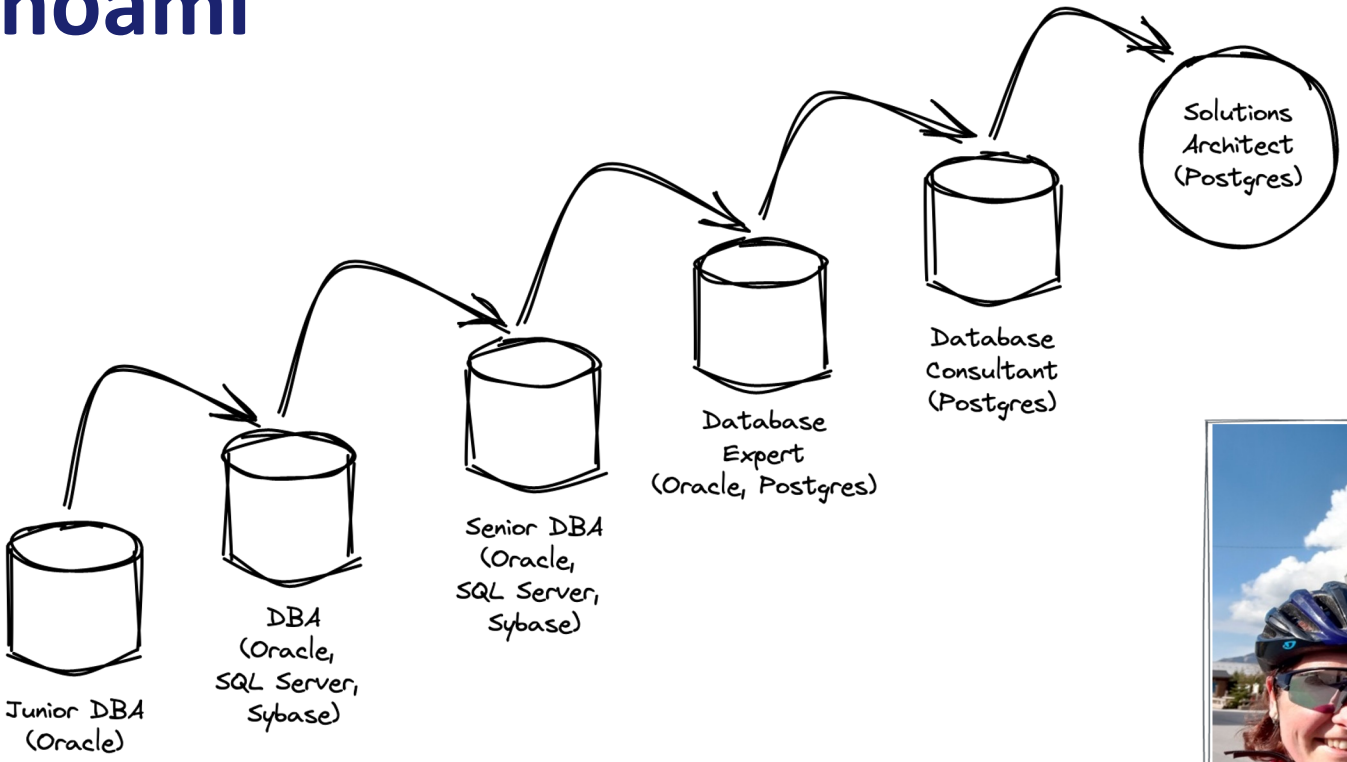
Tuning PostgreSQL to work even better

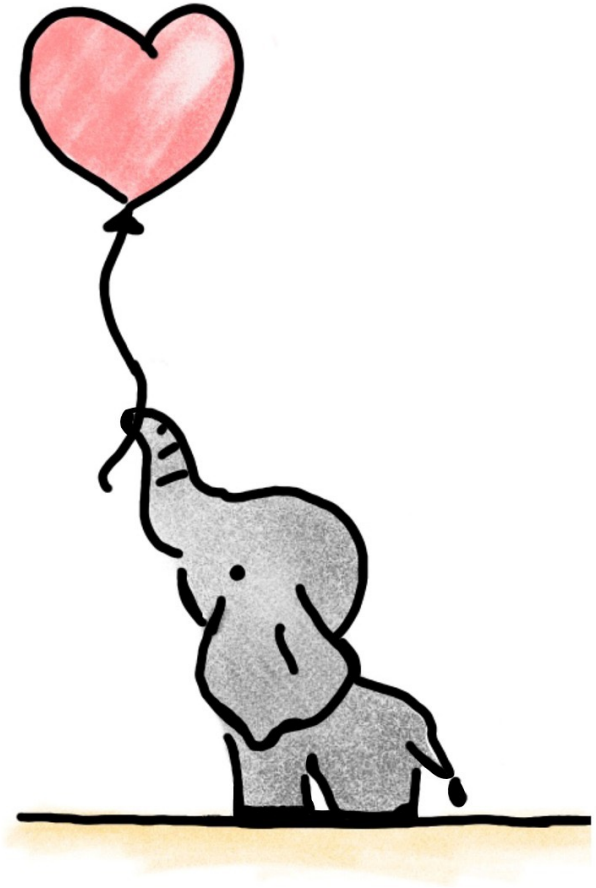
Karen Jex | Senior Solutions Architect @ Crunchy Data

