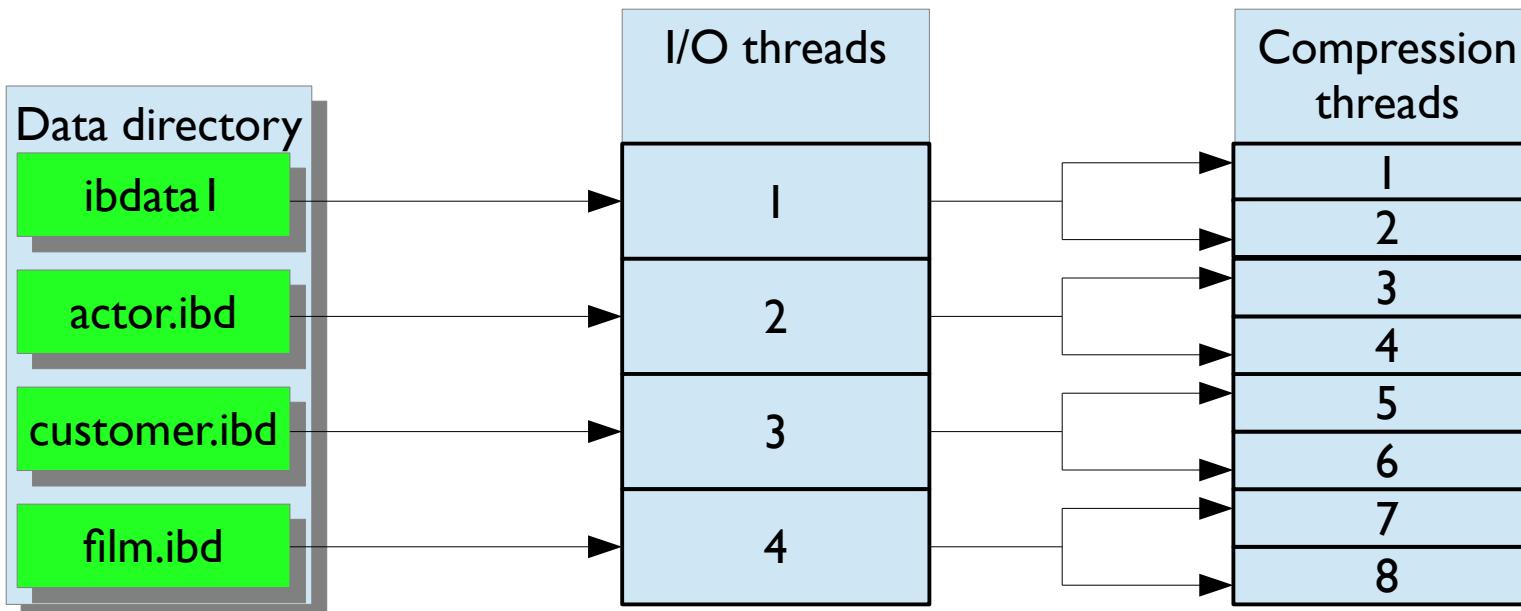
- `innobackupex --stream=tar ./ | gzip - > /data/backup.tar.gz`
 - `innobackupex --stream=tar ./ | gzip - | ssh user@host "cat - > /data/backup.tar.gz"`
 - gzip is single-threaded
 - pigz (parallel gzip) can do parallel compression, but decompression is still single-threaded
 - have to uncompress the entire .tar.gz even to restore a single table

Compression: XtraBackup 2.0

- new --compress option in both innobackupex and xtrabackup
- QuickLZ compression algorithm: <http://www.quicklz.com/>
 - “the world's fastest compression library, reaching 308 Mbyte/s per core”
 - combines excellent speed with decent compression (8x in tests)
 - more algorithms (gzip, bzip2) will be added later
- qpress archive format (the native QuickLZ file format)
- each data file becomes a single-threaded .qp archive
 - no need to uncompress entire backup to restore a single table as with .tar.gz

