PERCONA

Ferramentas e métodos para análise de performance em MySQL
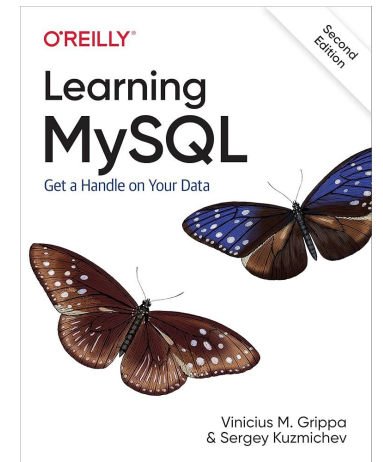
Sobre mim

Vinicius (vini) Grippa

Lead Database Engineer @ Percona

Oracle Ace Director

https://mysqlbr.com.br/

Expectativas

Essa apresentação não tem como foco:

- Ensinar os detalhes de cada ferramenta
- Explicar o funcionamento interno do MySQL
- Como resolver todos os problemas de performance com uma única solução

Foco da apresentação:

- Entender por onde começar uma análise de performance
- Apresentar ferramentas e comandos que temos à disposição
- Foco nas ferramentas de MySQL
- Exemplos

PERCONA

Agenda:

- Desafios
- Por onde começar uma análise de performance
- Entendendo o servidor
- Saturação de recursos
- Entendendo o banco de dados
- Investigando o MySQL

Troubleshooting

Desafios e como resolve-lôs

Desafios

- Muita informação (agulha num palheiro)

Solução

- Entender o problema!
- Análise sistemática

Método

- Qual problema estamos tentando resolver?
  - Perguntas básicas
- Entender a configuração do servidor e do MySQL
  - Recursos e configurações
- Achar a saturação
  - Sistema operacional
    - CPU
    - Memória
    - Disco
    - Rede
  - MySQL (análise direcionada)

Sistema Operacional

Entendendo o servidor - Percona toolkit (pt-summary)

```
# Percona Toolkit System Summary Report ######################
        Date | 2025-08-21 20:04:06 UTC (local TZ: -03 -0300)
    Hostname | XXXXX
      Uptime | 3 days, 19:25,  8 users,  load average: 6.03, 7.00, 7.63
      System | Dell Inc.; PowerEdge R7525; vNot Specified (Rack Mount Chassis)
 Service Tag | BBN2QJ3
    Platform | Linux
     Release | Ubuntu 22.04.5 LTS (jammy)
      Kernel | 5.15.0-151-generic
Architecture | CPU = 64-bit, OS = 64-bit
    Threading | NPTL 2.35
      SELinux | No SELinux detected
  Virtualized | VMWare
# Processor #############################################
   Processors | physical = 2, cores = 48, virtual = 96, hyperthreading = yes
       Speeds | 1x1472.438, 1x1489.703, 50x1500.000, 17x2000.000, 21x2650.000, 1x3464.894, 1x3466.334,
  1x3615.760, 1x362.414, 1x480.046, 1x587.042
       Models | 96xAMD EPYC 7413 24-Core Processor
```

PERCONA

Entendendo o servidor - Percona toolkit (pt-summary)

```
        Caches | 48x25344 KB
# Memory #################################################
         Total | 376.6G
          Free | 12.5G
          Used | physical = 169.7G, swap allocated = 2.0G, swap used = 2.0M, virtual = 169.7G
        Shared | 26.4G
       Buffers | 194.4G
        Caches | 178.5G
         Dirty | 16040 kB
       UsedRSS | 156.1G
    Swappiness | 1
    DirtyPolicy | 10, 10
[...]
# Disk Schedulers And Queue Size #############################
          dm-0 | 128
          dm-1 | 128
       nvme0n1 | [none] 1023
           sda | [mq-deadline] 256
           sdb | [mq-deadline] 64
[...]
# Memory mamagement ###################################
Transparent huge pages are currently disabled on the system.
# The End ###################################################
```

Antes de continuar

```
shell > dmesg -T

[....]

[Thu Nov 23 09:02:45 2023] XFS (dm-17): metadata I/O error in
"xfs_alloc_read_agfl+0x7c/0xc0 [xfs]" at daddr 0x3 len 1 error 74

[Thu Nov 23 09:02:45 2023] XFS (dm-17): page discard on page 0000000063ba1b5d, inode
0x84, offset 0.

[Thu Nov 23 09:02:45 2023] XFS (dm-17): Metadata CRC error detected at
xfs_agfl_read_verify+0xa5/0xf0 [xfs], xfs_agfl block 0x3

[Thu Nov 23 09:02:45 2023] XFS (dm-17): Unmount and run xfs_repair
```
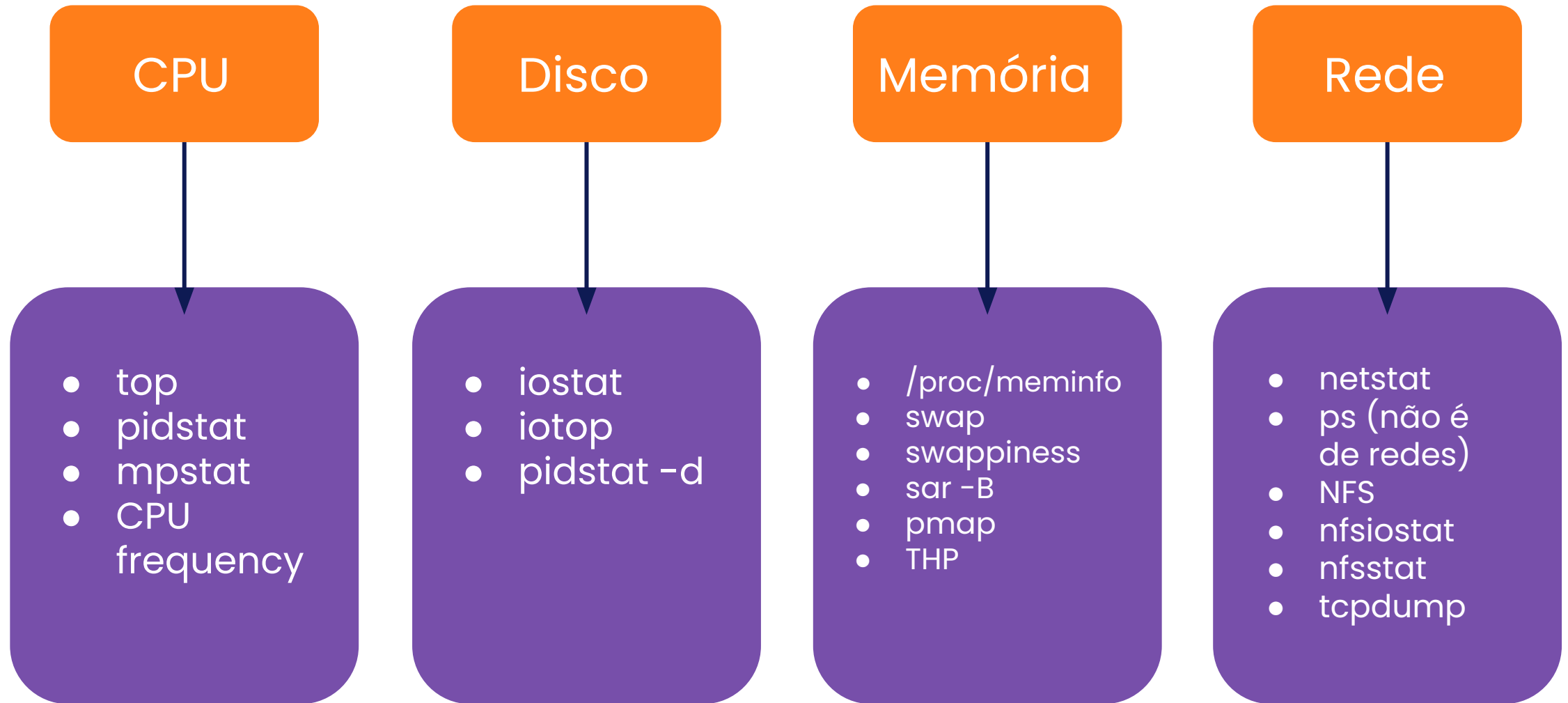
PERCONA