DjangoCon Europe | Edinburgh | May 2023

whoami


Junior DBA
(Oracle)
DBA
(Oracle,
SQL Server,
Sybase)
Senior DBA
(Oracle,
SQL Server,
Sybase)
Database
Expert
(Oracle, Postgres)
Database
Consultant
(Postgres)
Solutions
Architect
(Postgres)

Introduction



Introduction

- Minimal footprint by default
- Need to tune for production
- ~350 parameters
- Don’t need to know all of the parameters!

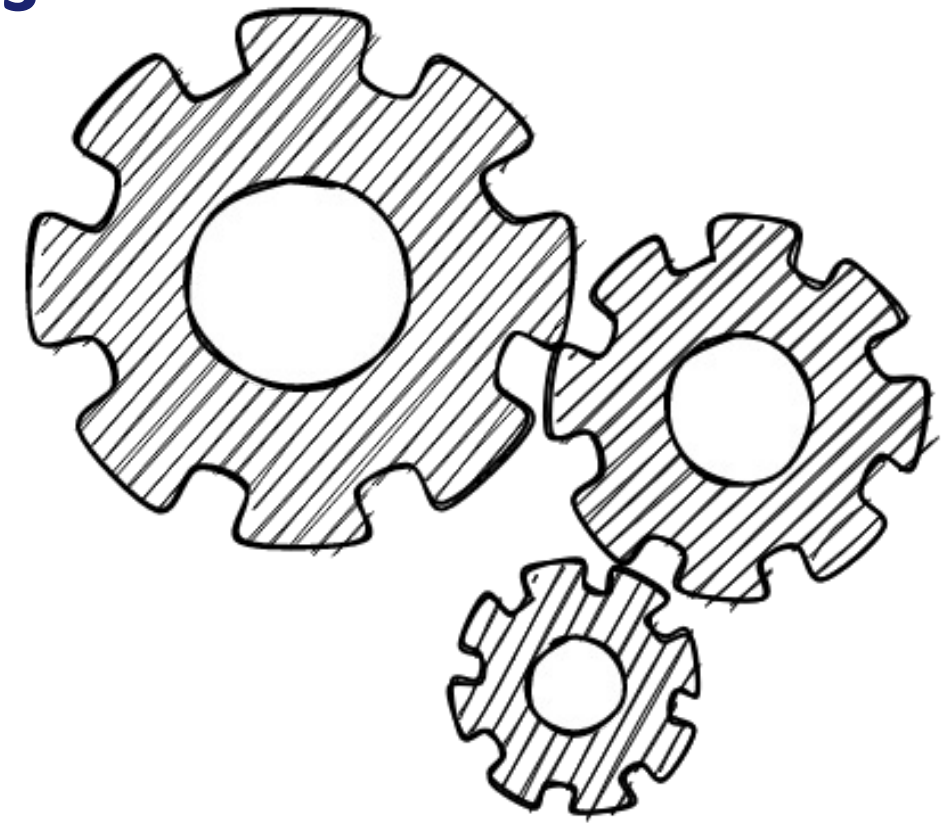
Agenda

- Setting and Viewing Parameters
- Most Important Parameters
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - Query Tuning
- Summary

Agenda

- **Setting and Viewing Parameters**
- Most Important Parameters
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - Query Tuning
- Summary

