PostgreSQL úvod do vývoje

… jak napsat první patch

Tomáš Vondra / EDB
tomas.vondra@postgresql.org

EDB

Proč tenhle workshop?

Agenda

- základní organizace vývoje

  - struktura komunity, commitfesty, nástroje

  - jaký patch napsat / udělat review?

- workflows:

  - základ: git clone, configure, make, tests, ...

  - pokročilé: CF app, github CI, buildfarm, cfbot, ...

- základy/specifika PostgreSQL kódu

  - Datum, Node, memory kontexty, error handling, ...

EDB

Jak komunita funguje?

- vzniklo organicky a stále se vyvíjí

- není dominováno jednou firmou
  - EDB, Crunchy Data, Microsoft, Amazon, Cybertec, data egret, dbi, Dalibo, Fujitsu, Percona, ...
  - typické: firma používající Postgres interně / poskytující služby
  - pravidla proti "dominanci" jedné firmy

- core team (7 lidí, spíš takový "steering committee / BoD")

- committeři (~30 lidí, právo na commit do git repositáře)

- (major) kontributoři, ...

EDB

Kde je komunita?

- e-mailové listy

    - pgsql-general, pgsql-performance, ...

    - pgsql-hackers (https://www.postgresql.org/list/pgsql-hackers/)

- IRC

    - https://www.postgresql.org/community/irc/

- slack, discord, ...

- konference

    - https://www.postgresql.org/about/events/

EDB

Jak je vývoj strukturovaný?

- roční "major" releasy (~září)

  - feature freeze in dubnu

- kvartální "minor" releasy

  - bugfixy, v případě závažné chyby dříve

- commitfesty

  - cyklus: 1 měsíc práce, 1 měsíc pauza

  - "soft" pravidlo (nikdo nebrání pracovat / commitovat kdykoliv)

  - https://commitfest.postgresql.org/

EDB

životní cyklus patchů
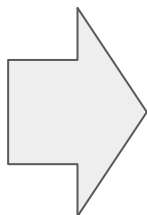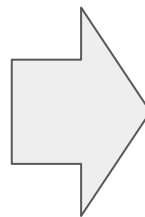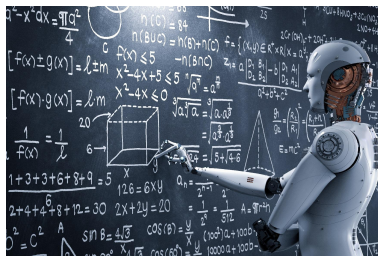
EDB

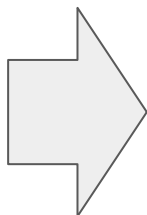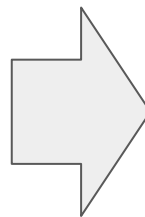# životní cyklus (autorů) patchů

nápad
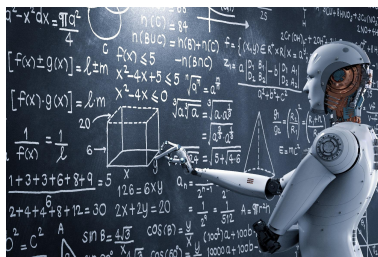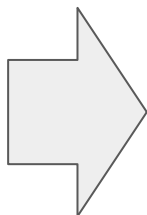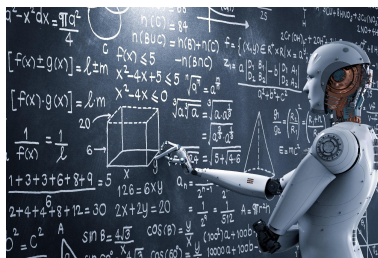
EDB

nápad

psaní patche
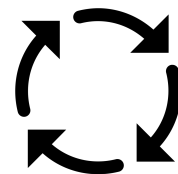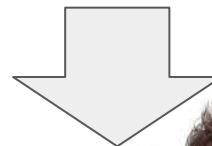
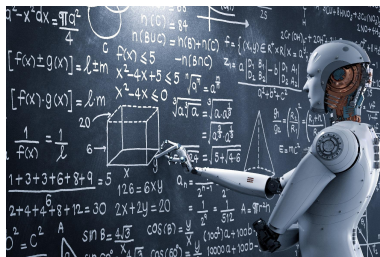EDB

nápad

psaní patche
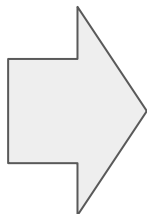
zaslání patche
pgsql-hackers