Encontrando a saturação

Linux Performance Observability Tools

http://www.brendangregg.com/linuxperf.html 2014
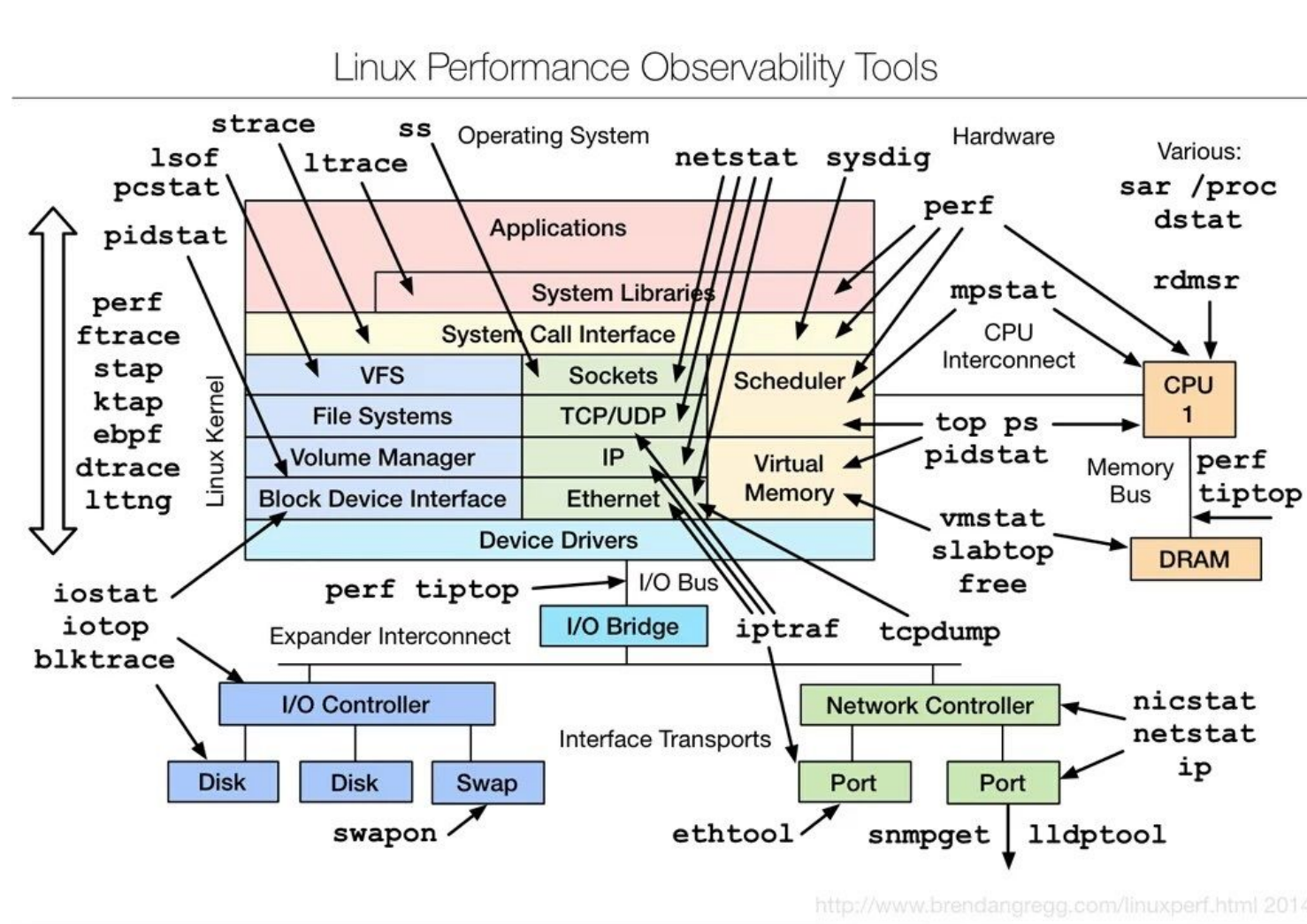
O pt-stalk:

```
2025_08_16_15_32_05-df            2025_08_16_15_32_05-mpstat         2025_08_16_15_32_05-procvmstat

2025_08_16_15_32_05-disk-space    2025_08_16_15_32_05-mpstat-overall 2025_08_16_15_32_05-ps

2025_08_16_15_32_05-diskstats     2025_08_16_15_32_05-mysqladmin     2025_08_16_15_32_05-slabinfo

2025_08_16_15_32_05-hostname      2025_08_16_15_32_05-netstat        2025_08_16_15_32_05-slave-status

2025_08_16_15_32_05-innodbstatus1 2025_08_16_15_32_05-netstat_s      2025_08_16_15_32_05-sysctl

2025_08_16_15_32_05-innodbstatus2 2025_08_16_15_32_05-numastat       2025_08_16_15_32_05-top

2025_08_16_15_32_05-interrupts    2025_08_16_15_32_05-opentables1    2025_08_16_15_32_05-transactions

2025_08_16_15_32_05-iostat        2025_08_16_15_32_05-opentables2    2025_08_16_15_32_05-trigger

2025_08_16_15_32_05-iostat-overall 2025_08_16_15_32_05-output        2025_08_16_15_32_05-variables

2025_08_16_15_32_05-lock-waits    2025_08_16_15_32_05-processlist    2025_08_16_15_32_05-vmstat

2025_08_16_15_32_05-meminfo       2025_08_16_15_32_05-procstat       2025_08_16_15_32_05-vmstat-overall
```

PERCONA

Mais detalhes!