Setting Parameters

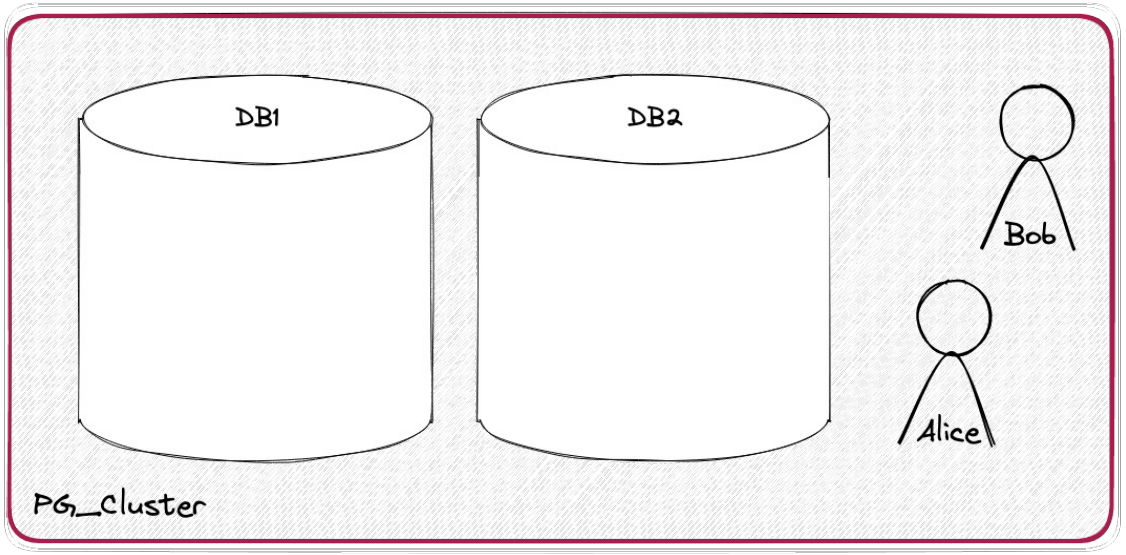


Setting Parameters: Cluster level

postgresql.conf

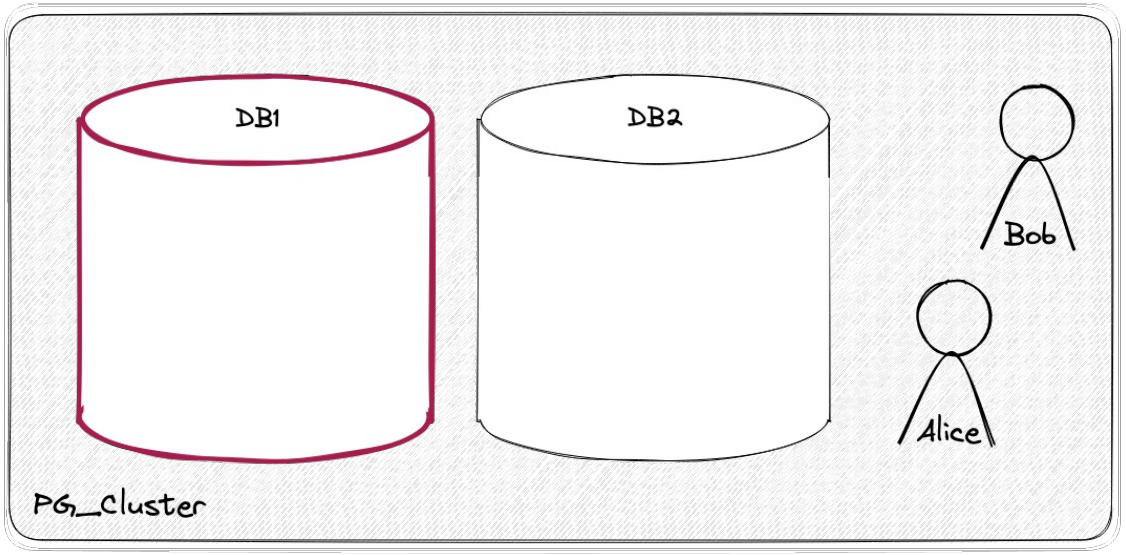
or

ALTER SYSTEM
SET parameter=value;