EDB

nápad
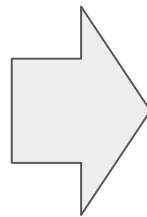
psaní patche

zaslání patche
pgsql-hackers
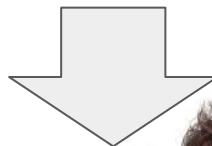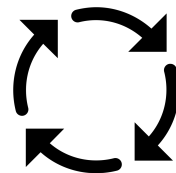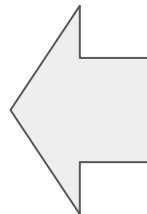
reviews / testing

EDB

nápad

psaní patche

zaslání patche
pgsql-hackers

committed

reviews / testing

EDB

nápad → psaní patche → zaslání patche pgsql-hackers → reviews / testing → committed → buildfarm gets red
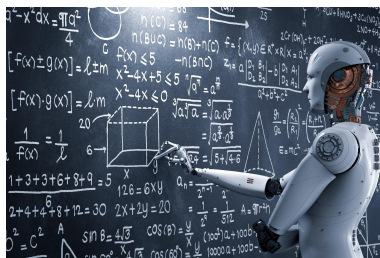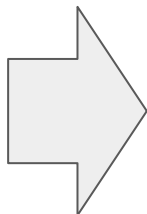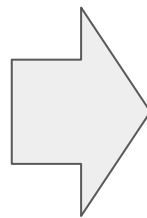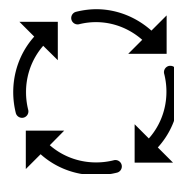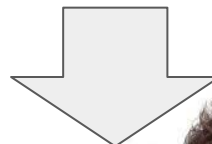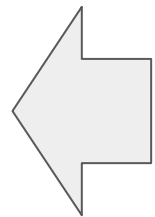
nápad

psaní patche

zaslání patche
pgsql-hackers

buildfarm gets red

committed

reviews / testing

EDB

nástroje

https://git.postgresql.org [link]