Compression: XtraBackup 2.0

- parallel! --compress-threads=N
- can be used together with parallel file copying:
`xtrabackup --backup --parallel=4
--compress --compress-threads=8`

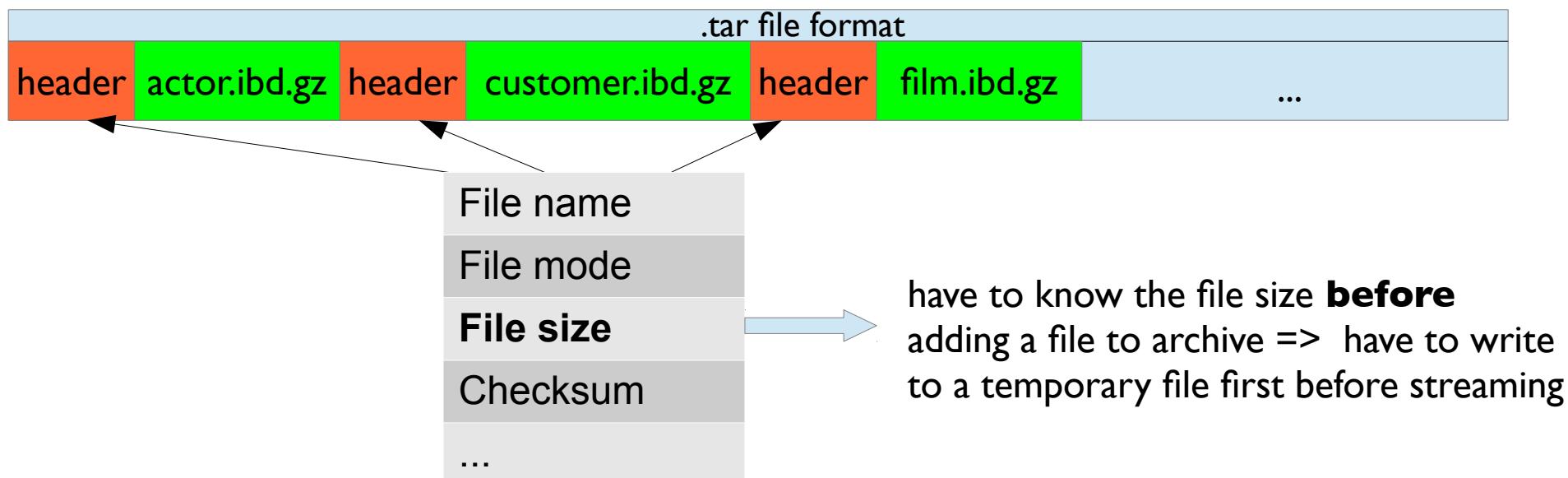


Compression: XtraBackup 2.0

- qpress file archiver
 - with --compress every file in the backup is a single-file qpress archive (e.g. customer.ibd.qp)
 - available from <http://www.quicklz.com/>
 - multi-threaded decompression:
 - qpress -d -TN customer.ibd.qp ./
- support for gzip/bzip2 to be added later

xbstream

- new streaming format in XB 2.0 in addition to --stream=tar
- Problems with traditional archive formats (tar, cpio, ar):
 - simultaneous compression and streaming is impossible



xbstream (XtraBackup 2.0)

- Problems with traditional archive formats:
 - no parallel streaming => can only backup single file at a time
- meant as file archiving utilities:
 - not really suitable for dynamically generated data streams like backups
 - XtraBackup 1.6 didn't support compression, parallel streaming, so using TAR streams was possible

xbstream (XtraBackup 2.0)



- the XBSTREAM format
 - files are written in chunks => can stream dynamically generated files
 - support interleaving chunks from different files => can do parallel file streaming
 - CRC32 checksums => can be optimized to use CPU instructions

xbstream (XtraBackup 2.0)

