

# ??I/O??????

MySQL replica

```
“ MySQL /IO  
InnoDB flush / fsync
```

-----

## ? ????????

“ ”

- Seconds\_Behind\_Master
- Relay\_Log\_Space Exec\_Master\_Log\_Pos
- IO %util
- CPU

## ? ????? MySQL ??????

### 1?? ???????????

```
SHOW SLAVE STATUS\G
```

MySQL

Seconds_Behind_Master		
Slave_IO_Running / Slave_SQL_Running		No

項目	説明	状態
Read_Master_Log_Pos / Exec_Master_Log_Pos	一致	正常 SQL 実行
Relay_Log_Space	relay log 使用量	正常範囲内

- IO スレッド → SQL スレッド (SQL\_THREAD)
- relay log スレッド → IO スレッド (IO\_THREAD)

?? ????????????

2?? ?? SQL\_THREAD ????????

```
SHOW PROCESSLIST;
```

- スレッドが "Update", "Insert" を実行中 (SQL スレッド)
- スレッドが "Waiting for semi-sync ACK" 状態 (SQL スレッド)
- スレッドが "Waiting for master to send event" 状態 (SQL スレッド)
- スレッドが "Reading event from the relay log" 状態 (IO スレッド)

? ???????? I/O ??????

3?? ?? iostat ????????

```
iostat -dxm 1 10
```

- %util > 90% (HDD 使用率)
- w\_await > 5~10ms (書き込み待ち時間)
- r/s, w/s (I/O 回数)

```

" sdb %util 95% w_await=2.5ms I/O
  HDD

```

## 4?? ?? iotop ??????

```
iotop -oPa
```

□□□□

```
mysqld --defaults-file=/etc/my.cnf  
[jbd2/sdb1-8]
```

→ □□ MySQL □□□□□□□□ + □□□□ journaling □□□□

---

## ? ??InnoDB ??????

## 5?? ?? InnoDB ??????

```
SHOW ENGINE INNODB STATUS\G
```

□□□□□

- **Log sequence number / Log flushed up to**
    - □□□□□□□□□□ flush □□□□
  - **Pending writes / OS file reads/writes**
    - □ pending writes □□□□□□□□□□
  - **Buffer pool hit rate**
    - □□□ 99%□□□□□□□□□□ I/O□
- 

## ? ??????????????

## 6?? ??????????????

```
mount | grep sdb
```

```
relatime, data=ordered → □□□□□□□□
```

□□□□□

```
noatime,nodiratime,data=writeback
```

# 7?? ???????

```
lsblk -d -o NAME,ROTA
```

- ROTA=1 →
- ROTA=0 → SSD

HDD  relay log replay  fsync

# ? ??????????????

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<code>innodb_flush_log_at_trx_commit</code>	2	<input type="checkbox"/> I/O
<input type="checkbox"/>	<code>sync_binlog</code>	0	<input type="checkbox"/> fsync
InnoDB <input type="checkbox"/>	<code>innodb_buffer_pool_size</code>	<input type="checkbox"/> 60%-70%	<input type="checkbox"/> I/O
InnoDB <input type="checkbox"/>	<code>innodb_log_file_size</code>	≥512M	<input type="checkbox"/> checkpoint flush
<input type="checkbox"/>	ext4	<code>noatime,nodiratime,data=writeback</code>	<input type="checkbox"/> metadata <input type="checkbox"/>
I/O <input type="checkbox"/>	<code>mq-deadline</code>		<input type="checkbox"/> I/O
<input type="checkbox"/>	<code>slave_parallel_workers</code>	≥4 MySQL5.7+	<input type="checkbox"/>

# ? ??????????????

<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<code>SHOW SLAVE STATUS\G</code>
<input type="checkbox"/> I/O <input type="checkbox"/>	<code>iostat -dxm 1</code>
<input type="checkbox"/>	<code>iotop -oPa</code>
<input type="checkbox"/> InnoDB flush <input type="checkbox"/>	<code>SHOW ENGINE INNODB STATUS\G</code>
<input type="checkbox"/>	<code>cat /sys/block/sdb/queue/scheduler</code>
<input type="checkbox"/>	<code>SHOW GLOBAL STATUS LIKE 'Innodb_buffer_pool%';</code>

# ? ????????

??	??????	????
MySQL????	□	SHOW SLAVE STATUS
?? I/O	????? %util=95%??	iostat / iotop
InnoDB flush	□	SHOW ENGINE INNODB STATUS
??	??? IO????? relay???	—

?? ?????