CPU

- top
- pidstat
- mpstat
- CPU frequency

Disco

- iostat
- iotop
- pidstat -d

Memória

- /proc/meminfo
- swap
- swappiness
- sar -B
- pmap
- THP

Rede

- netstat
- ps (não é de redes)
- NFS
- nfsiostat
- nfsstat
- tcpdump

Por onde começar? Vmstat!

procs		memory				swap		io		system		cpu				
r	b	swpd	free	buff	cache	si	so	bi	bo	in	cs	us	sy	id	wa	st
6	1	7855772	2738416	1030560	14434456	2	7	2410	3663	2	2	3	2	94	1	0
2	3	7841080	2649620	1030572	14499056	14736	0	14736	12648	14339	18965	4	5	79	12	0
2	1	7831380	2561088	1030628	14582280	9908	0	9912	536708	15789	21347	5	8	74	13	0
3	1	7817288	2460360	1030680	14670372	14180	0	14180	11628	16563	20716	5	7	76	12	0
6	1	7807036	2401968	1030704	14722064	10352	0	10352	10808	11244	14896	4	5	79	12	0
2	0	7788156	2286776	1030732	14815840	18916	0	18916	12008	18277	24335	6	7	76	11	0
5	0	7774144	2153444	1030800	14930096	14092	0	14092	11932	17075	21561	6	8	73	12	0
3	1	7758948	2032568	1030840	15038872	15228	0	15228	10140	17226	21524	6	8	74	12	0
2	2	7742324	1969736	1030852	15070872	16696	0	16696	226732	12879	17663	3	6	81	11	0
2	1	7733772	2916060	1030884	14112948	8624	0	8624	325984	12317	16447	5	10	71	14	0
2	1	7719412	2791012	1030932	14227000	14824	0	14824	10436	17696	21634	7	8	74	12	0
3	1	7703388	2688812	1030980	14300700	16204	0	16204	9828	14894	19540	5	6	78	11	0
2	1	7687968	2594900	1031064	14371124	15528	0	15528	9892	14939	19890	5	6	78	11	0
2	0	7671992	2490496	1031132	14460112	16136	0	16136	10380	16252	21255	6	7	76	11	0
3	1	7652988	2414100	1031204	14522084	19212	0	19212	10260	16720	22323	4	7	78	11	0
2	2	7644948	2349800	1031232	14588544	8048	0	8048	9276	10848	13851	4	5	79	12	0
2	1	7630584	2232148	1031304	14692068	14380	0	14380	9696	16172	20611	6	7	75	12	0
3	1	7613368	2097384	1031320	14807428	17268	0	17268	11644	18684	23692	7	8	73	12	0
3	1	7593860	2023476	1031344	14861360	19492	0	19492	8308	15854	21407	4	6	79	11	0
5	1	7579868	1946236	1031376	14926072	14052	0	14052	8144	14055	18682	5	6	78	11	0
3	1	7566964	1782620	1031460	15076976	12960	0	12960	36660	21848	29820	18	20	50	12	0

Alguns exemplos e o quê estamos procurando - CPU

```
top - 19:40:45 up 23 days,  1:04,  4 users,  load average: 1.18, 1.09, 1.12
Tasks: 251 total,   3 running, 248 sleeping,   0 stopped,   0 zombie
%Cpu(s):  9.0 us, 10.5 sy,  0.0 ni, 66.9 id, 11.3 wa,  0.0 hi,  2.3 si,  0.0 st
MiB Mem : 128812.0 total,   2648.4 free, 111039.2 used,  15124.4 buff/cache
MiB Swap:   8192.0 total,    522.8 free,   7669.2 used.  16543.8 avail Mem

    PID USER       PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
1217314 mysql      20   0  200.5g 107.5g  20088 S  87.5  85.5 556:37.28 mysqld
    249 root        0 -20       0      0      0 R   6.2   0.0   2:54.73 kworker+
1291049 root       20   0    8368   3440   2084 S   6.2   0.0   0:00.01 bash
1291259 root       20   0   10616   3968   3360 R   6.2   0.0   0:00.01 top
      1 root       20   0  167664   8940   5724 S   0.0   0.0   3:35.30 systemd
      2 root       20   0       0      0      0 S   0.0   0.0   0:00.38 kthreadd
```

