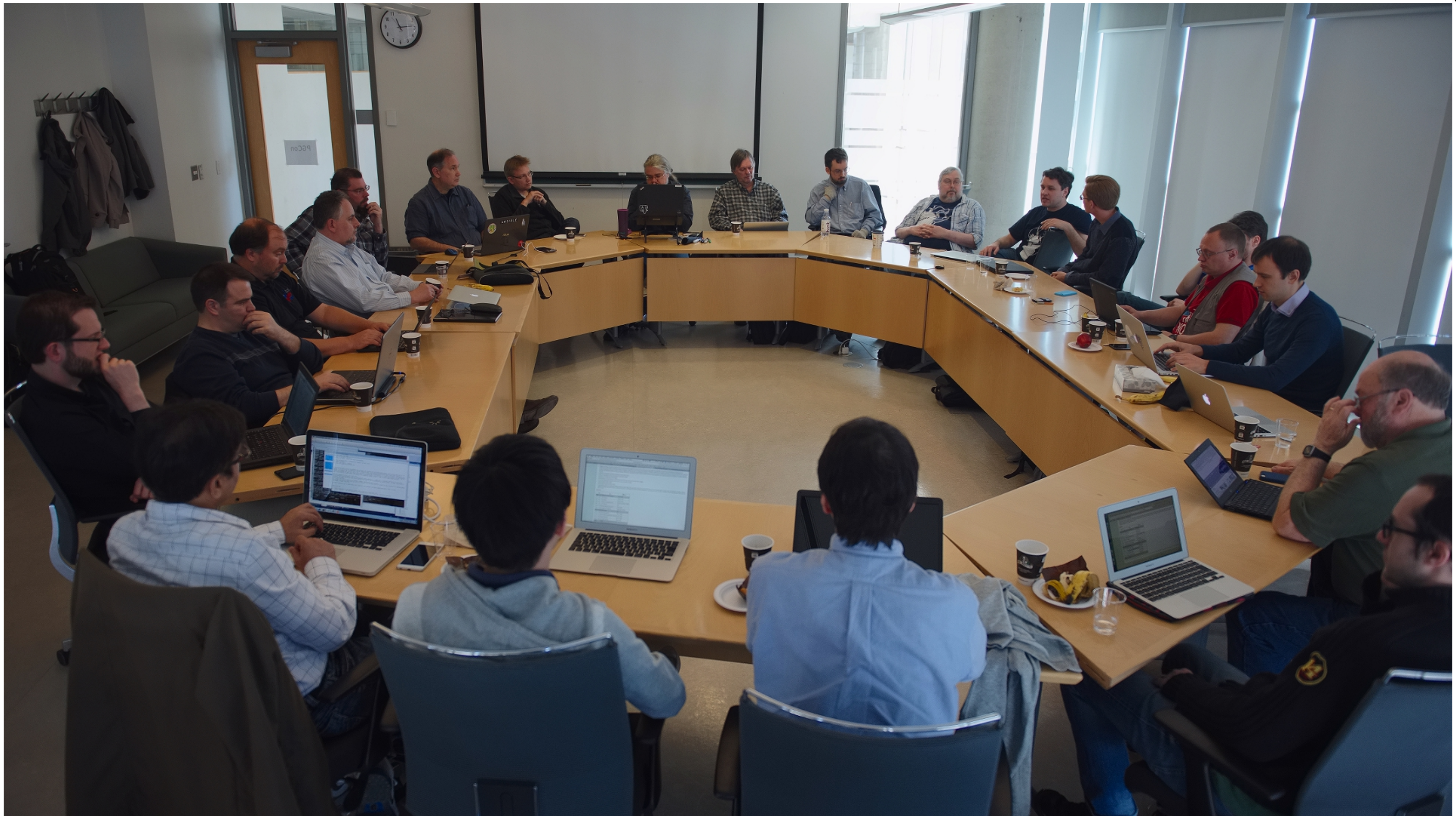
PGCon-2016: Последние новости

Oleg Bartunov, Alexander Korotkov
Postgres Professional, Moscow University

RIT++, June 1, 2017

PostgreSQL Developers Meeting, May 17, 2016, Ottawa, Canada

- Oleg Bartunov, PostgreSQL Professional
- Josh Berkus, Red Hat
- Joe Conway, Crunchy Data
- Jeff Davis, AWS
- Andrew Dunstan
- Peter Eisentraut, 2nd Quadrant
- Andres Freund, Citus Data
- Stephen Frost, Crunchy Data
- Etsuro Fujita, NTT
- Kevin Grittner, EnterpriseDB
- Robert Haas, EnterpriseDB
- Magnus Hagander, Redpill Linpro
- Heikki Linnakangas, Pivotal
- Amit Kapila, EnterpriseDB
- Alexander Korotkov, PostgreSQL Professional
- Tom Lane, Crunchy Data
- Noah Misch
- Dave Page, EnterpriseDB
- Michael Paquier, VMWare
- Simon Riggs, 2nd Quadrant
- Masahiko Sawada, NTT
- Teodor Sigaev, PostgreSQL Professional