The xbstream utility

- can be used to extract files from streams produced by xtrabackup
- tar-like interface:
 - `xbstream -x [-C directory)`
read stream from the standard input and extract files to the current (or specified with -C) directory

xbstream (XtraBackup 2.0)

xbstream usage example

```
$ innobackupex --parallel=8  
--compress --compress-threads=4  
--stream=xbstream ./ |  
ssh user@host "xbstream -x -C /data/backup"
```

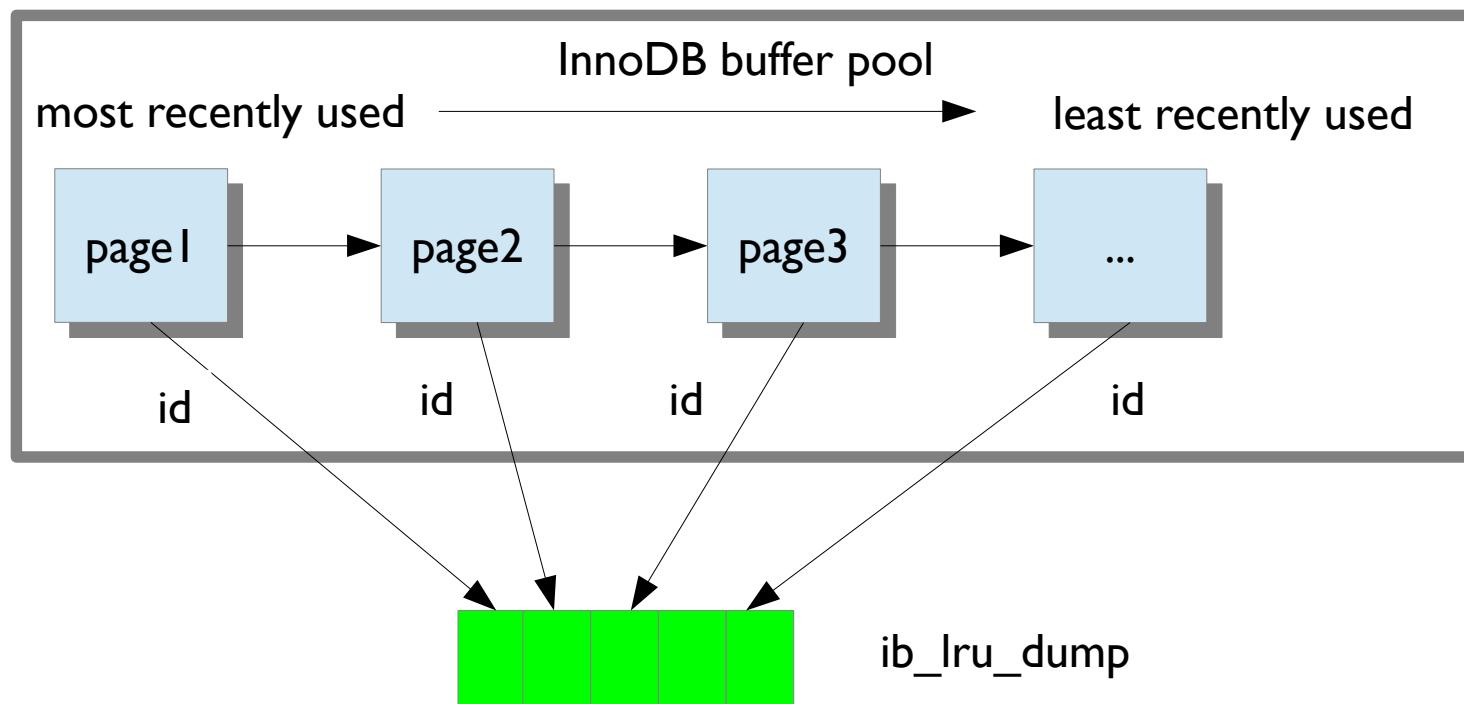
...

```
xtrabackup: Starting 8 threads for parallel data files  
transfer
```

```
[01] Compressing and streaming ./ibdata1  
[02] Compressing and streaming ./sakila/actor.ibd  
[03] Compressing and streaming ./sakila/address.ibd  
[04] Compressing and streaming ./sakila/category.ibd  
[05] Compressing and streaming ./sakila/city.ibd  
[06] Compressing and streaming ./sakila/country.ibd  
[07] Compressing and streaming ./sakila/customer.ibd
```