Mais detalhes - mpstat

11:45:14 AM	CPU	%usr	%nice	%sys	%iowait	%irq	%soft	%steal	%guest	%gnice	%idle
11:45:15 AM	all	18.71	0.00	2.52	0.55	0.00	0.38	0.00	0.00	0.00	77.84
11:45:15 AM	0	46.46	0.00	4.04	3.03	0.00	4.04	0.00	0.00	0.00	42.42
11:45:15 AM	1	35.71	0.00	4.08	2.04	0.00	0.00	0.00	0.00	0.00	58.16
11:45:15 AM	2	51.52	0.00	3.03	1.01	0.00	0.00	0.00	0.00	0.00	44.44
11:45:15 AM	3	53.00	0.00	2.00	2.00	0.00	1.00	0.00	0.00	0.00	42.00
11:45:15 AM	4	32.00	0.00	4.00	3.00	0.00	0.00	0.00	0.00	0.00	61.00
11:45:15 AM	5	44.44	0.00	3.03	2.02	0.00	0.00	0.00	0.00	0.00	50.51
11:45:15 AM	6	65.66	0.00	17.17	0.00	0.00	1.01	0.00	0.00	0.00	16.16
11:45:15 AM	7	52.53	0.00	3.03	1.01	0.00	1.01	0.00	0.00	0.00	42.42
11:45:15 AM	8	34.69	0.00	3.06	2.04	0.00	1.02	0.00	0.00	0.00	59.18
11:45:15 AM	9	47.96	0.00	5.10	2.04	0.00	1.02	0.00	0.00	0.00	43.88
11:45:15 AM	10	62.89	0.00	1.03	1.03	0.00	1.03	0.00	0.00	0.00	34.02
11:45:15 AM	11	54.55	0.00	2.02	1.01	0.00	1.01	0.00	0.00	0.00	41.41
11:45:15 AM	12	7.92	0.00	3.96	0.00	0.00	0.00	0.00	0.00	0.00	88.12
11:45:15 AM	13	4.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	94.00
11:45:15 AM	14	4.04	0.00	2.02	0.00	0.00	0.00	0.00	0.00	0.00	93.94
11:45:15 AM	15	3.03	0.00	2.02	0.00	0.00	0.00	0.00	0.00	0.00	94.95
11:45:15 AM	16	3.03	0.00	2.02	0.00	0.00	0.00	0.00	0.00	0.00	94.95
11:45:15 AM	17	3.00	0.00	4.00	0.00	0.00	0.00	0.00	0.00	0.00	93.00
11:45:15 AM	18	2.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	96.00
11:45:15 AM	19	4.00	0.00	2.00	1.00	0.00	0.00	0.00	0.00	0.00	93.00
11:45:15 AM	20	4.08	0.00	1.02	0.00	0.00	0.00	0.00	0.00	0.00	94.90
11:45:15 AM	21	5.05	0.00	1.01	1.01	0.00	0.00	0.00	0.00	0.00	92.93

[...]

Disco - iostat

Device:	rrqm/s	wrqm/s	r/s	w/s	rkB/s	wkB/s	avgrq-sz	avgqu-sz	await	r_await	w_await	svctm	%util
sda	1.36	8.03	1.49	10.87	20.69	100.32	19.58	0.01	0.89	0.22	0.98	0.04	0.05
sdb	0.04	840.30	69.58	2515.21	3156.37	32166.43	27.33	0.07	0.03	0.44	0.01	0.03	6.79
drbd0	0.00	0.00	69.62	3520.52	3156.18	32106.45	19.64	0.09	0.02	0.44	0.02	0.02	8.03
Device:	rrqm/s	wrqm/s	r/s	w/s	rkB/s	wkB/s	avgrq-sz	avgqu-sz	await	r_await	w_await	svctm	%util
sda	4.00	0.00	15.00	1.00	164.00	4.00	21.00	0.01	0.38	0.40	0.00	0.19	0.30
sdb	0.00	748.00	60.00	2028.00	960.00	23532.00	23.46	0.25	0.12	0.08	0.12	0.03	5.80
drbd0	0.00	0.00	60.00	2885.00	960.00	23660.00	16.72	8.60	2.92	0.08	2.97	0.14	40.50
Device:	rrqm/s	wrqm/s	r/s	w/s	rkB/s	wkB/s	avgrq-sz	avgqu-sz	await	r_await	w_await	svctm	%util
sda	0.00	63.00	3.00	152.00	76.00	860.00	12.08	0.11	0.72	1.00	0.72	0.02	0.30
sdb	0.00	759.00	184.00	2169.00	2944.00	24648.00	23.45	0.54	0.23	0.09	0.24	0.03	7.90
drbd0	0.00	0.00	185.00	3009.00	2960.00	24468.00	17.17	18.49	5.79	0.09	6.14	0.13	41.10

IO por processo - pidstat -d

```
07:46:41 AM     UID         PID    kB_rd/s    kB_wr/s kB_ccwr/s iodelay  Command
07:46:42 AM       0        1368      =1.00      =1.00     -1.00       3  auditd
07:46:42 AM       0     2480669      -1.00      -1.00     -1.00       1  sssd_kcm
07:46:42 AM       0     3240123      -1.00      -1.00     -1.00       1  sssd_be
07:46:42 AM    1921     3717915    1627.72    7398.02      0.00       0  mysqld

07:46:42 AM     UID         PID    kB_rd/s    kB_wr/s kB_ccwr/s iodelay  Command
07:46:43 AM       0        1368      =1.00      =1.00     -1.00       3  auditd
07:46:43 AM    1921     2482304       0.00       4.00      0.00       0  pidstat
07:46:43 AM    1921     3717915       0.00    4216.00      0.00       0  mysqld

07:46:43 AM     UID         PID    kB_rd/s    kB_wr/s kB_ccwr/s iodelay  Command
07:46:44 AM       0        1368      =1.00      =1.00     -1.00       3  auditd
07:46:44 AM    1921     2481653       0.00       4.00      0.00       0  pidstat
07:46:44 AM    1921     3717915       0.00   11228.00      0.00       0  mysqld
```

Memória - /proc/meminfo:

```
MemTotal:         792068956 kB

MemFree:            2385772 kB

MemAvailable:      72430296 kB

Buffers:                 28 kB

Cached:            69941412 kB

[...]

AnonHugePages:     53016576 kB

CmaTotal:                 0 kB
```