- Oleg Bartunov, PostgreSQL Professional
- Josh Berkus, Red Hat
- Joe Conway, Crunchy Data
- Jeff Davis, AWS
- Andrew Dunstan
- Peter Eisentraut, 2nd Quadrant
- Andres Freund, Citus Data
- Stephen Frost, Crunchy Data
- Etsuro Fujita, NTT
- Kevin Grittner, EnterpriseDB
- Robert Haas, EnterpriseDB
- Magnus Hagander, Redpill Linpro
- Heikki Linnakangas, Pivotal
- Amit Kapila, EnterpriseDB
- Alexander Korotkov, PostgreSQL Professional
- Tom Lane, Crunchy Data
- Noah Misch
- Dave Page, EnterpriseDB
- Michael Paquier, VMWare
- Simon Riggs, 2nd Quadrant
- Masahiko Sawada, NTT
- Teodor Sigaev, PostgreSQL Professional

Developers meeting

Agenda

Time	Item	Presenter
09:00 - 09:10	*Welcome and introductions*	*Dave Page*
9:10 - 9:25	Contributors	Robert Haas
9:25 - 9:40	Core team	Robert Haas
9:40 - 9:55	Release management team	Robert Haas/Simon Riggs
9:55 - 10:05	9.7 Schedule	All
10:05 - 10:20	Shared testing infrastructure	Amit Kapila
10:20 - 10:30	Postgres Pro Roadmap	Alexander Korotkov
10:30 - 10:45	*Coffee break*	*All*
10:45 - 11:00	Sharding/clustering	Simon Riggs
11:00 - 11:15	Cross-node transactional consistency	Simon Riggs
11:15 - 11:30	Logical replication	Simon Riggs
11:30 - 12:00	Any other business	Dave Page
12:00	*Lunch*	

Version Numbering

*   8 голосов за X.Y.Z
* 13 голосов за X.Y
   X меняется каждый год
   Y — минорный релиз
* 9.7 → 10.0

- Как номинировать на Major contributors
  Команда волонтеров займется этим
  (Josh Stephen, Joe, Dave)
  - Emeritus list для долгонеактивных
- Core team медленно меняется
  - Нет ясного механизма
- RMT — хорошая практика, так держать
  - За качество отвечают коммитеры !
- Next Release Schedule
  * CF1: September 1 to 30th
  * CF2: November 1 to 30th
  * CF3: January 1 to 31st.
  * CF4: March 1 to 31st.
  * Feature Freeze: March
  * 31st May 16th: Beta
- Нужен общий пул железа для тестирования
- Спор вокруг шардинга
  - FDW sharding ? EDB отреклась - это все Брюс, это его проект.
  - Distributed transactions в ядро ?
- Postgres-XL и Postgres Professional будут вести отдельные сессии на Developers Unconference
- Pglogical форкнут к 1 августу для включения в ядро
- CFM — Commitfest Management Team (2-3 человека) на каждый CF за три недели до старта (CORE team задача)

https://wiki.postgresql.org/wiki/PgCon_2016_Developer_Meeting

PostgreSQL Roadmap

- На самом деле планов развития у постгреса нет

  https://www.postgresql.org/developer/roadmap/

  "PostgreSQL is a non-commercial, all volunteer, free software project, and as such there is no formal list of feature requirements required for development. We really do follow the mantra of letting developers scratch their own itches."
- Однако, постгрес давно используется большим бизнесом и вокруг него сложились компании, которые занимаются поддержкой, разработкой как самого постгреса, так и приложений. Всем нужны планы !
- В России сложились специфические условия для развития постгреса, нас даже заметил Оракл :) Пришло много бизнеса, много новых людей.
- Поэтому Postgres Professional поставил вопрос о планах

PostgreSQL Roadmap

https://goo.gl/eCZOdX

On 04/12/2016 01:07 PM, Oleg Bartunov wrote:
>
> Our roadmap http://www.postgresql.org/developer/roadmap/ is the problem.
> We don't have clear roadmap and that's why we cannot plan future feature
> full release.