???? I/O ?????????????? ext4 journaling ????????? MySQL  
 ?????????????? → ??????????????

????? iostat ??????????????  
 ????????????????????? sdb????????????????  
 ?????????????

# ? ?????????????? sdb?

??	??	???	??
%util	?????	<b>94-96%</b>	????????? I/O ??????????
w/s	?????	□ <b>370-380</b>	???????????????? binlog ???
wMB/s	?????	□ <b>2.8-3.0 MB/s</b>	???????????????? ???????????????? □ SATA HDD□
w_await	?????????	<b>2.5-2.7 ms</b>	???????????????? ????????????
wrqm/s	?????	□ <b>300/s</b>	???????????????? ?????????
aqu-sz	?????????	□ <b>1.0</b>	????????????

?? ???

“ sdb ?? I/O ?????? MySQL ?????? relay log ??? I/O  
 ?????????????

?? ?????????????????

1?? ?? MySQL ????????

□□□

```
SHOW SLAVE STATUS\G
```

□□□□□

- Seconds\_Behind\_Master --- □□□□
  - Slave\_IO\_Running / Slave\_SQL\_Running --- □□□□ Yes
  - Relay\_Log\_Space --- relay log □□□□□□□□
- IO□□□□□ SQL□□□□□ □□□□□□□□□□ I/O □□□□□□□□

2?? ?? MySQL ? InnoDB ?????

```
SHOW ENGINE INNODB STATUS\G
```

□□□□□□□

- Log sequence number □ Last checkpoint at □□□□□□ → □□□□□□
- I/O thread □ pending writes □□□□
- Buffer pool □□□□ Dirty pages □□□□□□

3?? ?????? I/O ??

□□□□□□□□□□

```
fio --name=write_test --filename=/home/shuncom/mysql/data/testfile --size=2G --bs=16k --rw=write --ioengine=libaio --direct=1 --numjobs=1 --runtime=60 --group_reporting
```

□□□□□□□□□□ MB/s □ IOPS□□□□□□□□□□□□

? ??????????????

? 1. MySQL ?????

□ /etc/mysql/mysql.conf.d/mysqld.cnf □ /etc/my.cnf □



? ??????????

Item	Value	Notes
Slave_IO_Running	Yes	binlog
Slave_SQL_Running	Yes	SQL relay log
Slave_SQL_Running_State	waiting for handler commit	SQL commit fsync
Seconds_Behind_Master	5,495,423 (~63)	SQL
Relay_Log_Space	618,787,289,903 bytes (~576 GB)	Relay log 576GB
Read_Master_Log_Pos vs Exec_Master_Log_Pos	50GB+	SQL
SQL_Delay	3600	1
Slave_SQL_Running_State	waiting for handler commit	I/O fsync InnoDB flush

? ??????????

```

“ I/O
SQL SQL_THREAD
“waiting for handler commit” 100% InnoDB + fsync
  
```

|||||

```

“ binlog
apply relay log
  
```

?? ?????????? iostat?

- `iostat` `sdb` `%util≈95%` → `████████`
- `wMB/s≈2.8~3MB/s` `████████` SSD `████` MB/s`██`
- `[jbd2/sdb1-8]` `████` ext4 `████████`
- Relay log `██` 500+ GB`█`
- SQL\_THREAD “waiting for handler commit”`█`

`██ ██████████` `██ I/O ██████████` `████`

? ??????????????????

1`█ ██████████`

```
SHOW PROCESSLIST;
```

`██████████` `Slave_SQL` `██████████`  
`waiting for handler commit` `██`

2`██` InnoDB flush `██`

```
SHOW ENGINE INNODB STATUS\G
```

`█` “LOG” `██████████`

```
Log sequence number .....
Log flushed up to .....
Last checkpoint at .....
```

`██████████████████` MB~GB`██████████████████`

?? ??????????????????