PERCONA

MySQL

Análise direcionada

Entendendo o banco de dados (pt-mysql-summary)

```
# Report On Port 3306 ####################################
           User | root@localhost
           Time | 2025-02-29 19:39:59 (UTC)
       Hostname | db-replica
        Version | 8.0.34-26 Percona Server (GPL), Release '26', Revision '0fe62c85'
       Built On | Linux x86_64
        Started | 2025-02-29 06:45 (up 0+12:54:30)
[...]
# Processlist ####################################

  Command                          COUNT(*) Working SUM(Time) MAX(Time)
  -------------------------------- -------- ------- --------- ---------
  Connect                                 1       1     45000     45000
  Daemon                                  1       1     45000     45000
  Query                                  10      10       450       200
[...]
```
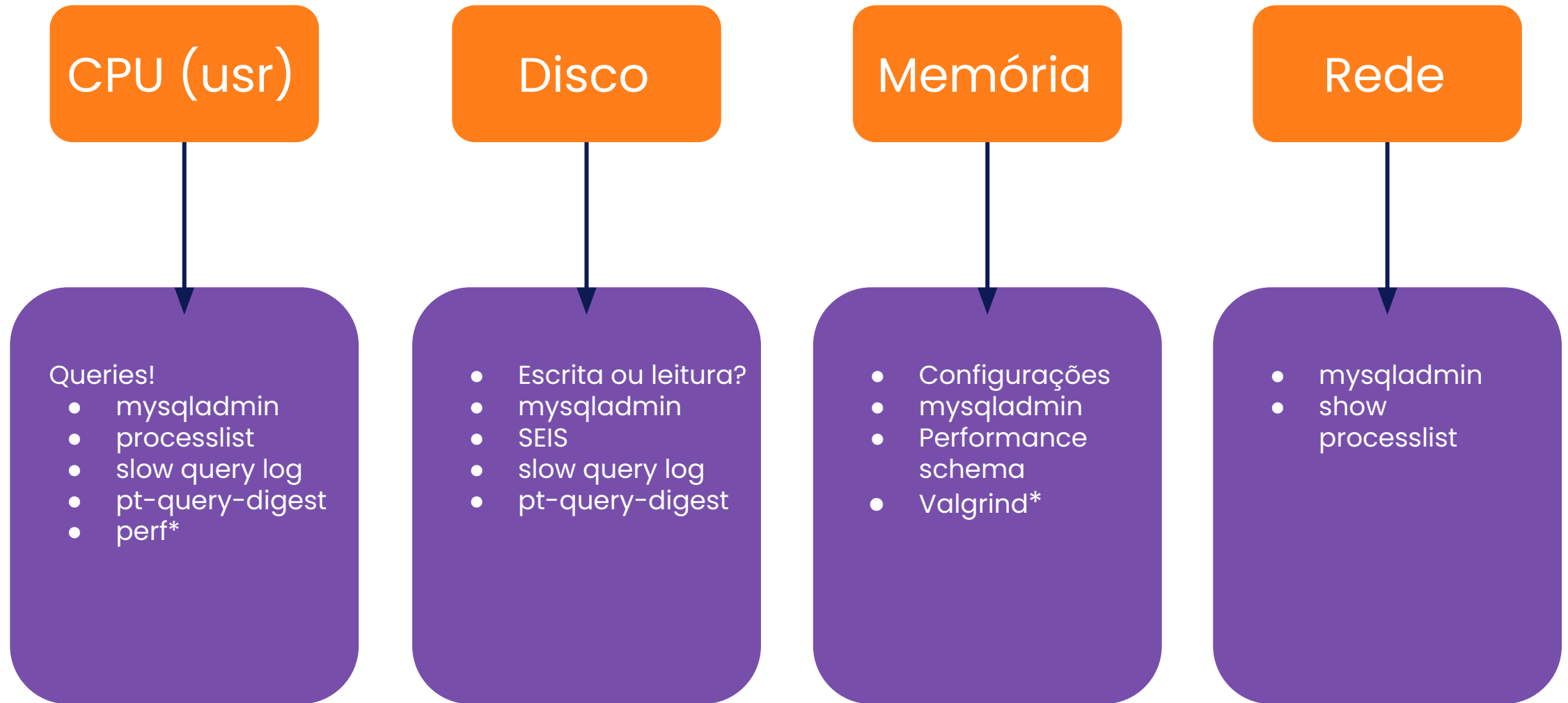
Entendendo o banco de dados (pt-mysql-summary)

```
# InnoDB #####################################################
             Version | 8.0.34-26
    Buffer Pool Size | 184.0G
    Buffer Pool Fill | 50%
   Buffer Pool Dirty | 4%
      File Per Table | ON
       Log File Size | 2 * 8.0G = 16.0G
     Log Buffer Size | 16M
        Flush Method | fsync
 Flush Log At Commit | 1
      Checkpoint Age | 3G
    History List Len | 394
   max_heap_table_size | 16G
       tmp_table_size | 16G
```

Direcionando a nossa análise

CPU (usr)

Queries!
- mysqladmin
- processlist
- slow query log
- pt-query-digest
- perf*

Disco

- Escrita ou leitura?
- mysqladmin
- SEIS
- slow query log
- pt-query-digest

Memória

- Configurações
- mysqladmin
- Performance schema
- Valgrind*

Rede

- mysqladmin
- show processlist

CPU (usr)

- Muitas queries
  - mysqladmin
    - Threads_running
    - Queries
- Queries com alto processamento (group by, order by, agregações)
  - mysqladmin
    - Handler_read_*
    - Rows_sorted
- SHOW PROCESSLIST
- Slow query log
- perf*

A ferramenta mysqladmin:

mysqladmin (SHOW GLOBAL STATUS):

```
+------------------------------------+----------------------
| Variable_name                      | Value
+------------------------------------+----------------------
| Aborted_clients                    | 341
| Aborted_connects                   | 58
| Aborted_connects_preauth           | 58
| Access_denied_errors               | 0
| Acl_column_grants                  | 0
| Acl_database_grants                | 1
| Acl_function_grants                | 0
| Acl_procedure_grants               | 0
| Acl_package_spec_grants            | 0
[...]
```

Usando mysqladmin com pt-mext:

```
shell > pt-mext -r -- cat 2025_08_16_15_32_05-mysqladmin >
2025_08_16_15_32_05-mysqladmin.mext
```

Aborted_clients	341	0	0	0	0	1	0	0	1	0	0	0	0	0	0
Aborted_connects	58	0	0	0	0	0	0	0	0	0	0	0	1	0	13
Aborted_connects_preauth	58	0	0	0	0	0	0	0	0	0	0	0	1	0	13
Access_denied_errors	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Acl_column_grants	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Acl_database_grants	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
Acl_function_grants	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

[...]

CPU - mysqladmin.mext

Threads_running	**35**	-11	**15**	**40**	**30**	-3	-12	-22
Queries	**35257685**	30105	**474350**	**410880**	**401232**	36849	37101	34370
Rows_read	**9084319798**	122884	**171800**	**9783766**	**803720**	152813	192471	733985
Sort_rows	**141471818**	158818	31381	33813	25874	16861	33254	33474
Innodb_row_lock_time	**882360**	0	22	726	10	9	0	1

CPU - SHOW PROCESSLIST

```
*************************** 1. row ***************************
      Id: 1
    User: system user
    Host:
      db: NULL
 Command: Daemon
    Time: NULL
   State: InnoDB purge coordinator
    Info: NULL
Progress: 0.000
*************************** 2. row ***************************
      Id: 2
    User: system user
    Host:
      db: NULL
 Command: Daemon
    Time: NULL
   State: InnoDB purge worker
    Info: NULL
Progress: 0.000

[...]
```

