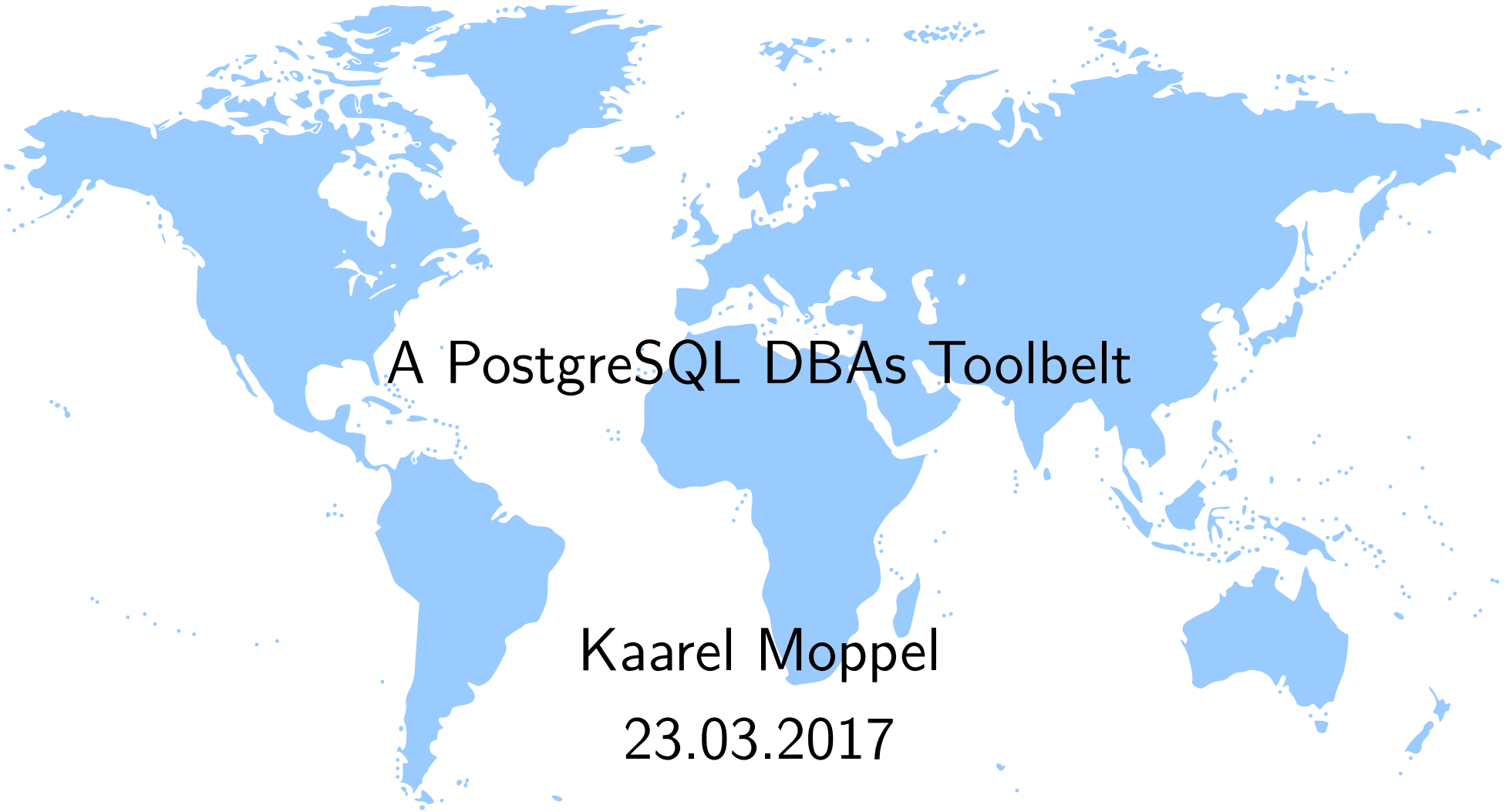

A PostgreSQL DBAs Toolbelt

Kaarel Moppel

23.03.2017

Fields of interest for DBAs

- Setting up Postgres
- Monitoring
- Daily DBA tasks
- Navigating the DB-land
- Data integration
- Developing
- Scaling

Setting up Postgres

First stop before any tools

Tools are great but understanding the concepts is still more important. Means provided by the Postgres project should be quite sufficient in half of the cases.

- Official documentation
- Mailing lists
- wiki.postgresql.org

pgtune

Reads postgresql.conf file and outputs a new, recommended configuration based on your hardware and workload type (Web, OLTP, DW, mixed).

- Online version available - http://pgtune.leopard.in.ua/

More hints about tuning:

- https://wiki.postgresql.org/wiki/Performance_Optimization
- https://wiki.postgresql.org/wiki/Category:Benchmarking

pgbench

Standard tool for quick hardware benchmarking. Bundeled with server binaries.

- Measures throughput in terms of TPS
- TPC-B like simple banking schema
- Custom workloads
- Foreign keys for more realistic numbers
- Fillfactor

pgbench-tools

A wrapper around pgbench to run test sets with different scale/concurrency levels and provides nice summaries/graphs

- Good for finding out the tipping point for concurrent users or cache ratio
- Would be even more useful if one could add the dimension of postgresql.conf settings

Backups

- wal-e - cloud oriented
- pgbarman - incremental backup, continous WAL streaming
- pgbackrest - incremental / differential backups, geared for speed and >TB DBs

Monitoring (ad hoc)

Unix/Linux tooling

Get acquainted with some as not all are available on all platforms - https://wiki.postgresql.org/wiki/Performance_Analysis_Tools

- top / htop
- iotop / dstat

pg_activity

A top like application for PostgreSQL server activity monitoring.
Works both locally and remotely.

