

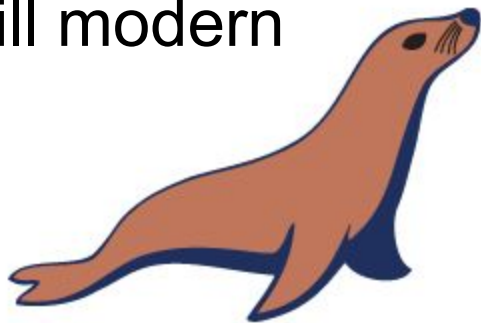
MariaDB PaaS using systemd multi-instance
+ socket activation

Bare metal, but still modern

Daniel Black

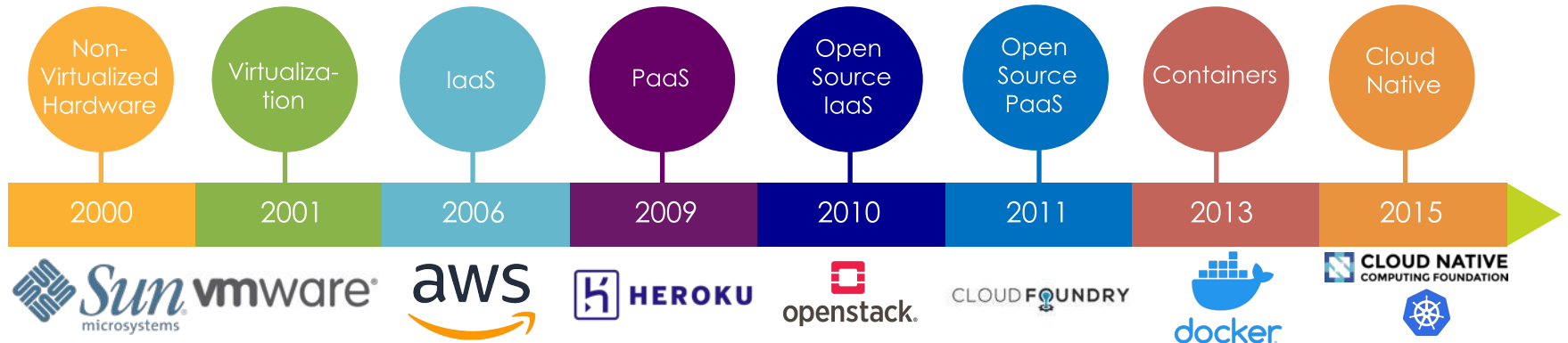
Chief Innovation Officer

(Making MariaDB Better)



MariaDB
Foundation

From Virtualization to Cloud Native



Cloud Requirements

Extract from **CNCF Cloud Native Definition v1.0**

- Loosely coupled systems
- Resilient
- Manageable
- Observable
- Robust automation

MariaDB on Shared Hosting?

It's a tradeoff:

Good:

- Database priv/user separation
- Simple

Weaknesses:

- Single mariadb process
- Settings are global or session (no per database/user)
- Shared items, table cache, innodb threads/history

MariaDB's systemd story

- 2015-10-15 MariaDB-10.1.8 Type=notify
- 2016-2018 compatibility packaging on sysv/systemd pain
- 2017-12-12 MariaDB-10.4.6 Flexible multi-service

Systemd

- 2017-12-17 v236 notify message to delay startup/runtime/stop timeout limits
- 2019-02-18 v242 killmode=cgroup|mixed, SendSIGKILL=no services are not multiprocess

Multi-instance MariaDB

- Separate data directory (where databases are stored)
- Unique TCP listen port
- Unique Unix socket

Current Systemd Implementation

/usr/lib/systemd/system/mariadb@.service

Initialization

ExecStartPre=.../mariadb-install-db **\$MYSQLD_MULTI_INSTANCE**

Execution

ExecStart=.../mariabdb **\$MYSQLD_MULTI_INSTANCE** \$MYSQLD_OPTS

Demo Time

Socket Activation (serverless)

Add one file:

```
[root@grit /]# cat /etc/systemd/system/mariadb@.socket

[Unit]
Description=MariaDB Sockets

[Socket]

SocketUser=%I
SocketMode=777

ListenStream=@mariadb-%I

Backlog=150
```

Socket Activation (serverless)

Add one file:

```
[root@grit /]# systemctl start mariadb@tom.socket
[root@grit /]# systemctl start mariadb@mary.socket
[root@grit /]# systemctl start mariadb@jan.socket
[root@grit /]# systemctl start mariadb@fran.socket
[root@grit /]# systemctl status mariadb@fran.socket
● mariadb@fran.socket - MariaDB Sockets
   Loaded: loaded (/etc/systemd/system/mariadb@.socket; static; vendor preset: disabled)
   Active: active (listening) since Wed 2021-01-20 21:44:04 AEDT; 2min 40s ago
 Triggers: ● mariadb@fran.service
   Listen: @mariadb-fran (Stream)
    Tasks: 0 (limit: 38291)
   Memory: 0B
     CPU: 862us
    CGroup: /system.slice/system-mariadb.slice/mariadb@fran.socket

Jan 20 21:44:04 grit systemd[1]: Starting MariaDB Sockets.
Jan 20 21:44:04 grit systemd[1]: Listening on MariaDB Sockets.
```