LRU dump backup (XtraBackup 2.0)

LRU dumps in Percona Server:



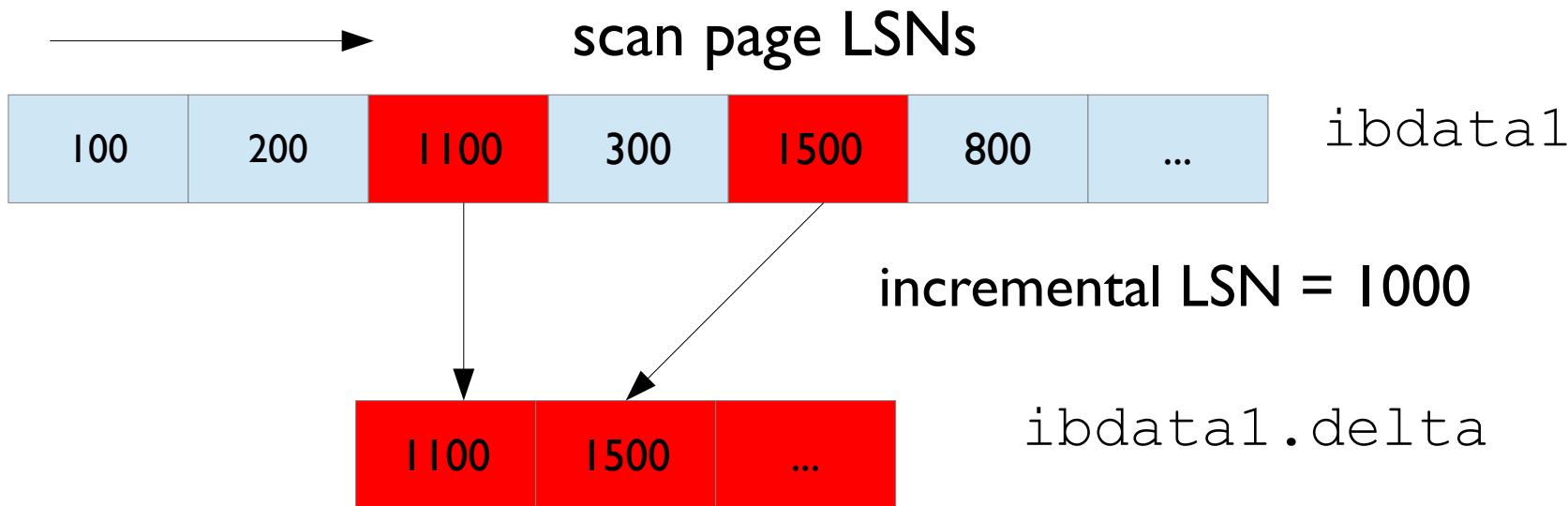
- Reduced warmup time by restoring buffer pool state from `ib_lru_dump` after restart

LRU dump backup (XtraBackup 2.0)

- XtraBackup 2.0 discovers `ib_lru_dump` and backs it up automatically
 - buffer pool is in the warm state after restoring from a backup!
 - make sure to enable buffer pool restore in `my.cnf` after restoring on a different server
 - `innodb_auto_lru_dump=1` (**PS 5.1**)
 - `innodb_buffer_pool_restore_at_startup=1` (**PS 5.5**)

True incremental backups (XtraBackup 2.x)