SHOW PROCESSLIST

```
shell > grep -h State *processlist | sort | uniq -c  | sort -nr
16374    State:
  987    State: Sending data
  840    State: Opening tables
  296    State: Update
  264    State: Creating sort index
[...]

shell > grep -h Command *processlist | sort | uniq -c  | sort -nr
16392  Command: Sleep
 2214  Command: Query
 1080  Command: Daemon
  112  Command: Connect
    1  Command: Ping

shell > grep -h "Command: Query" *processlist  -A1  | grep Time  | sort -nrk2   | head -n 6
    Time: 113
    Time: 112
    Time: 111
    Time: 111
    Time: 110
    Time: 110
```

Disco

- Escrita
  - Tamanho de Redo log
  - mysqladmin
    - Innodb_os_log_written
    - Innodb_checkpoint_age
    - Innodb_data_written
    - Innodb_buffer_pool_pages_flushed
    - Innodb_rows_inserted, _deleted, _updated
    - Innodb_os_log_pending_fsyncs, _writes
    - Innodb_data_pending_fsyncs, _writes
    - Created_tmp_disk_tables
    - Binlog_cache_disk_use
  - SHOW ENGINE INNODB STATUS
  - SHOW PROCESSLIST

Escrita - SHOW ENGINE INNODB STATUS:

---TRANSACTION 18548610466, **ACTIVE 11 sec**
mysql tables in use 1, locked 1
748 lock struct(s), heap size 123000,**71229 row lock(s), undo log entries 70938**
MySQL thread id 11, OS thread handle 139794744915520, query id 12673731**Applying batch of
row changes (delete)**
[...]
---TRANSACTION 18548487502, ACTIVE (PREPARED) 0 sec committing
mysql tables in use 1, locked 1
1 lock struct(s), heap size 1128, 0 row lock(s), undo log entries 1
MySQL thread id 3605448, OS thread handle 140364097828416, query id 126599638 <IP> <USER>
waiting for handler commit
INSERT IGNORE INTO [redacted]
[...]
Pending normal aio reads: [0, 0, 0, 0] ,**aio writes: [6, 10, 2, 4] ,**

Disco

- Leitura
  - mysqladmin
    - Innodb_data_read
    - Innodb_buffer_pool_reads
    - Innodb_buffer_pool_pages_data, _dirty, _free
    - Innodb_data_pending_reads
    - Handler_read_*
    - Handler_read_rnd_next
    - Created_tmp_tables
  - SHOW PROCESSLIST
  - Slow Query Log

Slow Query Log

- Auditoria de query
  - slow_query_log=1
  - long_query_time=0
  - Percona Server for MySQL
    - log_slow_rate_limit
    - log_slow_rate_type (session, query)
    - slow_query_log_use_global_control
- Relatório
  - pt-query-digest

Relatório (pt-query-digest)

```
# 1.8s user time, 20ms system time, 63.80M rss, 390.15G vsz
# Current date: Fri Dec  1 19:21:14 2023
# Hostname: XXXXX
# Files: slow_13_April.txt
# Overall: 1.54k total, 57 unique, 0.01 QPS, 0.36x concurrency _____
# Time range: 2025-04-11T00:06:21 to 2023-04-13T12:25:33
# Attribute          total     min     max     avg     95%  stddev  median
# ============     ======= ======= ======= ======= ======= ======= =======
# Exec time         78204s     10s  30837s     51s     63s    788s     17s
# Lock time             2s    19us   189ms     1ms     1ms     8ms   761us
# Rows sent         16.44M       0   3.25M  10.95k  223.14 166.68k       0
# Rows examine       5.07G     179   1.51G   3.38M   9.30M  39.23M 345.04k
# Rows affecte       6.73M       0   1.95M   4.49k  719.66  84.09k  381.65
# Bytes sent         1.02G       0 293.74M 693.62k  38.40k  12.61M       0
# Query size        26.81M      29   2.15M  17.86k  18.47k  92.35k  18.47k
```

Relatório (pt-query-digest)

```
# Profile
# Rank Query ID              Response time     Calls R/Call     V/M   Item
# ==== ================= ================ ===== ========== ===== ========
#    1 0x61FC30FD48BDB2BD 30974.8336 39.6%        2 15487.4168 30... SELECT UNION tb1 tb2 tb3
#    2 0xBA73B33390C045C8 26380.0881 33.7%     1024    25.7618 17.94 INSERT SELECT table4
#    3 0xD5EFE2CCEF8B7E1E  2746.2674  3.5%        3   915.4225 79... SELECT [redacted]
#    4 0xB7B1F398523D42C9  1934.9007  2.5%       74    26.1473  7.23 SELECT UNION [redacted]
#    5 0xD803B305A2A92E25  1882.2888  2.4%      108    17.4286  4.09 SELECT [redacted]
#    6 0x5B06045903FB685B  1653.3620  2.1%       52    31.7954  2.24 SELECT [redacted]
#    7 0x361D39E89A10AEAD  1570.7166  2.0%        3   523.5722 10... CALL `function`
[...]
```

Relatório (pt-query-digest)

```
# Query 1: 0.00 QPS, 0.26x concurrency, ID 0x61FC30FD48BDB2BD at byte 26217073
# This item is included in the report because it matches --limit.
# Scores: V/M = 30425.85
# Time range: 2023-04-11T21:12:17 to 2023-04-13T05:43:57
# Attribute     pct    total      min      max      avg      95%   stddev   median
# ============ === ======= ======= ======= ======= ======= ======= =======
# Count           0       2
# Exec time      39   30975s     138s   30837s   15487s   30837s   21708s   15487s
# Lock time       0      4ms      1ms      3ms      2ms      3ms    980us      2ms
# Rows sent       0   10.61k    5.20k    5.41k    5.31k    5.41k   152.03    5.31k
# Rows examine   29    1.52G   12.35M    1.51G  777.92M    1.51G    1.06G  777.92M
# Rows affecte    0        0        0        0        0        0        0        0
# Bytes sent      0    1.00M  501.92k  523.32k  512.62k  523.32k   15.13k  512.62k
# Query size      0    6.18k    3.09k    3.09k    3.09k    3.09k        0    3.09k
[...]
# Tables
#    SHOW TABLE STATUS FROM `database1` LIKE 'table1'\G
#    SHOW CREATE TABLE `database1`.`table1`\G
[...]
<redacted query>
```

Memória

- Configuração básica
  - Innodb_buffer_pool_size
  - Tabelas temporárias:
    - max_heap_size
    - max_tmp_table_size
    - internal_tmp_mem_storage_engine
    - temptable_max_ram
    - temptable_max_mmap
      - temptable_use_mmap
- Instrumentação de memória do Performance Schema
- Valgrind*