Socket Activation (serverless)

Systemd is listening:

```
[root@grit /]# lsof -p 1 | grep mariadb
lsof: WARNING: can't stat() fuse.gvfsd-fuse file system /run/user/1000/gvfs
Output information may be incomplete.
systemd  1 root  55u    unix 0x000000009f94ad9b    0t0    817417 @mariadb-tom type=STREAM
systemd  1 root  57u    unix 0x00000000e558ba4a    0t0    814430 @mariadb-mary type=STREAM
systemd  1 root  58u    unix 0x00000000c7c62c57    0t0    814437 @mariadb-jan type=STREAM
systemd  1 root  59u    unix 0x000000000c074d39    0t0    813487 @mariadb-fran type=STREAM
```

Socket Activation (serverless)

Systemd is listening:

```
[root@grit /]# lsof -p 1 | grep mariadb
lsof: WARNING: can't stat() fuse.gvfsd-fuse file system /run/user/1000/gvfs
Output information may be incomplete.
systemd  1 root  55u    unix 0x000000009f94ad9b      0t0      817417 @mariadb-tom type=STREAM
systemd  1 root  57u    unix 0x00000000e558ba4a      0t0      814430 @mariadb-mary type=STREAM
systemd  1 root  58u    unix 0x00000000c7c62c57      0t0      814437 @mariadb-jan type=STREAM
systemd  1 root  59u    unix 0x000000000c074d39      0t0      813487 @mariadb-fran type=STREAM
```

Socket Activation (serverless)

Connect and...

```
[root@grit /]# systemctl status mariadb@tom.socket
● mariadb@tom.socket - MariaDB Sockets
   Loaded: loaded (/etc/systemd/system/mariadb@.socket; static; vendor preset: disabled)
   Active: active (running) since Wed 2021-01-20 21:44:57 AEDT; 5min ago
     Triggers: ● mariadb@tom.service
   Listen: @mariadb-tom (Stream)
    Tasks: 0 (limit: 38291)
   Memory: 0B
     CPU: 779us
   CGroup: /system.slice/system-mariadb.slice/mariadb@tom.socket

Jan 20 21:44:57 grit systemd[1]: Starting MariaDB Sockets.
Jan 20 21:44:57 grit systemd[1]: Listening on MariaDB Sockets.
[root@grit /]# systemctl status mariadb@tom.service
● mariadb@tom.service - MariaDB 10.5.8 database server (multi-instance tom)
   Loaded: loaded (/usr/lib/systemd/system/mariadb@.service; disabled; vendor preset: disabled)
  Drop-In: /etc/systemd/system/mariadb@.service.d
           └─user.conf
   Active: active (running) since Wed 2021-01-20 21:49:40 AEDT; 23s ago
TriggeredBy: ● mariadb@tom.socket
     Docs: man:mariadb(8)
           https://mariadb.com/kb/en/library/systemd/
   Process: 100019 ExecStartPre=/bin/sh -c [ -f /home/tom/.my.cnf ] || sed -e "s/USER/tom/g" -e "s/PORT/${id -u t
   Process: 100020 ExecStartPre=/usr/bin/mkdir -p /home/tom/mariadb-datadir (code=exited, status=0/SUCCESS)
   Process: 100021 ExecStartPre=/usr/bin/mariadb-install-db --sysconf=/etc/my.cnf --multi-instance --no-auth-root-authenticati
```

MariaDB Socket Activation (serverless)

WIP - MDEV-5536 Coming in 10.6

```
569     585     }
570     586     #endif
571     587
588 + /** mysql_socket_fd */
589 +
590 + static inline MYSQL_SOCKET
591 + inline_mysql_socket_fd
592 + (
593 + #ifdef HAVE_PSI_SOCKET_INTERFACE
594 +     const char *src_file, uint src_line,
595 + #endif
596 +     int fd)
597 + {
598 +     MYSQL_SOCKET mysql_socket= MYSQL_INVALID_SOCKET;
599 +     mysql_socket.fd= fd;
600 +
601 +     if (likely(mysql_socket.fd != INVALID_SOCKET))
602 +     {
603 + #ifdef HAVE_PSI_SOCKET_INTERFACE
604 +         /*     PSI_socket_locker_state state;
605 +         PSI_socket_locker *locker;
606 +         locker= PSI_SOCKET_CALL(init_socket)
607 +         mysql_socket.m_psi= PSI_SOCKET_CALL(init_socket)
608 +         (key, (const my_socket*)&mysql_socket, addr, *addr_len); */
```

Systemd against Cloud Requirements

- Loosely coupled systems
 - Independent services, same OS
- Resilient
 - Hardware / OS SPOF
 - Independent services
- Manageable
 - Create user/manage service
- Observable
 - `journalctl -u mariadb@bob`
 - Distinct cgroups
- Robust automation
 - Systemd service restart automated

Questions