Setting Parameters: Database level

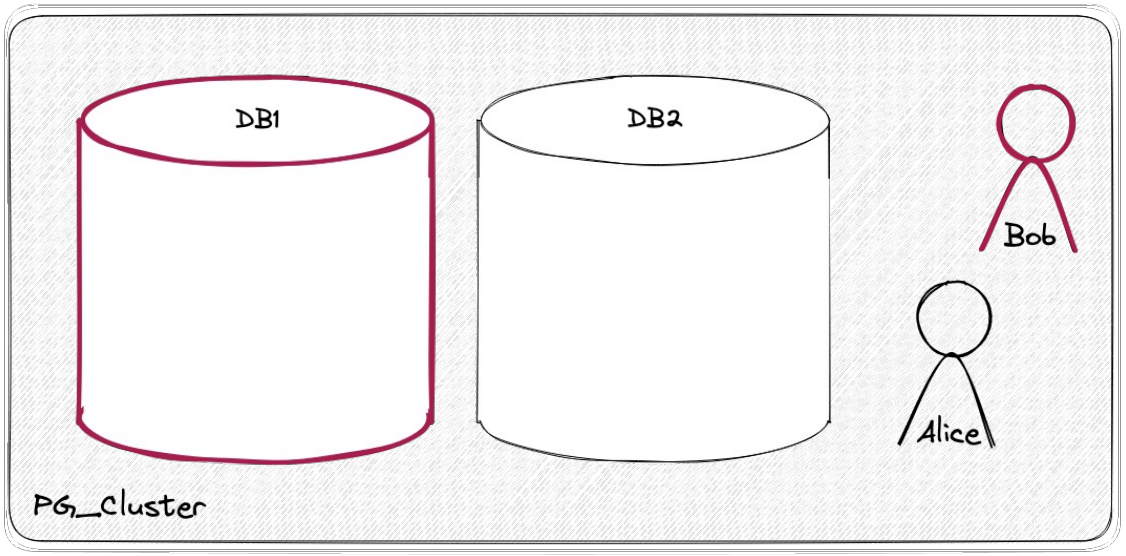
ALTER DATABASE db1
SET parameter=value;



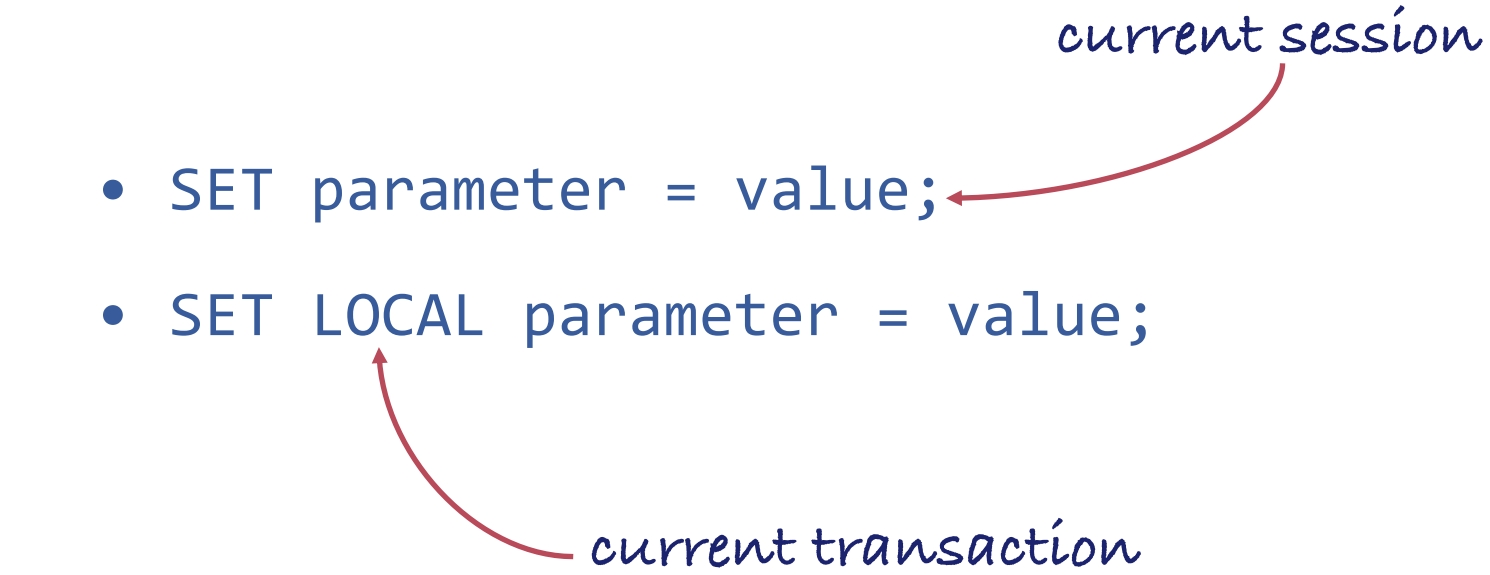
Setting Parameters: User/Role level

ALTER ROLE bob
SET parameter=value;

ALTER ROLE bob
IN DATABASE db1
SET parameter=value;



Setting Parameters: Session level



Setting Parameters: Context

```
postgres=# SELECT DISTINCT context FROM pg_settings;

      context
-------------------
 postmaster
 superuser-backend
 user
 internal
 backend
 sighup
 superuser
```