Performance Schema

- Dinâmico

```
UPDATE performance_schema.setup_instruments SET ENABLED = 'YES' WHERE
NAME LIKE 'memory/%';
```

- Restart

```
[mysqld]
```

```
performance-schema-instrument='memory/%=ON'
```

Instrumentação de memória - Performance Schema

```
SELECT * FROM sys.memory_global_by_current_bytes;
SELECT * FROM sys.memory_by_host_by_current_bytes;
SELECT * FROM sys.memory_by_thread_by_current_bytes;
SELECT * FROM sys.memory_by_user_by_current_bytes;
SELECT * FROM performance_schema.memory_summary_global_by_event_name;
```

Instrumentação de memória - Performance Schema

```
mysql> select event_name, CURRENT_NUMBER_OF_BYTES_USED from
performance_schema.memory_summary_global_by_event_name order by CURRENT_NUMBER_OF_BYTES_USED desc limit
10;
+----------------------------------------------------------+------------------------------+
| event_name                                               | CURRENT_NUMBER_OF_BYTES_USED |
+----------------------------------------------------------+------------------------------+
| memory/innodb/buf_buf_pool                               |                 123884044288 |
| memory/innodb/hash0hash                                  |                    3301798144 |
| memory/innodb/buf0buf                                    |                     978383808 |
| memory/sql/THD::main_mem_root                            |                     415611096 |
| memory/sql/TABLE                                         |                      94581184 |
| memory/innodb/memory                                     |                      84150712 |
| memory/performance_schema/events_errors_summary_by_host_by_error    |          43767808 |
| memory/performance_schema/events_errors_summary_by_account_by_error |          43767808 |
| memory/performance_schema/events_statements_summary_by_digest       |          41600000 |
| memory/performance_schema/table_handles                 |                      38010880 |
```

Redes

- mysqladmin + Processlist
  - Bytes_received
    - Queries grandes (DML com Large Objects)
  - Bytes_sent
    - Queries retornando muitos dados
      - Queries retornando Large Objects
      - mysqldump
      - clone

Próximos passos

Como continuar se não encontrar saturação?

- MySQL error log
- innodb_thread_concurrency
- SHOW ENGINE INNODB STATUS (SEIS)
  - Locks
  - Mutexes
  - History List Length
  - Transações de longa duração
  - Pending I/O operations

SEIS – Mutexes

```
----------
SEMAPHORES
----------
OS WAIT ARRAY INFO: reservation count 74141362449
--Thread 139618702718720 has waited at btr0sea.ic line 128 for 0.00 seconds the semaphore:
S-lock on RW-latch at 0x5583cc0cfea8 created in file btr0sea.cc line 230
number of readers 6, waiters flag 0, lock_word: 1fffffffa
Last time read locked in file btr0sea.ic line 128
Last time write locked in file
/mnt/workspace/percona-xtradb-cluster-5.7-debian-binary/label_exp/min-bionic-x64/percona-xtradb
-cluster-5.7-5.7.26-31.37/storage/innobase/include/btr0sea.ic line 90
[...]
--Thread 139795325380160 has waited at ha_innodb.cc line 19920 for 4 seconds the semaphore:
Mutex at 0x56415cba0e60, Mutex RW_LOCK_LIST created sync0debug.cc:1674, locked by
139772063462976
```

PERCONA

SEIS - Transactions

```
------------
TRANSACTIONS
------------
[...]
```
History list length 151
```
LIST OF TRANSACTIONS FOR EACH SESSION:
[...]
```
---TRANSACTION 4418712625, **ACTIVE 18 sec starting index read**
```
mysql tables in use 3, locked 3
LOCK WAIT 2 lock struct(s), heap size 1128, 1 row lock(s)
MySQL thread id 70729626, OS thread handle 140637547157248, query id 7996430202
<ip> <database> updating
```
UPDATE [redacted]
------- TRX HAS BEEN WAITING 18 SEC FOR THIS LOCK TO BE GRANTED:
```
RECORD LOCKS space id 30714 page no 42997 n bits 128 index PRIMARY of table
`<database>`.`<table>` trx id 4418712625
```
lock_mode X locks rec but not gap waiting
```
Record lock, heap no 19 PHYSICAL RECORD: n_fields 53; compact format; info bits 0
```

Visão Macro - mysqladmin

- Threads_connected, Threads_running
- Innodb_buffer_pool_reads
- Innodb_data_read
- Handler_read_*
- Innodb_data_written
- Innodb_*pending*
- Binlog_cache_disk_use

Acesso à arquivos

- Table_open_cache_hits, _misses
  - table_open_cache
- Open_files, _tables, Opened_files, _tables
  - open_files_limit
- Innodb_num_open_files
  - innodb_open_files
- Open_table_definitions, Opened_table_definitions
  - table_definition_cache

PERCONA

Caso exemplo

```
Opened_tables                  1038776101  13310   13812  13228  13356  14060   13468
Table_open_cache_misses        938174187   12718   2417   12135  12767  12222   13479
```

Original:

```
    PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM    TIME+ COMMAND
 128838 mysql     20   0  166.9g 131.1g  39524 S  1057   8.7  11457:29 mysqld
 128838 mysql     20   0  166.9g 131.2g  39540 S  947.9  8.7  11457:58 mysqld
 128838 mysql     20   0  166.9g 131.1g  39540 S  913.5  8.7  11458:27 mysqld
```