`████` `██████` `██████████████████`

? 1. ??????????????????

`█` `/etc/mysql/mysql.conf.d/mysqld.cnf` `█` `/etc/my.cnf` `█`

```
[mysqld]
# 
innodb_flush_log_at_trx_commit = 2
sync_binlog = 0
innodb_doublewrite = 0
innodb_flush_method = 0_DIRECT
```



## ? 2. ????????????

□□

```
mount | grep sdb
```

□□ `/etc/fstab` □

```
/dev/sdb1 /home ext4 defaults,noatime,nodiratime,barrier=0,data=writeback 0 2
```

□□

```
mount -o remount /home
```

□□□□□ journal □□□□

## ? 3. ?????????? I/O ???

```
cat /sys/block/sdb/queue/scheduler
echo deadline > /sys/block/sdb/queue/scheduler
```

□□□ `/etc/default/grub` □

```
GRUB_CMDLINE_LINUX_DEFAULT="quiet splash elevator=deadline"
update-grub && reboot
```

## ? 4. ?? relay log ???????

Relay log [ ]

```
relay_log = /mnt/fastdisk/relay-log
relay_log_info_repository = TABLE
relay_log_recovery = ON
```

## ? 5. ?????????MySQL 5.7+?

```
STOP SLAVE;
SET GLOBAL slave_parallel_workers = 8;
START SLAVE;
```

[ ]

```
slave_parallel_type = LOGICAL_CLOCK
slave_parallel_workers = 8
```

“ [ ]

## ? 6. ??? HDD?????

- [ ] relay log [ ] SSD [ ]
- [ ] MySQL [ ] SSD [ ]
- [ ] innodb\_buffer\_pool\_size [ ] 60-70% [ ]

## ? ??????????

[ ]	[ ]	[ ]
[ ]	[ ]	[ ]
[ ] SQL [ ]	<b>waiting for handler commit</b>	I/O flush [ ]
Relay log [ ]	576GB	[ ]

MySQL I/O	%util ≈ 95%	SSD
MySQL	fsync	MySQL

MySQL

MySQL `shell + SQL`

- MySQL
- I/O
- relay log
- InnoDB flush
- MySQL

MySQL `InnoDB I/O` + replication SQL

thread fsync

??????????

MySQL	MySQL	MySQL
Seconds_Behind_Master: 5495423	63	I/O SQL
Slave_SQL_Running_State: waiting for handler commit	SQL fsync	I/O
Pending flushes (fsync) log: 1; buffer pool: 97886	fsync 97,886	I/O
OS file writes: 8,575,785,215		
Log flushed up to Added dirty pages up to	log buffer flush	redo
jbd2/sdb1-8 I/O	ext4	

??????????

MySQL SQL

- MySQL → (commit)
- InnoDB fsync() redo log





```
/dev/sdb1 on /home type ext4 (rw,nosuid,nodev,relatime,data=ordered)
```

```
data=ordered [ ] [ ] [ ] [ ] ext4[ ] [ ] [ ] [ ] commit[ ] [ ] [ ] flush[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]
```

```
data=writeback,noatime,nodiratime,barrier=0
```

```
[ ] [ ] [ ]
```

```
mount -o remount,noatime,nodiratime,data=writeback,barrier=0 /home
```

```
“△ [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] UPS [ ] [ ] barrier=0 [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]
```

## ? 5. ??????????????????????

```
[ ] [ ] [ ] [ ] [ ] MySQL [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] SQL [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] backlog [ ]
```

```
STOP SLAVE SQL_THREAD;
-- [ ]I/0 [ ] [ ] [ ] [ ] [ ]
START SLAVE SQL_THREAD;
```

## ? ??????????????????

```
[ ] [ ] [ ] [ ] [ ]
```

```
SHOW SLAVE STATUS\G
SHOW ENGINE INNODB STATUS\G
```

```
[ ] [ ] [ ]
```

- Seconds\_Behind\_Master [ ] [ ]
- Slave\_SQL\_Running\_State [ ] [ ] [ ] *waiting for handler commit*
- Pending flushes [ ] [ ] [ ] [ ] [ ] [ ]

## ? ??????????????????

OS	DB	IO
MySQL	MySQL SSD	I/O
innodb_flush_log_at_trx_commit=2 + sync_binlog=0		
mount data=writeback,noatime		ext4 journal
binlog		fsync

```

/home sdb1 + my.cnf

```

```

iostat -x 1 10 MySQL " "

```

???????