Setting Parameters - context: internal

```
      context
-------------------

 internal                set at compile time, can't be changed
                         e.g. block_size, server_version, wal_segment_size
```

Setting Parameters - context: postmaster

```
      context
-------------------
 postmaster
```

requires restart of server
e.g. archive_mode, max_connections

Setting Parameters - context: sighup

```
      context
-------------------
```

sighup — *requires reload of configuration file*
e.g. archive_command, max_wal_size

Setting Parameters - context: superuser-backend

context	
superuser-backend	requires reload of configuration file doesn't change in a session after it has started can be set by superuser

Setting Parameters - context: backend

```
      context
------------------
```

backend

requires reload of configuration file
doesn't change in a session after it has started

Setting Parameters - context: superuser

context	
superuser	can be set by superuser for the server or a session will affect existing sessions

Setting Parameters - context: **user**

context	
user	can be set for the server any user can set in their own session e.g. search_path, work_mem, random_page_cost

Information about Parameters

- Postgres documentation
- Default postgresql.conf
- pg_settings view

Info about Parameters: Postgres Documentation

https://www.postgresql.org/docs/current/runtime-config.html

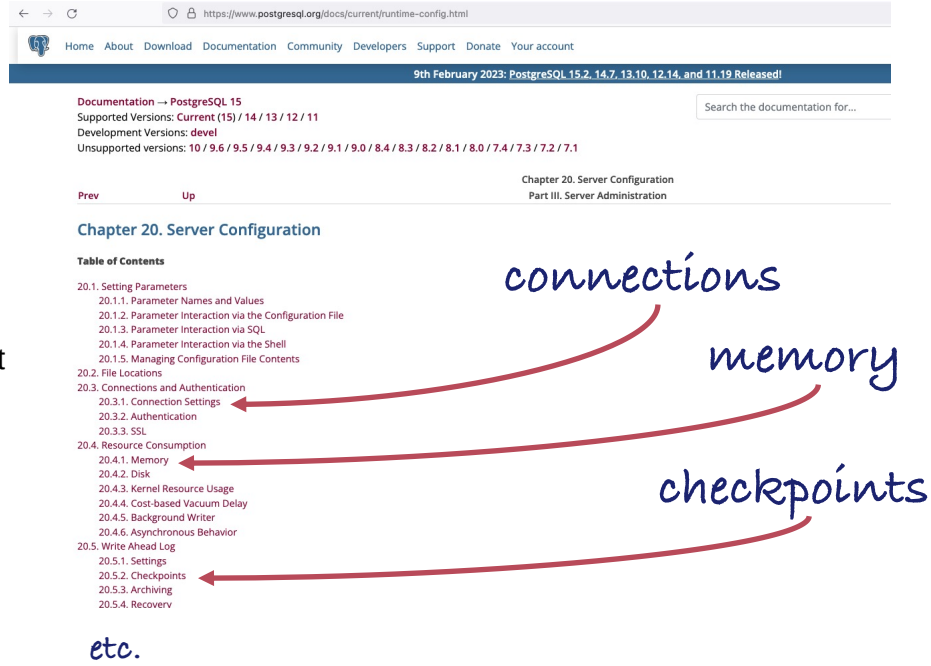
20.3.1. Connection Settings

...

max_connections (integer)

Determines the maximum number of concurrent connections to the database server. The default is typically 100 connections, but might be less if your kernel settings will not support it (as determined during initdb). This parameter can only be set at server start.

When running a standby server, you must set this parameter to the same or higher value than on the primary server. Otherwise, queries will not be allowed in the standby server.



Info about Parameters: postgresql.conf

```
# -----------------------------
# PostgreSQL configuration file
# -----------------------------
#
# This file consists of lines of the form:
#
#   name = value
#
# (The "=" is optional.)  Whitespace may be used.  Comments are introduced with
# "#" anywhere on a line.  The complete list of parameter names and allowed
# values can be found in the PostgreSQL documentation.
#
# The commented-out settings shown in this file represent the default values.
# Re-commenting a setting is NOT sufficient to revert it to the default value;
# you need to reload the server.
#
# This file is read on server startup and when the server receives a SIGHUP
```

Info about Parameters: postgresql.conf

default values

```
#------------------------------------------------------------------------------
# CONNECTIONS AND AUTHENTICATION
#------------------------------------------------------------------------------

# - Connection Settings -

#listen_addresses = 'localhost'         # what IP address(es) to listen on;
                                        # comma-separated list of addresses;
                                        # defaults to 'localhost'; use '*' for all
                                        # (change requires restart)
#port = 5432                            # (change requires restart)
max_connections = 100                   # (change requires restart)
#superuser_reserved_connections = 3     # (change requires restart)
#unix_socket_directories = '/tmp'       # comma-separated list of directories
                                        # (change requires restart)
#unix_socket_group = ''                 # (change requires restart)
```