As someone who's worked at multiple proprietary software companies,
having a roadmap doesn't magically make code happen.

> There are several postgres-centric companies, which have
> most of developers, who do all major contributions. All these companies
> has their roadmaps, but not the community. I think 9.6 release is
> inflection point, where we should combine our roadmaps and release the
> one for the community. Than we could plan releases and our customers
> will see what to expect. I can't say for other companies, but we have
> big demand for many features from russian customers and we have to
> compete with other databases. Having community roadmap will helps us to
> work with customers and plan our resources.

It would be good to have a place for the companies who do PostgreSQL
feature work would publish their current efforts and timelines, so we at
least have a go-to place for "here's what someone's working on". But
only if that information is going to be *updated*, something we're very
bad at. And IMHO, a "roadmap" which is less that 50% accurate is a
waste of time.

There's an easy way for you to kick this off though: have PostgresPro
publish a wiki page or Trello board or github repo or whatever with your
roadmap and invite other full-time PostgreSQL contributors to add their
pieces.

Developers meeting (Roadmaps)

Postgres Professional Roadmap

1 Multimaster cluster with sharding
2 JIT-compilation of queries
3 Pluggable storages
4 Effective partitioning
5 Adaptive query planning
  5.1 Query plans management
  5.2 Machine learning
  5.3 Execution-time planning
6 Backup
  6.1 Block-level incremental backup
  6.2 Backup validation
  6.3 Partial backup and partial restore
7 Connection pooling
8 Page-level data compression
9 Wait events monitoring
10 Better temporary tables
  10.1 Temporary tables outside of system catalog
  10.2 Temporary tables on standbys
11 Native querying for jsonb with indexing support