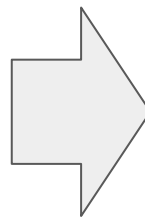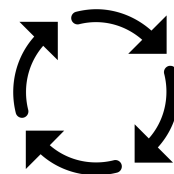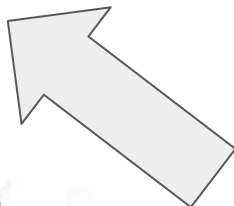| Project	Description	Owner	Last Change				
2ndquadrant_bdr.git	Obsolete repo for Bi-direction...	Abhijit Menon-Sen	7 years ago	summary	shortlog	log	tree
advocacy.git	This repository contains pictu...	Jonathan Katz	11 years ago	summary	shortlog	log	tree
automirror.git	The postgresql.org automirror...	Dave Page	11 years ago	summary	shortlog	log	tree
bucardo.git	Bucardo		7 years ago	summary	shortlog	log	tree
cfbot.git	cfbot repo	Stephen Frost	No commits	summary	shortlog	log	tree
check_postgres.git	check_postgres		6 years ago	summary	shortlog	log	tree
docbot.git	pg_docbot source code, website...	Andreas 'ads' Scherbaum	5 years ago	summary	shortlog	log	tree
fosdem2021-content.git	Public content for the Postgre...	Andreas 'ads' Scherbaum	21 months ago	summary	shortlog	log	tree
fosdem2021-static.git	Static content for the Postgre...	Andreas 'ads' Scherbaum	2 years ago	summary	shortlog	log	tree
hamn.git	This is the code behind http...	Magnus Hagander	2 months ago	summary	shortlog	log	tree
json-datatype.git	Repository for Google Summer...		12 years ago	summary	shortlog	log	tree
libusual.git	Low-level utility library...	Marko Kreen	8 years ago	summary	shortlog	log	tree
pg_adviser.git	Moving the pgAdviser project...	Gurjeet Singh	13 years ago	summary	shortlog	log	tree
pg_bsd_indent.git	Postgres' version of the BSD...	Tom Lane	6 months ago	summary	shortlog	log	tree
pg_filedump.git	pg_filedump provides facilitie...	Christoph Berg	8 weeks ago	summary	shortlog	log	tree
pgapt.git	Code and repository config...	Christoph Berg	26 hours ago	summary	shortlog	log	tree
pgarchives.git	PostgreSQL mailinglist archive...	Magnus Hagander	2 months ago	summary	shortlog	log	tree
pgbench-tools.git	Tools for automating multiple...	Greg Smith	12 years ago	summary	shortlog	log	tree
pgbouncer.git	PgBouncer master repo	Marko Kreen	8 years ago	summary	shortlog	log	tree
pgcac-website.git	PostgreSQL Community Associati...	Dave Page	3 months ago	summary	shortlog	log	tree
pgcommitfest.git	CommitFest management code.	Robert Haas	9 years ago	summary	shortlog	log	tree
pgcommitfest2.git	New generation pg commitfest...	Magnus Hagander	5 months ago	summary	shortlog	log	tree
pgdaychicago-test.git	PGDay Chicago Test Site	Mark Wong	2 days ago	summary	shortlog	log	tree
pgdayuk.git	Websites PGDay UK events.	Dave Page	8 weeks ago	summary	shortlog	log	tree
pgdu2018-test.git	PgDU 2018 Test	Gary Evans	5 years ago	summary	shortlog	log	tree

EDB

pgsql-hackers [https://www.postgresql.org/list/pgsql-hackers/]

CF app [https://commitfest.postgresql.org/]

					13:25	17:01
Make it possible to add custom options to EXPLAIN	Needs review		Konstantin Knizhnik (knizhnik)		1	2023-11-02 13:07	2023-10-21 12:16

Performance

More scalable multixacts buffers and locking	Needs review		Kyotaro Horiguchi (horiguti), Andrey Borodin (x4m), Ivan Lazarev (avaness)		18	2023-10-02 17:34	2023-01-11 17:13
Parallelize correlated subqueries that execute within each worker	Needs review		James Coleman (jcole1989)		11	2023-10-02 14:22	2023-07-04 01:26
In-place persistence change of a relation (fast ALTER TABLE ... SET LOGGED with wal_level=minimal)	Needs review		Kyotaro Horiguchi (horiguti)	Nathan Bossart (bossartn), Justin Pryzby (justinpryzby), Jakub Wartak (jakub.wartak.tomtom)	10	2023-10-02 11:04	2023-09-04 08:37
AcquireExecutorLocks() and run-time pruning	Needs review		Amit Langote (amitlan)	Tom Lane (tgl)	10	2023-10-02 11:03	2023-09-28 08:26
Speed up releasing of locks	Needs review		Andres Freund (andresfreund), David Rowley (davidrowley)	Heikki Linnakangas (heikki)	10	2023-10-02 11:04	2023-11-09 16:18
nbtree performance improvements through specialization on key shape	Needs review		Matthias van de Meent (mmeent)	David Christensen (davidchristensen)	8	2023-10-02 11:05	2023-10-30 16:19
Add sortsupport for range types and btree_gist	Needs review		Christoph Heiss (christoph.heiss)		8	2023-10-02 09:45	2022-11-30 17:25
Improve dead tuple storage for lazy vacuum	Waiting on Author		Masahiko Sawada (masahikosawada), John Naylor (john.naylor)		8	2023-10-02 10:56	2023-10-28 08:56
Reducing planning time when tables have many partitions	Waiting on Author		Yuya Watari (watari)	Thom Brown (thom), Ashutosh Bapat (ashutoshbapat), Andrei Lepikhov (lepikhov)	8	2023-10-02 10:56	2023-09-27 07:28
ALTER TABLE and CLUSTER fail to use a BulkInsertState for toast tables	Waiting on Author	17	Justin Pryzby (justinpryzby)		7	2023-11-07 16:17	2023-11-07 16:17
Reduce timing overhead of EXPLAIN ANALYZE using rdtsc	Needs review	16	Andres Freund (andresfreund), Lukas Fittl (lfittl), David Geier (geidav)	Lukas Fittl (lfittl), Hannu Krosing (hannukrosing)	7	2023-10-02 11:05	2023-02-20 10:36
asynchronous execution support for Custom Scan	Needs review		KaiGai Kohei (kaigai), kazutaka onishi (onishi)		7	2023-10-02 09:46	2022-12-01 23:35
Add semi-join pushdown to postgres_fdw	Needs review		Alexander Pyhalov (a.pyhalov)	Yuki Fujii (fujiiyuki)	7	2023-10-02 11:05	2023-10-31 11:07