Info about Parameters: pg_settings

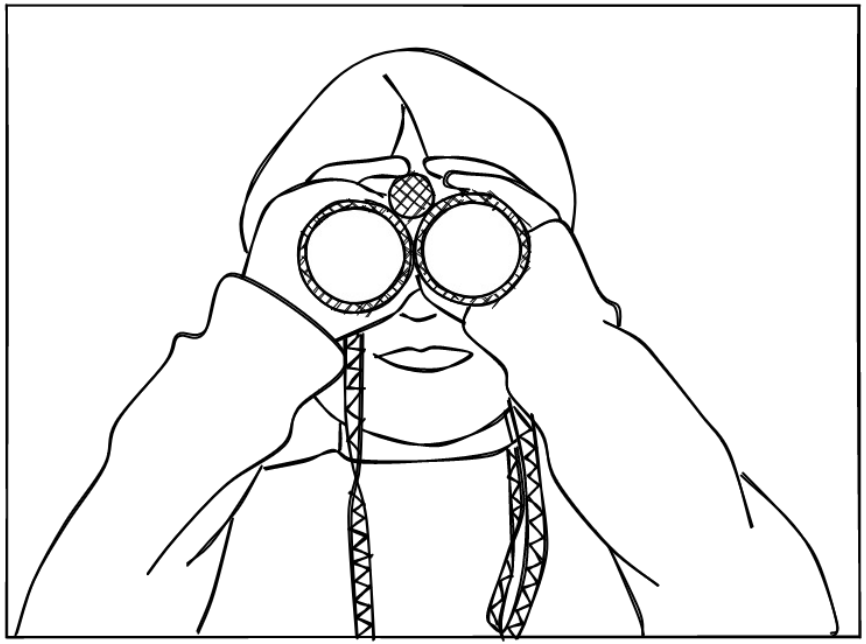
https://www.postgresql.org/docs/current/view-pg-settings.html

```
postgres=# SELECT name, context, unit, min_val, max_val
postgres-# FROM pg_settings
postgres-# WHERE name IN ('max_connections','shared_buffers','work_mem');

       name       |  context   | unit | min_val |  max_val
------------------+------------+------+---------+------------
 max_connections  | postmaster |      | 1       | 262143
 shared_buffers   | postmaster | 8kB  | 16      | 1073741823
 work_mem         | user       | kB   | 64      | 2147483647
```

Viewing Parameter Values

- SHOW parameter
- pg_settings view



Viewing Parameter Values: SHOW parameter

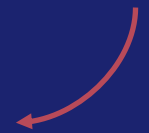
```
postgres=# SHOW max_connections;
 max_connections
-----------------
 200
(1 row)
```

```
postgres=# SHOW work_mem;
 work_mem
----------
 2MB
(1 row)
```

Viewing Parameter Values: pg_settings

```
postgres=# select * from pg_settings where name='max_connections';
-[ RECORD 1 ]---+-----------------------------------------------------
name            | max_connections
setting         | 100
context         | postmaster
vartype         | integer
source          | configuration file
boot_val        | 100
reset_val       | 100
sourcefile      | /Users/karen/homebrew/var/postgres/postgresql.conf
sourceline      | 65
pending_restart | f
```

postgresql.auto.conf
if changed via ALTER SYSTEM



Viewing Parameter Values: pg_settings

```
postgres=# select * from pg_settings where name='work_mem';
-[ RECORD 1 ]---+-----------------------------------------------------
name            | work_mem
setting         | 4096
unit            | kB
context         | user
source          | default
boot_val        | 4096
reset_val       | 4096
sourcefile      |
sourceline      |
pending_restart | f
```

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - Query Tuning
- Summary

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - **Connections/Sessions**
  - Memory
  - Logging
  - WAL
  - Query Tuning
- Summary

Connection parameters: listen_addresses

https://www.postgresql.org/docs/current/runtime-config-connection.html#GUC-LISTEN-ADDRESSES

```
description     | Sets the host name or IP address(es) to listen to.
suggested_value | '*'
context         | postmaster
default_value   | localhost
```

If you want to accept connection requests from all available IP interfaces

- "Is the server running on that host and accepting TCP/IP connections?"
- 0.0.0.0 listen for all IPv4 addresses
- :: listen for all IPv6 addresses
- Control who can connect via pg_hba.conf