Device: sdb

Device: sdb		
w/s = 380~394		
wkB/s ≈ 2900~3100		
%util ≈ 90~94%		
w_await ≈ 2.4ms		
%iowait ≈ 18~22%		

Device: sdb

OS	DB	IO
%util ≈ 90~94%		100% CPU " "
%iowait ≈ 20%	CPU IO	mysqld IO
w_await ≈ 2.4ms		2.4ms w/s 380+ IO
wrqm/s 300 %wrqm ≈ 44%		

# ????????????

MySQL 数据库 “ ”

1. **relay log** SQL binlog relay log
2. **SQL** **innodb**
3. **binlog** **log-bin**
4. **redo log/fsync** `innodb_flush_log_at_trx_commit` 1

# ?? ??????????

# 1?? ?? MySQL ?????

???

```
SHOW GLOBAL STATUS LIKE 'Innodb_%writes%';
SHOW GLOBAL STATUS LIKE 'Innodb_data%';
SHOW GLOBAL STATUS LIKE 'Innodb_os_log%';
SHOW GLOBAL STATUS LIKE 'Slave_running%';
SHOW SLAVE STATUS\G
```

????

- `Innodb_data_pending_writes` ???? ?
- `Innodb_data_fsyncs` ???? ?
- `Seconds_Behind_Master` ???? ?
- `Slave_SQL_Running_State` ????????? ?

# 2?? ?????????????

???

```
lsblk -o NAME,ROTA,MOUNTPOINT
mount | grep mysql
```

- `ROTA=1` `XXXXXXXXXXXX`
- `MySQL` `XXXXXXXXXXXX` `barrier` `noatime` `XXX`
  - `XXXX` `/etc/fstab` `XXXX`

```
defaults,noatime,nodiratime,barrier=0
```

`barrier=0` `UPS` `SSD` `XXXXXXXXXXXX`

## 3?? ?? InnoDB ???

`XXX`

```
SHOW VARIABLES LIKE 'innodb_flush_log_at_trx_commit';
SHOW VARIABLES LIKE 'sync_binlog';
SHOW VARIABLES LIKE 'innodb_flush_method';
```

- `innodb_flush_log_at_trx_commit=1` `XXXXXXXX` `fsync` `XXXXXX`
- `XXXX`

```
SET GLOBAL innodb_flush_log_at_trx_commit=2;
```

`my.cnf`

```
[mysqld]
innodb_flush_log_at_trx_commit=2
sync_binlog=0
```

`XXXXXXXXXXXXXXXXXXXXXXXXXXXX`

## 4?? ?? relay log ????

`relay log` `XXXXXXXX` `/home/shuncom/mysql/data` `XXXX` `relay log` `XXXX`  
`SSD`

```
relay_log=/data/relay/relay-bin
```

????????????

<code>XXX</code>	<code>XXXX</code>
------------------	-------------------

SSD	SSD NVMe sdb relay log/data
InnoDB	innodb_flush_log_at_trx_commit=2 sync_binlog=0
	mount noatime,nodiratime,barrier=0
	binlog log_bin=0
IO	deadline noop SSD noop

```
echo deadline > /sys/block/sdb/queue/scheduler
```

??????????

SSD

```
# 1. SSD
lsblk -o NAME,ROTA,MOUNTPOINT

# 2. IO
cat /sys/block/sdb/queue/scheduler

# 3.
mount | grep mysql

# 4. InnoDB
mysql -e "SHOW VARIABLES LIKE 'innodb_flush_log_at_trx_commit';"
mysql -e "SHOW VARIABLES LIKE 'sync_binlog';"

# 5. Slave
mysql -e "SHOW SLAVE STATUS\G" | egrep "Seconds_Behind_Master|Running_State"
```

SSD

SSD

- SHOW VARIABLES LIKE 'innodb\_%';
- mount | grep mysql
- lsblk -o NAME,ROTA,SIZE,MOUNTPOINT
- /etc/mysql/my.cnf MySQL [mysqld]

SSD

MySQL IO

MySQL  
Seconds\_Behind\_Master = 5495423 63  
SQL  
SQL

? ??????????