buildfarm [https://buildfarm.postgresql.org]

cfbot [http://cfbot.cputube.org/]

PostgreSQL Patch Tester

Here lives an experimental bot that converts email threads that are registered in the Commitfest system into branches on Github, and collates test results from Cirrus CI. Key: ✔ or ✘ = new/recently changed, ✔ or ✘ = stable, ◖ = working.

Ready for Committer

45/4633	ALTER DEFAULT PRIVILEGES only affects objects created by the target role	Laurenz Albe	✉ H	✔✔✔✔✔✔✔
45/3659	Add SPLIT PARTITION/MERGE PARTITIONS commands	Dmitry Koval	⚙ ✉ H	✔✔✔✔✔✔✔
45/3511	Add non-blocking version of PQcancel	Jelte Fennema-Nio	✉ H	✔✔✔✔✔✔✔
45/4632	Clarify that you can create unique constraints on ONLY partitioned_table	Laurenz Albe	✉ H	✔✔✔✔✔✔✔
45/4240	Cross-database SERIALIZABLE safe snapshots	Thomas Munro	✉ H	✔✔✔✔✔✔✔
45/4209	Evaluate arguments of correlated SubPlans in the referencing ExprState	Andres Freund	✉ H	✔✔✔✔✔✔✔
45/4585	Fix a wrong comment in setrefs.c	Richard Guo	✉ H	✔✔✔✔✔✔✔
45/4593	Fix output of zero privileges in psql	Erik Wienhold	✉ H	✔✔✔✔✔✔✔
45/4127	Infinite interval values	Ashutosh Bapat, Joseph Koshakow, jian he	✉ H	✔✔✔✔✔✔✔

EDB

coverage [https://coverage.postgresql.org/]

LCOV - code coverage report

			Hit	Total	Coverage
Current view:	top level				
Test:	PostgreSQL 17devel	**Lines:**	393083	498534	78.8 %
Date:	2023-11-11 19:10:41	**Functions:**	23558	26430	89.1 %
Legend:	Rating: low: < 75 %	medium: >= 75 %	high: >= 90 %		