- Blocked/blocking queries view
- Terminating backends possible

pg_view

Activity view utility with very detailed system information.

- Mainly for “on the server” usage
- Linux only, but very lightweight
- Out-of-disk estimate alert
- Different output formats

pgcenter

Command line monitoring/management tool that does almost everything.

- "top" features
- pg_stat_* monitoring
- session termination
- server config editing
- logfile access

pg_stat_statements

A “must have” for quick problem detection. Enables ordering of often executed queries by avg. runtime, total time, blocks written/read, IO time.

- Contrib extension
- Performance penalty of couple of percent
- Somewhat secure

Monitoring frameworks

Generic monitoring frameworks

Plugins to generic monitoring frameworks and APM providers - Nagios, Icinga, Munin, New Relic, AppDynamics, etc.

- Comes built in with cloud providers like AWS RDS and Heroku
- Most commercial providers have black-box approach

check_postgres

A Perl script with around 50 checks (actions) to answer most important questions around Postgres - is DB answering, how many sessions, longest/idle tx time, last checkpoint, bloat percentage, etc.

- Integrates with Nagios and some other general systems monitoring frameworks
- Knows quite a bit about Postgres ecosystem (PgBouncer, PgAgent, Slony)
- Can be well used from command line for ad hoc stuff

check_pgactivity

Nagios plugin. Similar to check_postgres but with less checks (services).

- Actively maintained

PgObserver

Remote data collector daemon + webapp for graphing / tables.

- Almost all of pg_stat_* + some custom data gatherers
- Developer friendly - no extensions or special rights, Vagrant / Docker

pgwatch2

A new tool from Cybertec with minimalistic but flexible approach, relies heavily on state-of-the-art components like Grafana and InfluxDB.

- Almost all of pg_stat_* + some custom data gatherers
- User extensible metrics collection with SQL
- Developer friendly - no extensions or special rights, Docker first

pghero

Health checks, suggested indexes, and more

- Seems to be the most popular Postgres monitoring project on Github
- Thresholds and limits hardcoded
- Docker for quick testing

Tools for common tasks

Snippets

There's a ton of snippets from Postgresql Wiki to help with daily work. Might not always be 100% up to date but mostly nicely commented.

- Deleting duplicates
- Fixing sequences
- Finding / deleting dependencies
- Invalid / unused / duplicate indexes
- Unindexed foreign keys
- Bloat

Other script-sets

Activity, biggest tables, indexing problems, bloat, locking, foreign keys, ...

- https://github.com/pgexperts/pgx_scripts
- https://github.com/dataegret/pg-utils

Log analyzing

Kaarel Moppel
23.03.2017

Standard Linux tools

- grep
- awk

pgBadger

Number one log analyzer for Postgres producing nice HTML pages with graphs. Dozens and dozens of options.

- Slowest/most frequent/most time consuming queries
- Sessions / locking / temp files
- TPS / query type distribution
- autodetects “syslog” log format
- parallel/incremental processing
- PgBouncer logfiles support

Other approaches

- Cloud provider plugins - Loggly, Splunk
- redislog extension for the ELK stack
- Importing directly into Postgres tables (CVS format)
- Exposing logfiles via file_fdw

Indexing & bloat

Indexing

- Snippetes from wiki.postgresql.org
- hypopg - extension for building “hypothetical” indexes
- pghero - index suggestions. Uses query parser from https://github.com/lfittl/pg_query

Bloat

Spotting bloat

- pgstattuple extension
- Snippets from wiki.postgresql.org
- check_pgactivity - https://github.com/OPMDG/check_pgactivity
- Somekind of indicators from most monitoring tools

Removing bloat

- VACUUM FULL / CLUSTER
- pg_repack
- pg_squeeze
- pgcompacttable

Working with data

Kaarel Moppel
23.03.2017

Navigating Postgres DBs

- psql
  - .psqlrc (per version)
  - variable interpolation
  - crosstab
- libpq service files (~/.pg_service.conf)
  - psql_switch
- pgcli
  - widely popular and actively developed
  - code completes also functions
  - chokes noticably on big datasets
- niceupdate
  - controlled chunk-by-chunk updates
  - CPU load and space checking

Navigating Postgres DBs graphically

- pgadmin3/4
- DataGrip
- SqlPad
  - save and share queries + basic charting

Data integrations

- Foreign Data Wrappers
  - Your proprietary database / NoSQL store is most probably already supported
  - https://wiki.postgresql.org/wiki/Foreign_data_wrappers
- cstore_fdw
- pgloader - 1-click MySQL migrations, CSV
- https://github.com/begriffs/postgrest
- https://github.com/FRiCKLE/ngx_postgres

Helper tools for developers

- https://explain.depesz.com/
- sqlparse - A non-validating SQL parser module for Python. Supports plpgsql, tokenizing, extracting single statements, formatting with customizations.
- pgFormatter - https://github.com/darold/pgFormatter
- pg_partman / pgslice
- audit-trigger

Helper tools for developers

- plpgsql_check - extra validations for your sprocs
- plprofiler - per line code coverage and performance info
- pldebugger - step through code visually in PgAdmin
- pgTap - unit testing for Postgres
- sqitch - migration management

Architecture

Scaling

- PgBouncer
- Foreign table inheritance
- pgpool-II
- PL/Proxy

Thank you!

Contact us

Cybertec Schönig & Schönig GmbH
Kaarel Moppel
Gröhrmühlgasse 26
A-2700 Wiener Neustadt

www.postgresql-support.de

Follow us on Twitter: @PostgresSupport