Postgres monitoring review checklist

This document describes what every good Postgres monitoring must or should have. The document contains two sections: first, we describe the actual “monitoring” part, and then overview the list of alerts that are worth having.

> 👉 Follow-up doc: [PostgreSQL Troubleshooting & Monitoring – Checklist / Runbook]

Time-series graphs and tables

Color highlighting meanings:

- most important,
- very good to have,
- unclassified (yet),
- new, unclassified
- without highlighting: good to have, may be helpful for troubleshooting, but we can survive without it.

System

- Basic
  - CPU
  - LA
  - Memory usage
  - Memory errors
  - Swap usage
  - Swap I/O
- Disk
  - Read and write latency
  - IOPS (read and write)
  - Throughput (read and write)
  - % util / io usage
  - Disk space
  - Page cache hit/miss
- Network
  - Bandwidth

    - Errors, dropped packets
    - Packets/second
    - Connections (in / out; per status / per IP)
- Process analysis
    - Top-N by CPU load
    - Top-N by RAM load
    - Top-N by disk reads (IOPS and throughput)
    - Top-N by disk writes (IOPS and throughput)
    - Top-N by swap (amount + rates)
    - Top-N by process count and by threads count
    - Top-N by open files

Postgres

- Main
    - Average query time
    - Connections (by `state` in `pg_stat_activity`)
    - Top-N transactions by age
    - Total query time (by query group from `pg_stat_statements`)
    - Buffers hits and reads (by query group from `pg_stat_statements`)
    - Buffers written, dirtied (by query group from `pg_stat_statements`)
    - Transactions (committed vs rolled back)
    - Replication lags
- Autovacuum
    - Workers (by mode: main vs tx ID wraparound)
    - Top-N tables (and DBs) by autovacuum pending work
    - TXs left (tx ID wraparound), database transaction age
    - Autovacuum queue and progress https://gitlab.com/snippets/1889668
- Bloat (estimated)
    - Table bloat (estimated)
    - Index bloat (estimated)
- Checkpoints, bgwriter
    - Disk writes (by processes)
    - Pages dirtied by queries
    - Checkpoints issued
- Connections
    - # of connections by `state` + `max_connections`
    - Connections by client address (active + all)
    - Connections by app (active + all)
    - Connections by DB user
    - Connections by database name
    - Idle in transaction connections

- Locks, waits
  - Total number of locks acquired
  - Locks by time
  - Queries blocked longer than X seconds
  - Deadlocks
  - Types of waits
- Replication
  - Destination (follower)
    - Replication lag in bytes (for followers of the current node)
    - Replication lag in bytes (for the current node compared to the leader)
    - Replication lag in seconds (for the current node compared to the leader)
  - Origin (primary or replica with cascaded replication)
    - Unused replication slots / replication statuses
    - Amount of bytes for each replication slot
    - Replication lag phases (1 graph for each follower)
- Tables
  - Table sizes (total, heap, TOAST, indexes)
  - Estimated rows
  - Top-N by seqscan
  - Top-N by blocks read
  - Top-N by INSERT
  - Top-N by UPDATE
  - Top-N by DELETE
  - Top-N by size (tuples, bytes)
  - Top-N by bloat (estimated!)
  - Top-N by n_dead_tup
- Indexes
  - Index sizes
  - Index usage
  - Not valid indexes
  - Unused indexes
  - Redundant indexes
- Functions
  - Function usage (calls per second)
  - Average time of execution (total, self)
- WAL
  - pg_xlog/pg_wal size
  - Archiver statuses (fail/success)
  - WAL write rates, B/s
  - WAL files count (total, unarchived)
  - WAL directory size

    - WAL files which ready to be archived (count of files in pg_xlog/pg_wal/archive_status which end in ".ready")
- Transactions
    - Transactions per second (TPS)
    - Long-running transactions / max transaction age
- Query macro-analysis based on pg_stat_statements
    - Top-N by total_time
    - Top-N by mean_time
    - Top-N by calls
    - Top-N by CPU usage
    - Top-N by I/O timing
    - Top-N by I/O timing - writes
    - Top-N by block reads (page cache->buffer pool)
    - Top-N by blocks dirtied
    - Top-N by rows
    - Top-N by block hits (buffer pool)
    - Top-N by temporary files generated (bytes; blocks)
    - Top-N by block reads from disk
    - Top-N by block writes
    - --- ability to filter and/or aggregate by dbid (“no filter” also wanted)
    - --- ability to filter and/or aggregate by userid (“no filter” also wanted)
    - --- ability to filter and/or aggregate by “the first word” (SELECT/INSERT/…) (“no filter” also wanted)
    - --- ability to filter and/or aggregate by relations mentioned in query text (“no filter” also wanted)

    - For each query group from the top-N list -- personal graphs showing:
        - `mean_time`
        - `total_time` (?)
        - `calls`
        - block operations
        - rows
        - query stages: CPU, I/O read, I/O write
        - more info: pg_stat_kcache, pg_qualstats, pg_sortstats
- Macro-analysis based on wait events
    - Query groups by wait event types
    - Query groups by wait events
    - Top-N by time spent in wait event (agg. by type)
    - Top-N by time spent in wait event
    - For each query
        - History of the query group: wait event types
        - History of the query group: wait events
    - For each wait event type:

        - History of query groups
    - For each wait event:
        - History of query groups
        - Time spent in each event withing this type
- Log analysis
    - Critical events: restarts, crashes
    - Autovacuum activity
    - Checkpointer activity
    - Locks (>deadlock_timeout)
    - Deadlocks
    - Query examples
    - Plan examples
    - Connections, disconnections
- pgBouncer monitoring
    - From pgbouncer log:
        - Average query time
        - Average transaction time
        - Queries per second (QPS)
        - Transactions per second (TPS)
        - Traffic in and out, B/s
    - Number of connections by client addr
    - Connections between pgBouncer and Postgres, by state
    - Utilization for each pool
    - Waiting clients and waiting time
- Backups
    - ?
    - Time from last successful time (graph)
    - Last backup size (graph)

Alerts

WIP

Critical

- *Disk space (% or in GiB)*
- *Number of connections is close to max_connections (%)*
- *Number of idle-in-transaction connections > N*
- *Inactive replication slots*
- *Replication slot size is > X*

- *Autovacuum workers = autovacuum_max_workers*
- *Transaction ID wraparound risk*
- *Archives failed X times in Y seconds*
- *Long blocking session (e.g. >1min)*
- *Time from last successful backup (e.g. >30hours)*