Directory	Line Coverage ⬍			Functions ⬍	
contrib/adminpack		45.7 %	79 / 173	70.8 %	17 / 24
contrib/amcheck		67.0 %	971 / 1449	94.3 %	50 / 53
contrib/auth_delay		0.0 %	0 / 13	0.0 %	0 / 3
contrib/auto_explain		89.3 %	92 / 103	100.0 %	6 / 6
contrib/basebackup_to_shell		86.5 %	90 / 104	100.0 %	14 / 14
contrib/basic_archive		46.8 %	52 / 111	90.0 %	9 / 10
contrib/bloom		90.1 %	456 / 506	96.7 %	29 / 30
contrib/bool_plperl		100.0 %	11 / 11	100.0 %	5 / 5
contrib/btree_gin		93.3 %	181 / 194	100.0 %	202 / 202
contrib/btree_gist		94.7 %	2071 / 2186	97.6 %	520 / 533
contrib/citext		100.0 %	156 / 156	100.0 %	35 / 35
contrib/cube		85.8 %	939 / 1094	94.2 %	97 / 103
contrib/dblink		86.7 %	935 / 1078	97.4 %	75 / 77
contrib/dict_int		97.4 %	38 / 39	100.0 %	5 / 5
contrib/dict_xsyn		96.2 %	100 / 104	100.0 %	8 / 8
contrib/earthdistance		90.9 %	20 / 22	100.0 %	5 / 5
contrib/file_fdw		76.5 %	247 / 323	90.9 %	20 / 22
contrib/fuzzystrmatch		35.5 %	385 / 1083	87.2 %	41 / 47
contrib/hstore		83.7 %	1490 / 1780	76.7 %	132 / 172
contrib/hstore_plperl		100.0 %	58 / 58	100.0 %	6 / 6
contrib/hstore_plpython		97.0 %	64 / 66	100.0 %	6 / 6
contrib/intarray		89.3 %	1295 / 1450	92.6 %	138 / 149
contrib/isn		76.0 %	381 / 501	88.9 %	40 / 45
contrib/jsonb_plperl		94.3 %	115 / 122	100.0 %	10 / 10

EDB

struktura

adresářová struktura

```
-rw-r--r--  1 user user    1192 Sep 10 01:39 COPYRIGHT
-rw-r--r--  1 user user    4288 Oct 14 02:22 GNUmakefile
-rw-r--r--  1 user user    4288 Oct  7 21:59 GNUmakefile.in
-rw-r--r--  1 user user     277 Oct  7 21:59 HISTORY
-rw-r--r--  1 user user    1875 Oct  7 21:59 Makefile
-rw-r--r--  1 user user    1213 Oct  7 21:59 README
-rw-r--r--  1 user user     721 Oct  7 21:59 README.git
-rw-r--r--  1 user user     365 Oct  7 21:59 aclocal.m4
drwxr-xr-x  2 user user    4096 Oct  7 21:59 config
-rwxr-xr-x  1 user user  583147 Oct 12 21:40 configure
-rw-r--r--  1 user user   86760 Oct 12 21:40 configure.ac
drwxr-xr-x 61 user user    4096 Oct 12 21:40 contrib
drwxr-xr-x  4 user user    4096 Jul  2 16:53 coverage
drwxr-xr-x  3 user user    4096 Oct  7 21:59 doc
-rw-r--r--  1 user user  102720 Oct 14 02:22 meson.build
-rw-r--r--  1 user user    6266 Oct  7 21:59 meson_options.txt
-rw-r--r--  1 user user   14695 Jul  5  2018 messages.mo
drwxr-xr-x 16 user user    4096 Oct 14 02:22 src
```

EDB

adresářová struktura (src)

```
-rw-r--r--  1 user user   173 Jun 14  2020 DEVELOPERS
-rw-r--r--  1 user user  1858 Oct  7 21:59 Makefile
-rw-r--r--  1 user user 35873 Oct 14 02:22 Makefile.global
-rw-r--r--  1 user user 35984 Oct 12 21:40 Makefile.global.in
lrwxrwxrwx  1 user user    31 Oct 14 02:22 Makefile.port -> ../src/makefiles/Makefile.linux
-rw-r--r--  1 user user 15048 Oct 12 21:40 Makefile.shlib
drwxr-xr-x 31 user user  4096 Oct 14 02:23 backend
drwxr-xr-x 25 user user  4096 Oct  7 21:59 bin
drwxr-xr-x  4 user user 20480 Oct 14 02:22 common
drwxr-xr-x  3 user user  4096 Oct 14 02:23 fe_utils
drwxr-xr-x 34 user user  4096 Oct 14 02:22 include
drwxr-xr-x  4 user user  4096 Oct  7 21:59 interfaces
drwxr-xr-x  2 user user  4096 Oct  7 21:59 makefiles
-rw-r--r--  1 user user  1295 Oct  7 21:59 meson.build
-rw-r--r--  1 user user  6885 Oct  7 21:59 nls-global.mk
drwxr-xr-x  6 user user  4096 Oct  7 21:59 pl
drwxr-xr-x  3 user user 20480 Oct 14 02:22 port
drwxr-xr-x  2 user user  4096 Oct  7 21:59 template
drwxr-xr-x 16 user user  4096 Oct  7 21:59 test
drwxr-xr-x  5 user user  4096 Oct 14 02:22 timezone
drwxr-xr-x 10 user user  4096 Oct 13 19:29 tools
drwxr-xr-x  2 user user  4096 Oct  7 21:59 tutorial
```