Connection parameters: max_connections

https://www.postgresql.org/docs/current/runtime-config-connection.html#GUC-MAX-CONNECTIONS

```
description     | Sets the maximum number of concurrent connections.
suggested_value | no more than 500
context         | postmaster
default_value   | 100
```

May need to be much lower (<100) on a small system

- If max_connections is exceeded:
  `failed: FATAL:  sorry, too many clients already`
- Think about connection pooling above a few hundred connections

Session parameters:

idle_in_transaction_session_timeout

Session parameters:

idle_in_transaction_session_timeout

https://www.postgresql.org/docs/current/runtime-config-client.html#GUC-IDLE-IN-TRANSACTION-SESSION-TIMEOUT

```
description      | Maximum allowed idle time between queries, when in a transaction.
suggested_value  | 30 minutes
unit             | ms
context          | user
default_value    | 0
```

i.e. disabled

- Sessions that are in a transaction, but waiting for a query hold on to locks and block vacuum

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - **Memory**
  - Logging
  - WAL
  - Query Tuning
- Summary

Memory parameters: shared_buffers

https://www.postgresql.org/docs/current/runtime-config-resource.html#GUC-SHARED-BUFFERS

```
description     | number of shared memory buffers used by the server.
suggested_value | 25% - 40% available memory
unit            | 8kB
context         | postmaster
default_value   | 16384
```

i.e. 128MB

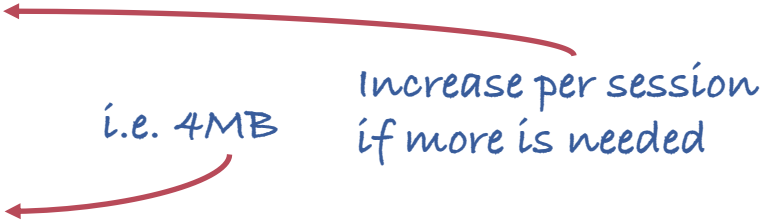
Maybe lower on systems with < 1GB RAM

- Hard allocation at startup
- Benchmark to get right setting

Memory parameters: work_mem

https://www.postgresql.org/docs/current/runtime-config-resource.html#GUC-WORK-MEM

```
description     | maximum memory to be used for query workspaces.
suggested_value | 10MB
unit            | kB
context         | user
default_value   | 4096
```



- Increase work_mem for sessions that perform large sort or hash operations
- Set log_temp_files=0 to check if operations are spilling to disk
- TAKE CARE! e.g. 50 users x 4 sorts x 10MB work_mem ≅ 2GB

Memory parameters: maintenance_work_mem

https://www.postgresql.org/docs/current/runtime-config-resource.html#GUC-MAINTENANCE-WORK-MEM

```
description     | maximum memory to be used for maintenance operations.
suggested_value | ~5% RAM
unit            | kB
context         | user
default_value   | 65536
```

- Higher value can improve performance of maintenance tasks
- Note: autovacuum will use up to 3 x maintenance_work_mem
  (with default values for autovacuum_work_mem & autovaccum_max_workers)

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - Memory
  - **Logging**
  - WAL
  - Query Tuning
- Summary

Logging parameters: log_min_duration_statement

https://www.postgresql.org/docs/current/runtime-config-logging.html#GUC-LOG-MIN-DURATION-STATEMENT

```
description     | min execution time above which all statements will be logged.
suggested_value | 1s
unit            | ms
context         | superuser
default_value   | -1
```

or whatever counts as "too long" in your system

i.e. disabled

- Helps track down unoptimized queries
- Consider setting just for debugging

Logging parameters: log_line_prefix

https://www.postgresql.org/docs/current/runtime-config-logging.html#GUC-LOG-LINE-PREFIX

```
description     | Controls information prefixed to each log line.
suggested_value | '%t:%r:%u@%d:[%p]: '
context         | sighup
default_value   | %m [%p]
```

timestamp
host connecting from
db user name
database connecting to
PID

- Add who, what, where, when etc.
- more information per log line = easier debugging

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - Memory
  - Logging
  - **WAL**
  - Query Tuning
- Summary

WAL/Checkpoint parameters: wal_buffers

https://www.postgresql.org/docs/current/runtime-config-wal.html#GUC-WAL-BUFFERS

```
description     | Sets the number of disk-page buffers in shared memory for WAL.
suggested_value | 32MB
unit            | 8kB
context         | postmaster
default_value   | -1
```

Calculated automatically as:
~3% of shared_buffers (up to max 16MB)

- Amount of memory available for WAL before it's synced to disk
- Increase value if you have many concurrent connections