Depois de ajustar `open_files_limit, innodb_open_files` e `table_open_cache`:

```
    PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM    TIME+ COMMAND
3863380 mysql     20   0  160.4g 126.9g  38952 S  100.0  8.4   6631:06 mysqld
3863380 mysql     20   0  160.4g 126.9g  38952 S  191.9  8.4   6631:12 mysqld
3863380 mysql     20   0  160.4g 126.9g  38952 S  191.0  8.4   6631:18 mysqld
```
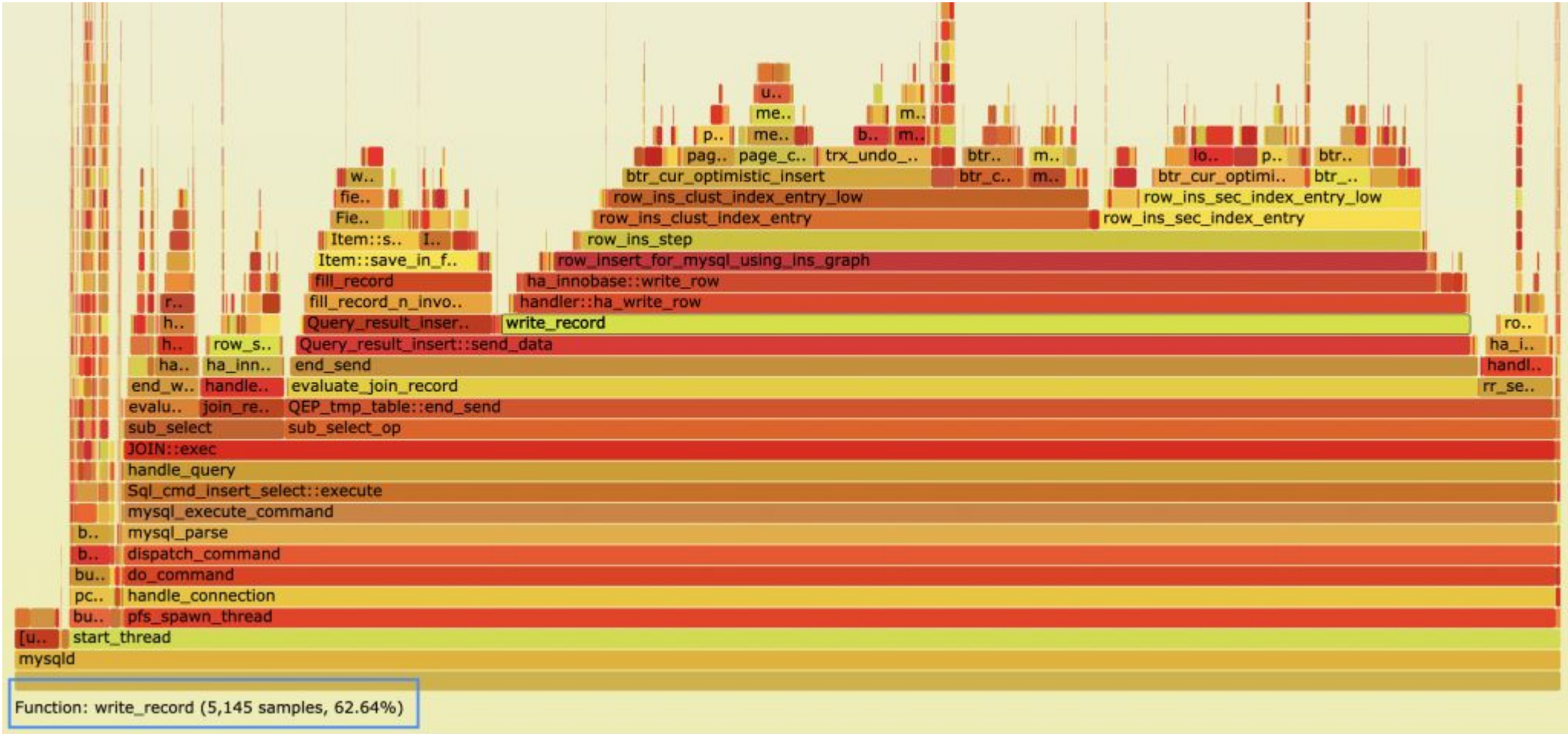
Se tudo mais falhar

- Perf + Flamegraph
  - https://github.com/brendangregg/FlameGraph

```
shell > sudo perf record -a -F 99 -g -p $(pgrep -x mysqld) -- sleep 10

shell > sudo perf script > perf.script

shell > ~/src/FlameGraph/stackcollapse-perf.pl perf.script |
~/src/FlameGraph/flamegraph.pl > flamegraph.svg
```

Se tudo mais falhar

Function: write_record (5,145 samples, 62.64%)

Lentidão vs. não responsivo

- Bugs!
  - pt-pmp
    - Somente o stack trace de cada thread
    - Necessário gdb e debug symbols
  - Core dump
    - gcore

Pt-pmp

```
shell > pt-pmp --pid $(pgrep -x mysqld) --save-samples pt-pmp.log


  17260 __lll_lock_wait(libpthread.so.0),_L_lock_1093(libpthread.so.0),pthread_mutex_lock(libpthread
 .so.0),THD::init,THD::THD,Channel_info::create_thd,Channel_info_tcpip_socket::create_thd,handle_conn
ection,pfs_spawn_thread,start_thread(libpthread.so.0),clone(libc.so.6)
   1175 _lll_lock_wait(libpthread.so.0),_L_lock_1093(libpthread.so.0),pthread_mutex_lock[...]
[...]
      1 __lll_lock_wait(libpthread.so.0),_L_lock_1093(libpthread.so.0),pthread_mutex_lock(libpthread
.so.0),System_variable::init,PFS_system_variable_cache::do_materialize_session, table_variables_by_thr
ead::rnd_next,ha_perfschema::rnd_next,handler::ha_rnd_next,filesort,QEP_TAB::sort_table,join_init_rea
d_record,sub_select,JOIN::exec,handle_query,execute_sqlcom_select,mysql_execute_command,mysql_parse,
dispatch_command,do_command,handle_connection,pfs_spawn_thread,start_thread(libpthread.so.0),clone(li
bc.so.6)
      1 __lll_lock_wait(libpthread.so.0),_L_lock_1093(libpthread.so.0),pthread_mutex_lock(libpthread
.so.0),cleanup_variables,plugin_thdvar_init,THD::init,THD::cleanup_connection,dispatch_command,do_co
mmand,handle_connection,pfs_spawn_thread,start_thread(libpthread.so.0),clone(libc.so.6)
```

PMM

Links úteis

- https://www.percona.com/software/database-tools/percona-monitoring-and-management
- https://pmmdemo.percona.com/graph/
- https://docs.percona.com/percona-server/8.0/slow-extended.html
- https://docs.percona.com/percona-toolkit/index.html
- https://www.percona.com/blog/profiling-software-using-perf-and-flame-graphs/
- https://www.percona.com/blog/inspecting-mysql-servers-part-1-the-percona-support-way/
- https://www.percona.com/blog/inspecting-mysql-servers-part-2-knowing-the-server/
- https://www.percona.com/blog/inspecting-mysql-servers-part-3-what-mysql/
- https://www.percona.com/blog/inspecting-mysql-servers-part-4-an-engine-in-motion/
- https://www.percona.com/blog/inspecting-mysql-servers-part-5-percona-monitoring-and-management/

https://mysqlbr.com.br/

Perguntas?

percona.com

Obrigado!

percona.com