EDB

https://www.postgresql.org/developer/backend/
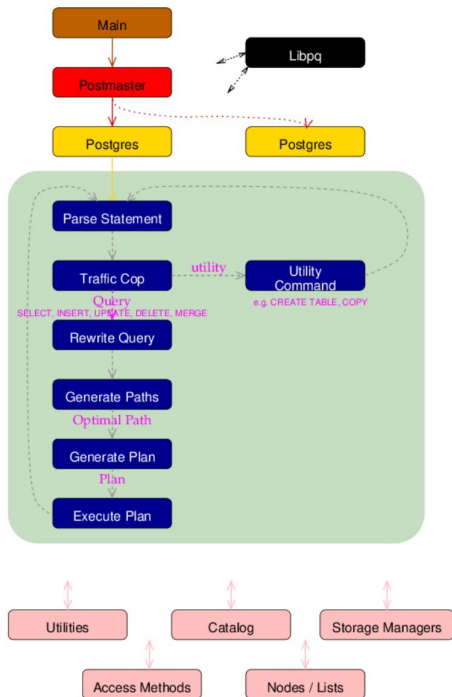
základy práce

- autotools
  - configure + make
  - configure --help
  - doporučené: ccache
- meson (cmake)
  - novější build systém
  - neumí všechno

EDB

git clone

git clone https://github.com/postgres/postgres.git

78 min ago	master	shortlog | log | tree
25 hours ago	REL_16_STABLE	shortlog | log | tree
34 hours ago	REL_15_STABLE	shortlog | log | tree
34 hours ago	REL_12_STABLE	shortlog | log | tree
34 hours ago	REL_13_STABLE	shortlog | log | tree
34 hours ago	REL_14_STABLE	shortlog | log | tree
3 days ago	REL_11_STABLE	shortlog | log | tree
9 months ago	REL9_2_STABLE	shortlog | log | tree

cirrus-ci.org

```
$ ls -la
total 1344
drwxr-xr-x  8 user user   4096 Nov 10 22:49 .
drwxr-xr-x 49 user user   4096 Oct 23 22:23 ..
-rw-r--r--  1 user user   1640 Oct 12 21:40 .cirrus.star
-rw-r--r--  1 user user  24808 Oct 12 21:40 .cirrus.tasks.yml
-rw-r--r--  1 user user   1980 Oct 12 21:40 .cirrus.yml
-rw-r--r--  1 user user    730 Sep 10 01:39 .dir-locals.el
-rw-r--r--  1 user user    183 Oct  7 21:59 .editorconfig
drwxr-xr-x  8 user user   4096 Nov 10 17:37 .git
-rw-r--r--  1 user user  11917 Oct 27 23:11 .git-blame-ignore-revs
-rw-r--r--  1 user user   1535 Oct  7 21:59 .gitattributes
-rw-r--r--  1 user user    674 Oct  7 21:59 .gitignore
-rw-r--r--  1 user user  19979 Nov 10 18:08 0001-v3.patch
-rw-r--r--  1 user user  12911 Nov 10 18:08 0002-review.patch
-rw-r--r--  1 user user  17388 Nov 10 18:08 0003-rework.patch
-rw-r--r--  1 user user   1192 Sep 10 01:39 COPYRIGHT
-rw-r--r--  1 user user   4195 Nov 10 17:41 GNUmakefile
```

EDB

cirrus-ci.org [src/tools/ci/README]

configure

```
configure --help

configure --prefix=$HOME/pgbuild \
    --enable-depend \
    --enable-tap-tests --enable-cassert \
    --with-lz4
```

make targets

```
make

make -s -j4 install

make (check | installcheck)

make check-world

make html

make coverage
```

Typy testů

regresní testy

- jednoduchý SQL skript + očekávaný výstup

izolační testy

- interakce více sessions (ACID, ...)