WAL/Checkpoint parameters: checkpoint_timeout

https://www.postgresql.org/docs/current/runtime-config-wal.html#GUC-CHECKPOINT-TIMEOUT

```
description     | Sets the maximum time between automatic WAL checkpoints.
suggested_value | 10 to 30 minutes
unit            | s
context         | sighup
default_value   | 300
```

i.e. 5 mins

- Checkpoints triggered by timeout = predictability
- Checkpoints are expensive and IO intensive
- Longer timeout = slower crash recovery & more space needed for WAL files

WAL/Checkpoint parameters: max_wal_size

https://www.postgresql.org/docs/current/runtime-config-wal.html#GUC-MAX-WAL-SIZE

```
description     | Sets the WAL size that triggers a checkpoint.
suggested_value | 1/2 to 2/3 disk space in WAL directory
unit            | MB
context         | sighup
default_value   | 1024
```

i.e. 1GB = 64 WAL files

But how big should the WAL directory be?!

- Prevents WAL directory from filling
- Monitor your logs for checkpoints triggered by max_wal_size

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - **Query Tuning**
- Summary

Query tuning parameters: effective_cache_size

https://www.postgresql.org/docs/current/runtime-config-query.html#GUC-EFFECTIVE-CACHE-SIZE

```
description     | Sets planner's assumption about total size of data caches.
suggested_value | 50% to 70% total memory
unit            | 8kB
context         | user
default_value   | 524288
```

i.e. 4GB

leave at least enough RAM for shared_buffers + 5% for OS

- NOT a memory allocation
- Guideline for the query planner
- Higher value = planner more likely to use indexes to speed up query

Query tuning parameters: random_page_cost

https://www.postgresql.org/docs/current/runtime-config-query.html#GUC-RANDOM-PAGE-COST

```
description     | estimate of the cost of a nonsequentially fetched disk page.
suggested_value | 1.1 for SSD, 2.0 for fast spinning disks
context         | user
default_value   | 4
```

- Indicates cost to seek random disk page (as multiple of sequential read)
- Lower value makes planner prefer index scans
- Set lower for fast disks, especially SSD

Agenda

- Setting and Viewing Parameters
- **Most Important Parameters**
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - Query Tuning
  - **Don’t Touch!**
- Summary

DON’T TOUCH!

- fsync
- autovacuum

DON'T TOUCH: fsync

https://www.postgresql.org/docs/current/runtime-config-wal.html#GUC-FSYNC

```
short_desc    | Forces synchronization of updates to disk.
extra_desc    | The server will use the fsync() system call in several
  places to make sure that updates are physically written to disk.
  This insures that a database cluster will recover to a consistent
  state after an operating system or hardware crash.

#fsync = on                    # flush data to disk for crash safety
                               # (turning this off can cause
                               # unrecoverable data corruption)
```

synchronous_commit

switching this off is **slightly** less scary!

DON'T TOUCH: autovacuum

https://www.postgresql.org/docs/15/runtime-config-autovacuum.html#AUTOVACUUM

```
description     | Starts the autovacuum subprocess.
suggested_value | on
context         | sighup
default_value   | on
```

- "optional but highly recommended"
- Executes VACUUM and/or ANALYZE as needed
- Set `log_autovacuum_min_duration=0` to monitor autovacuum

Agenda

- Setting and Viewing Parameters
- Most Important Parameters
  - Connections/Sessions
  - Memory
  - Logging
  - WAL
  - Query Tuning
  - Don’t Touch!
- **Summary**

Summary

Connections/Sessions

- listen_addresses
- max_connections
- Idle_in_transaction_session_timeout

Memory

- shared_buffers
- work_mem
- maintenance_work_mem

Logging

- log_min_duration_statement
- log_line_prefix

WAL

- wal_buffers
- checkpoint_timeout
- max_wal_size

Query Tuning

- effective_cache_size
- random_page_cost

Don’t Touch!

- fsync
- autovacuum

Summary – the lazy version

- shared_buffers
- work_mem
- maintenance_work_mem
- wal_buffers
- effective_cache_size

Conclusions

- PostgreSQL really does Just Work™
- Tune 13 parameters (and leave 2 alone)
- Suggested values are starting points, not directives
- Leave your database to look after itself

Thank You!

Karen Jex | @karenhjex | karen.jex@crunchydata.com

Image acknowledgements

- Elephant with balloon: Jan-Mikael Stjernberg at Pixabay
- Binoculars: Based on image by nightowl at Pixabay
- Reading: Based on image by Лариса Мозговая at Pixabay
- Danger: Based on image by 8ocho8 at Pixabay