- Slave\_IO\_Running: Yes Slave\_SQL\_Running: Yes →
- Slave\_SQL\_Running\_State: waiting for handler commit → SQL
- innodb\_flush\_log\_at\_trx\_commit = 2
- sync\_binlog = 1 binlog
- innodb\_flush\_method = 0\_DIRECT → OS OK
- Innodb\_data\_pending\_fsyncs = 97889 → △

“ I/O fsync redo log/fsync

? ?????????? SQL ????? commit

? ????? waiting for handler commit ???

InnoDB redo log fsync

- IOPS — SATA HDD
- flush
- MySQL redo
- sync\_binlog=1 I/O
- SQL relay log I/O

? ??????????

1?? ?????IO??

```
iostat -x 1 10
```

```
|||||
```

- %util ||||| 100%|
- await ||||| 10ms|
- svctm |||||  
|||||

```
|||||
```

```
pidstat -d 1 -p $(pidof mysqld)
```

```
|| mysqld |||||
```

## 2?? ?? relay log ????????

```
du -sh /var/lib/mysql/relay-log*
```

```
|| relay log || || GB || TB|| SQL||||| IO|||
```

```
|||| SQL |||||
```

```
SHOW SLAVE STATUS\G
```

```
-- || Read_Master_Log_Pos || Exec_Master_Log_Pos ||
```

## 3?? ?? InnoDB I/O ??

```
SHOW VARIABLES LIKE 'innodb_io_capacity%';
```

```
||||| 200|||||
```

```
innodb_io_capacity=2000
```

```
innodb_io_capacity_max=4000
```

## 4?? ?? redo log ??

```
SHOW VARIABLES LIKE 'innodb_log_file_size';
```





# ??1??? PowerShell ??????????

1. `Get-PhysicalDisk`
2. `Get-PhysicalDisk`

```
Get-PhysicalDisk | Select-Object FriendlyName, MediaType, CanPool, BusType, Size
```

- `BusType` → `Sata`, `Sas`, `Usb`, `Nvme`
- `MediaType` → `HDD`, `SSD`
- `Size` → `8 TB`

Hitachi HUS728

FriendlyName	MediaType	CanPool	BusType	Size
Hitachi HUS728	HDD	False	Sas	8 TB

`Get-PhysicalDisk | Where-Object { $_.BusType -eq 'Sas' }` → SAS  
`Get-PhysicalDisk | Where-Object { $_.BusType -eq 'Sata' }` → SATA

# ??2????????/??????

1. `Get-PhysicalDisk` → `Get-PhysicalDisk`
2. `Get-PhysicalDisk` → `Get-PhysicalDisk` → `Get-PhysicalDisk` ID
3. `Get-PhysicalDisk` SAS SATA

# ??3????????

Dell HPE Lenovo RAID /

- **Dell** → OpenManage Server Administrator (OMSA)
- **HPE** → Smart Storage Administrator (SSA)
- **Lenovo** → XClarity Administrator

RAID SAS/SATA

PowerShell BusType RAID





---

### 3. ?????????

- Virtual Hard Disk Cache
  - None →  I/O
  - Host Cache →
  - Write-Through / Write-Back →
- I/O  
  - Hyper-V  **Virtualization-based I/O (VSP/VSC)**

---

### 4. ????????

1.  
  - SATA HDD →  10~200 MB/s
  - SAS HDD →  150~300 MB/s
  - SSD/NVMe →  GB/s
2. **RAID**  
  - RAID1  RAID10 →
  - RAID5 →
3.  / 
  - HBA  PERC RAID

---

### 5. ?????????

- vCPU NUMA  I/O
- I/O
- Hyper-V **Dynamic Memory**

---

### 6. ????????

1.  **VHDX**
  2.  **SCSI**
  3.  **SSD**  **SAS RAID**
  4.  **vCPU**/
  5. **\*\*** Write-Back / Host Cache  **\*\***
  6.  **Hyper-V**  “ ”
-