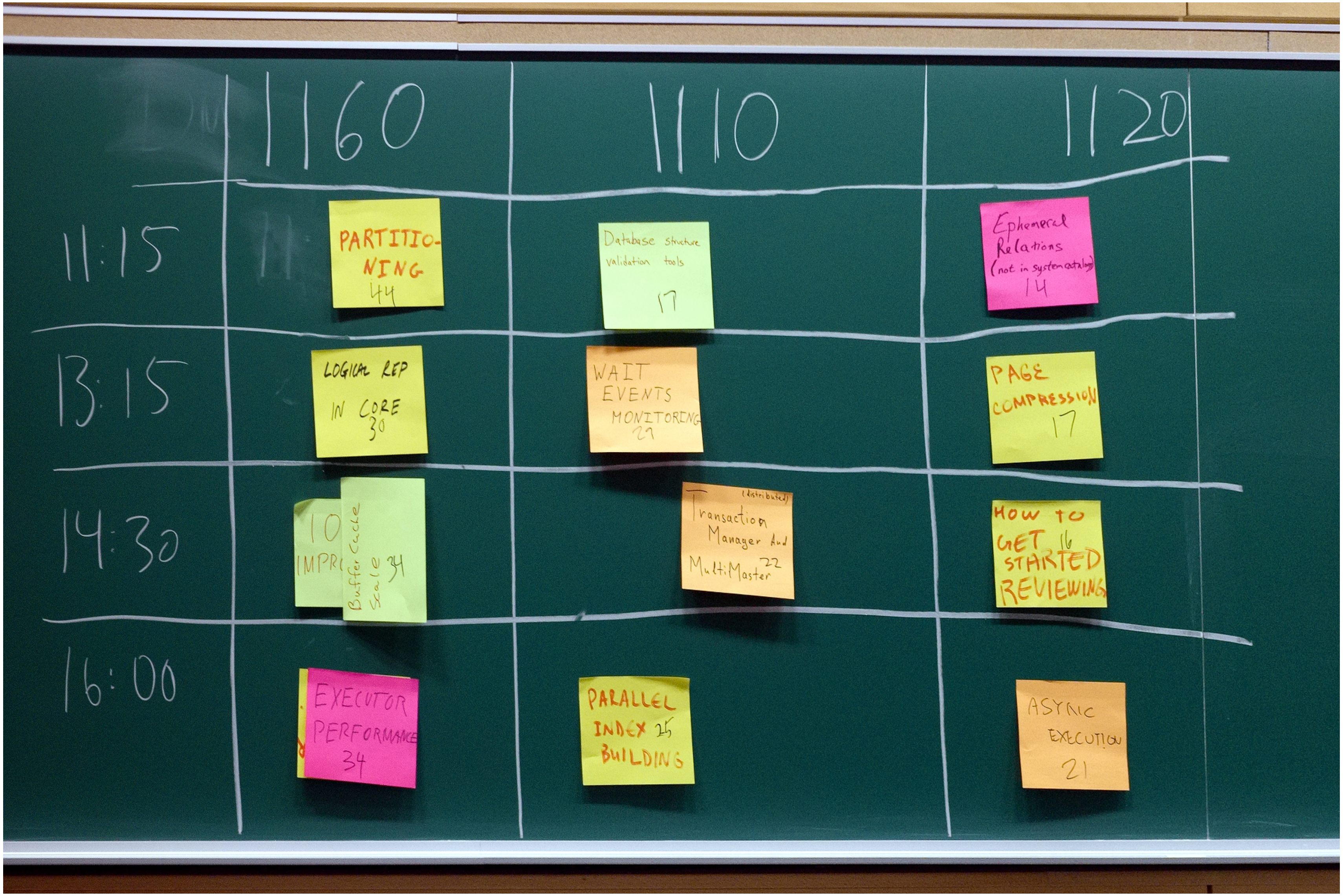
EnterpriseDB database server roadmap

1 Parallelism
  Hash node
  Bitmap Heap Scan
  Index Scan
  Vacuum
2 Partitioning
  Declarative partitioning
3 Foreign Data Wrappers
  Aggregate pushdown to postgres_fdw
  Async execution to executor
  Pluggable API for heap in core
4 Replication
  Causal reads
5 Vertical Scalability
  Speedup CLOG access
  Cache MVCC snapshots-reduce contention ProcArrayLock
6 Performance
  Asynchronous, Vectorized execution
  Hash indexes
7 SQL Features

https://wiki.postgresql.org/wiki/Postgres_Professional_roadmap
https://wiki.postgresql.org/wiki/EnterpriseDB_database_server_roadmap

Developers Unconference

5 sessions out of 12 by Postgres Professional



https://wiki.postgresql.org/wiki/PgCon_2016_Developer_Unconference

Developers Unconference

5 sessions out of 12 by Postgres Professional

Unconference schedule by Time and Room

Time	DMS1160	DMS1110	DMS1120
Wed 11:15-12:15	Partitioning	Database schema validating tools	Ephemeral relations
Wed 13:15-14:15	Logical replication in core	Wait events monitoring	Page compression
Wed 14:30-15:30	IO improvements + buffer cache scaling	Transaction manager + multimaster	How to get started reviewing
Wed 16:00-17:00	Executor performance	Parallel index building	Async execution

https://wiki.postgresql.org/wiki/PgCon_2016_Developer_Unconference

Конференция - 1

Parallel Query Has Arrived!
Where We've Been, Where We Are, and Where We're Going
Robert Haas
Track **9.6 Features**

Yandex.Mail success story
Experience of a big migration from Oracle to PostgreSQL
Vladimir Borodin
Track **DBA**

In-memory column store
Hari Babu
Track **Hacking**

ERROR: snapshot too old
(why having the option of seeing that is a good thing)
Kevin Grittner
Track **9.6 Features**

Unveiling pg_global_kv
A distributed key value store built on top Postgres

Index Internals
Heikki Linnakangas
Track **Hacking**

B-Tree - explore the heart of PostgreSQL.
Anastasia Lubennikova
Track **Hacking**

Autovacuum, explained for engineers
Ilya Kosmodemiansky
Track **DBA**

pg_paxos: Paxos-based table replication
Bringing the distributed c...
PostgreSQL
Marco Slot
Track **Scalin...**

PostgreSQL High-Availability and Geographic Locality using consul...
The website may be down, but the database is not.
Sean Chittenden
Track **Scaling Out**

Run Simple Query Faster
Dilip Kumar
Track **Performance**

FTS is dead ? Long live FTS !
New features of Full-text search in PostgreSQL
Oleg Bartunov, **Teodor Sigaev**
Track **Hacking**