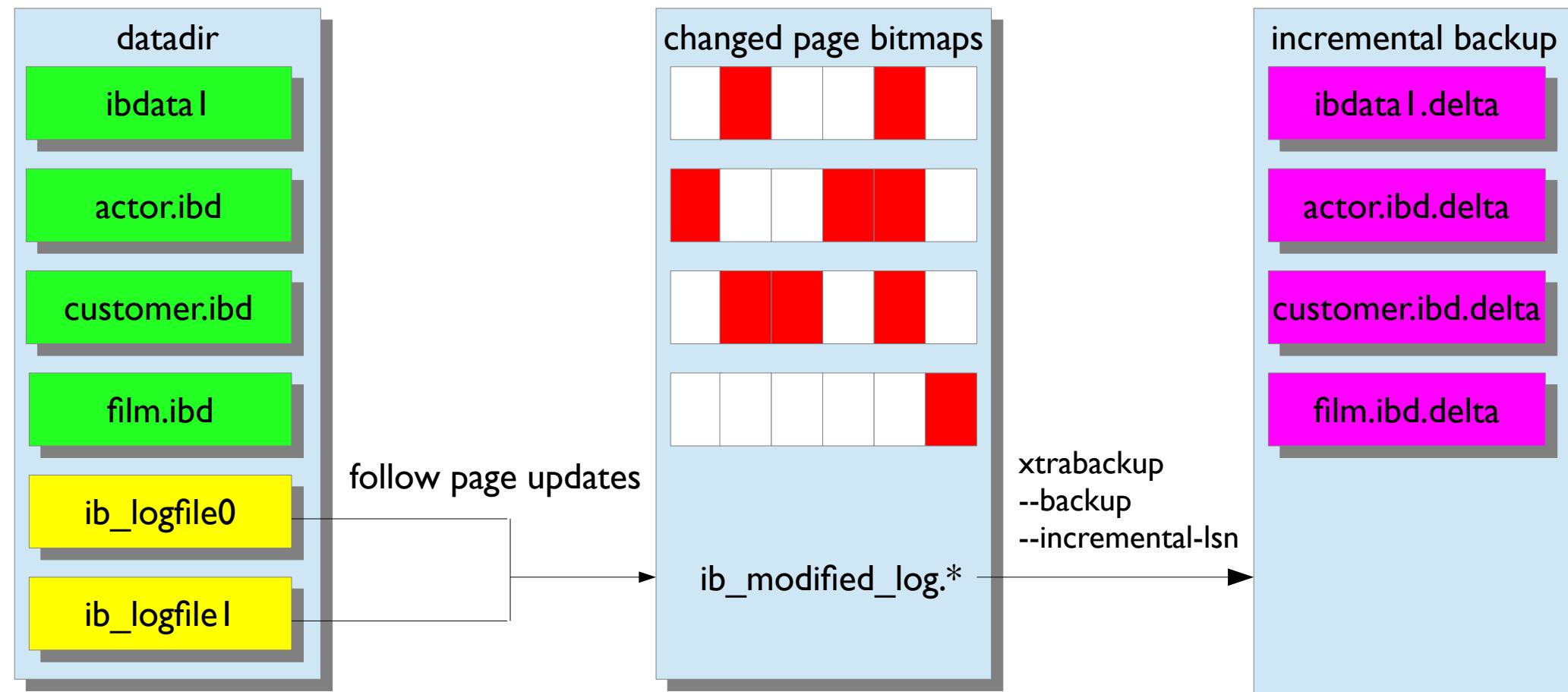
- slow incremental backups



- 2 solutions in future releases:
 - changed page tracking
 - log archiving

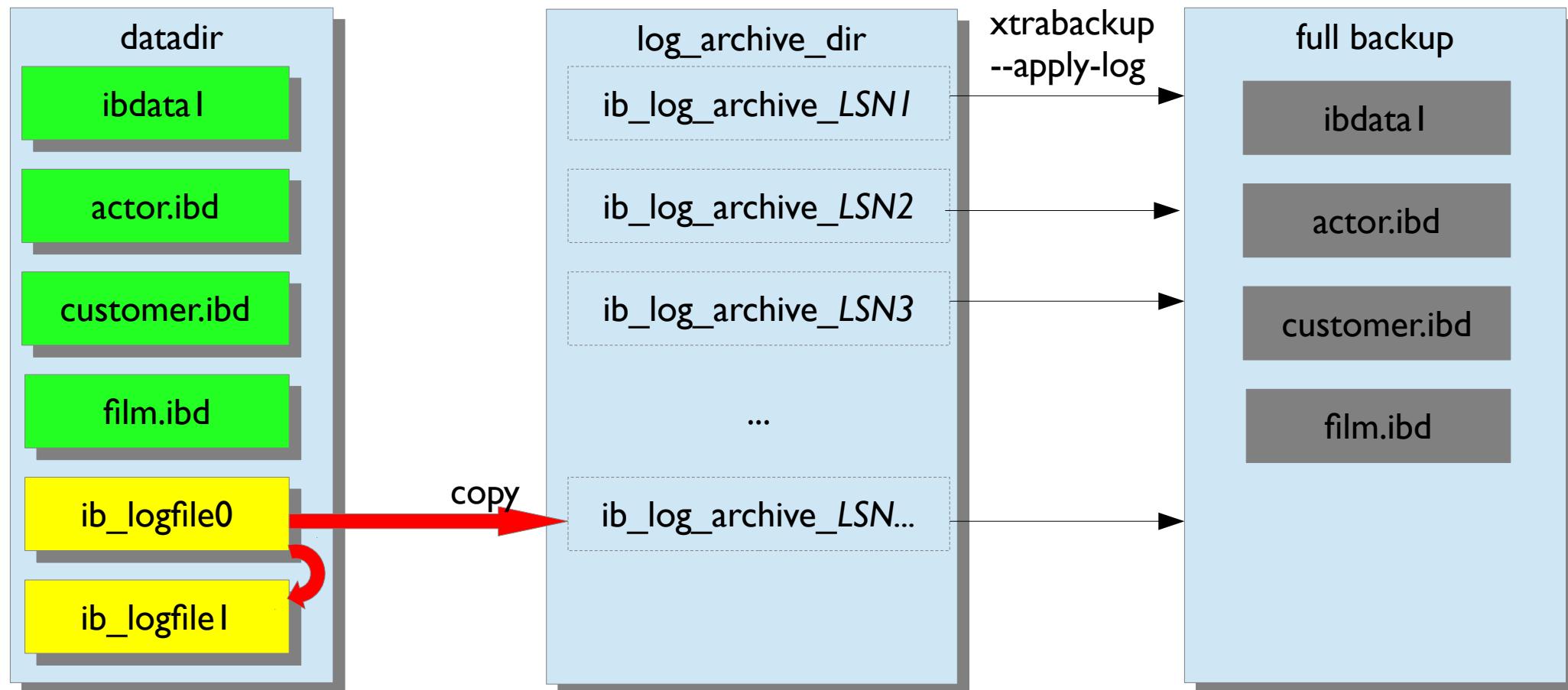
True incremental backups (XtraBackup 2.x)

- changed page tracking (new feature in Percona Server)



True incremental backups: log archiving

- log archiving (a new feature in Percona Server)