TAP testy

- složitější skripty
- replikace, pg_dump, pg_upgrade

Valgrind

idiomy

Datum

- abstraktní datový typ
  - nechceme být omezeni/navázáni na C datové typy
  - SQL má jiné datové typy než C
  - chceme umožnit rozšiřitelnost
  - nechceme závislost na 32/64 bitech (změna rozsahů)
- hodnota nebo pointer
  - až 8B hodnotou (64bit platformy)
  - pointer delší a variable-length hodnoty

EDB

PG_FUNCTION_ARGS

- volací konvence

- opět, nechceme být omezeni C (a máme jiné datové typy)

- chceme polymorfismus, default hodnoty, NULL (SQL), ...

```
Datum
random_integer(PG_FUNCTION_ARGS)
{
    PG_RETURN_INT32(42);
}
```

Memory management

- nechceme dělat malloc/free

- museli bychom sledovat všechny alokace (jinak memleak)

- MemoryContext

  - objekt / struktura která seskupuje alokace

  - ideálně alokace se stejnou životností, pro danou oblast kódu

  - lze uvolnit / resetovat jako celek (uvolní všechny alokace)

  - současně cache (malloc může být celkem pomalý)

- efektivní obrana proti memory leakům

- hierarchie memory contextů (top => transakce => request => ...)

Node [src/include/nodes/...]

- systém C struktur, v podstatě třídy

- abstrakce, umožňující polymorfismus

- Node je "base class" a všichni z toho dědí

```
typedef struct Node
{
    NodeTag  type;
} Node;
```

```
typedef struct Query
{
    NodeTag       type;
    ...
    Node     *utilityStmt;
    ...
```

EDB

Error handling

- elog
  - jednodušší na použití
  - primárně "shouldn't happen" případy
  - bez podpory pro překlady
- ereport
  - složitější, různé vychytávky (hinty, detaily, SQL errcode, ...)
  - podpora pro překlady
- PG_TRY / PG_CATCH / PG_END_TRY / PG_RE_THROW
  - víceméně výjimky
  - schopnost reagovat na chyby

EDB

DEMO

jednoduchá extenze

- generování random integerů
    - volací konvence SQL funkcí, Datum

jednoduchá extenze

- generování random integerů
    - volací konvence SQL funkcí, Datum
- generování random stringů
    - alokace paměti apod.

jednoduchá extenze

- generování random integerů
  - volací konvence SQL funkcí, Datum
- generování random stringů
  - alokace paměti apod
- memory kontext
  - vlastní memory kontext, statistiky, uvolnění

EDB

jednoduchá extenze

- generování random integerů
  - volací konvence SQL funkcí, Datum
- generování random stringů
  - alokace paměti apod
- memory kontext
  - vlastní memory kontext, statistiky, uvolnění
- hook / callback
  - jednoduchý executor callback

EDB

inspirace

jak vymyslet jaký patch psát?

- ideálně něco co zajímá / trápí vás
  - potřebujete to v práci, osobní zájem, ...
  - např. pomalý dotaz => vymyslím jak udělat rychleji
- projděte patche v CF aplikaci
  - review na existující patch, pomoc autorovi, související patch, ...
- coverage report
  - části mají nízkou coverage - napíšu pár testů, najdu/opravím chyby
- konkrétní příklad
  - zkusit memory contexty přepnout čistě na mmap()

EDB

zdroje

Zdroje

- Developer FAQ
  https://wiki.postgresql.org/wiki/Developer_FAQ

- Server programming
  https://www.postgresql.org/docs/16/server-programming.html

- Internals
  https://www.postgresql.org/docs/16/server-programming.html

- PostgreSQL 14 Internals / Egor Rogov [PDF, paperback, ...]
  https://postgrespro.com/community/books/internals

- The Internals of PostgreSQL / Hironobu SUZUKI
  for database administrators and system developers
  https://www.interdb.jp/pg/

tomas.vondra@postgresql.org

EDB