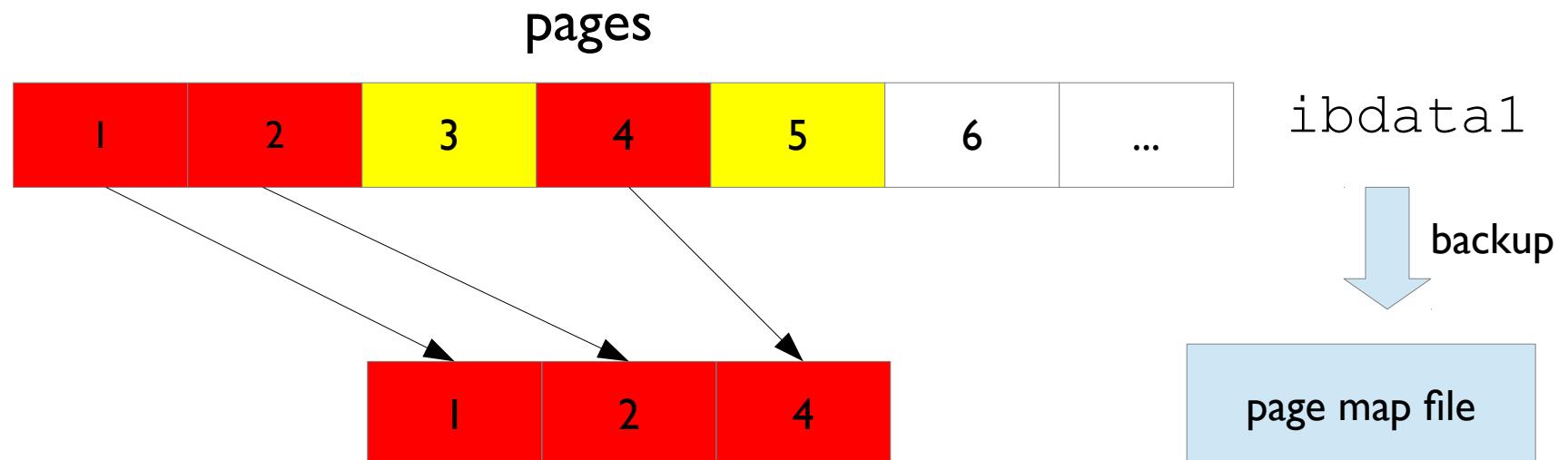


True incremental backups: which one should I use?

- “old” incremental backups:
 - standard MySQL server (w/o Percona Server features)
- changed page tracking:
 - less data to store/transfer (only changed page bitmaps)
 - faster recovery
 - updates full backup to a certain point in time
 - => more suitable for static backups
- log archiving:
 - more data to store/transfer (all changes / log updates)
 - longer recovery
 - updates full backup to an arbitrary point in time
 - => more suitable for “stand-by copy” of the main server

Compact backups

Idea: skip certain pages on backup, recreate on prepare



Legend:



primary key page

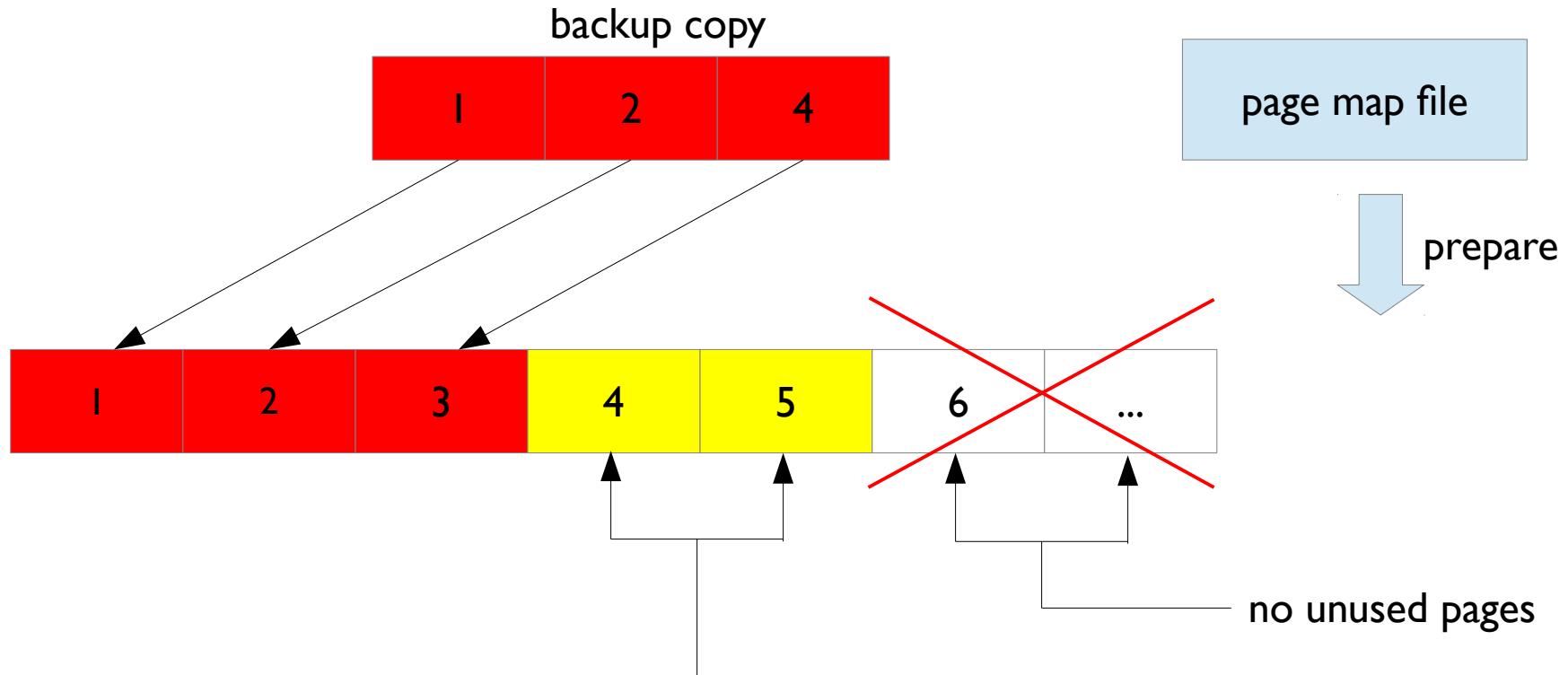


secondary key page



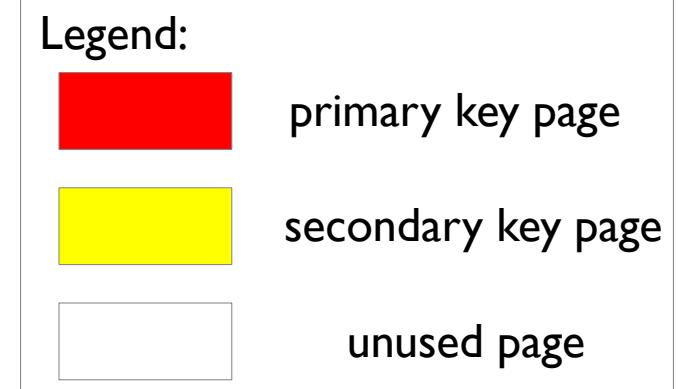
unused page

Compact backups



On prepare:

- recreate secondary keys
- fix PK page offsets
- use page map to translate page offsets when applying log records
- update pointers to root pages in data dictionary



Compact backups

- Pros:
 - smaller backups (sometimes a lot!)
 - reclaim unused space (a very old problem, MySQL bug #1341)
- Cons:
 - --prepare takes much longer

Merging innobackupex and xtrabackup

- problems with innobackupex:
 - legacy code inherited from InnoDB Hot Backup tool
 - duplicates features from xtrabackup binary
 - non-portable
- the plan is to rewrite it in C and merge with the xtrabackup binary (XtraBackup 2.x/3.0)

XtraBackup 2.x

- changed page tracking
- log archiving
- compact backups
- obsoleted innobackupex

Resources, further reading & feedback

- XtraBackup documentation:

<http://www.percona.com/doc/percona-xtrabackup/>

- Downloads:

<http://www.percona.com/software/percona-xtrabackup/downloads/>

- Google Group:

<http://groups.google.com/group/percona-discussion>

- #percona IRC channel on Freenode

- Launchpad project:

<https://launchpad.net/percona-xtrabackup>

- Bug reports:

<https://bugs.launchpad.net/percona-xtrabackup>

Questions ?