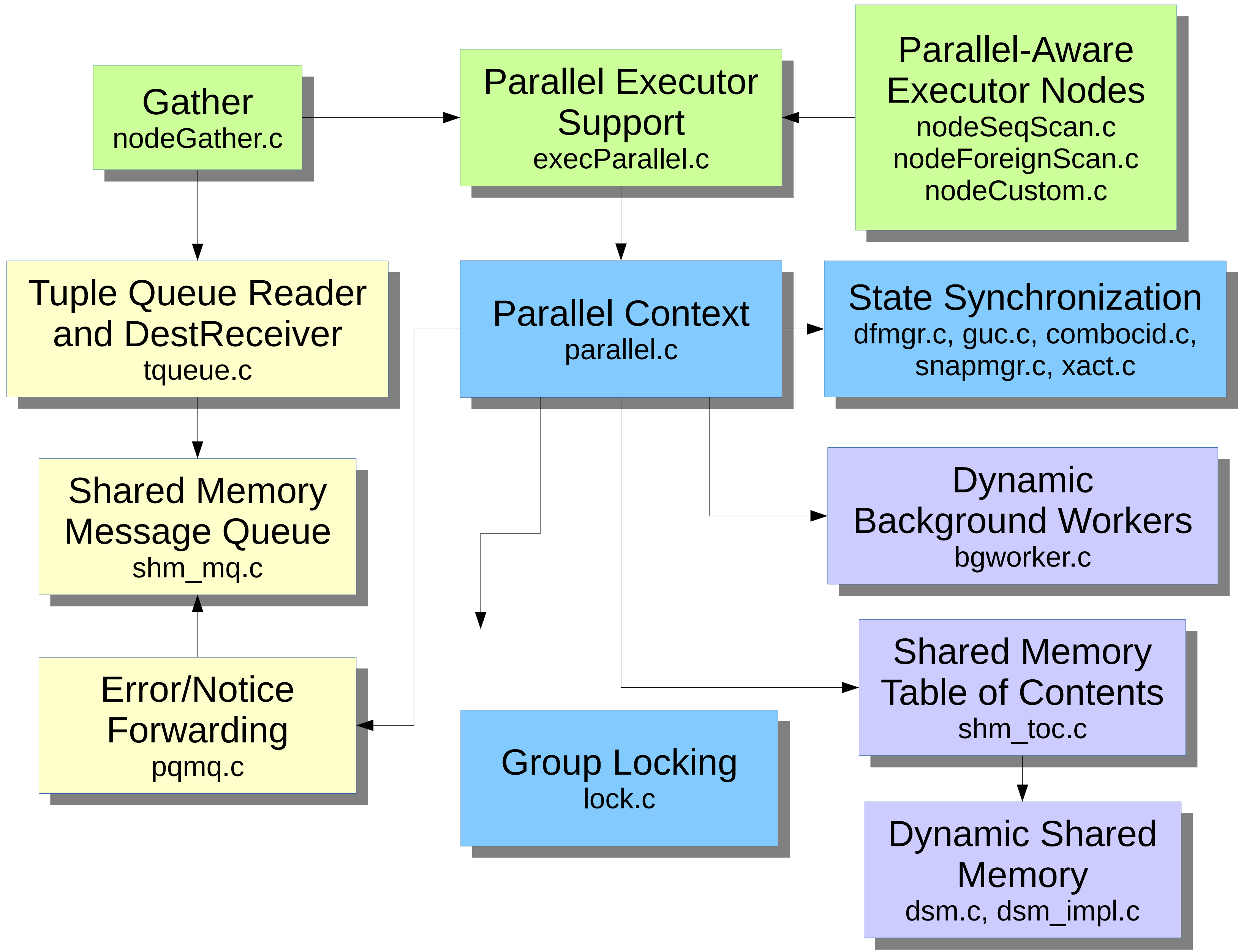
9.4 The Hard Way
Version upgrades with pg_upgrade and rsync using hard links
James Miller, **Samantha Billington**
Track **DBA**

PostgreSQL extendability: Origins and new horizons
Towards pluggable storage engines
Alexander Korotkov
Track **Hacking**

Experienc... with Postg... PostgreSQ...
Research an... with Postgre...
Gilberto Cas...
Track **Case S...**

Lightning talks
Short sharp descriptions of short topics
Magnus Hagander
Track **Plenary**

Parallel Query Architecture



Источник: http://www.pgcon.org/2016/schedule/events/913.en.html

TPC-H Testing: Results

- With parallel query enabled, 17 plans used parallelism. The other 5 query plans did not change.
- Of the 17 queries that used parallelism, 15 got faster, 1 was unchanged, and 1 got slower. (Increasing work_mem fixed the regression.)
- 3 queries ran at least 4x faster; 11 queries ran at least 2x faster.
- Full results at rhaas.blogspot.com

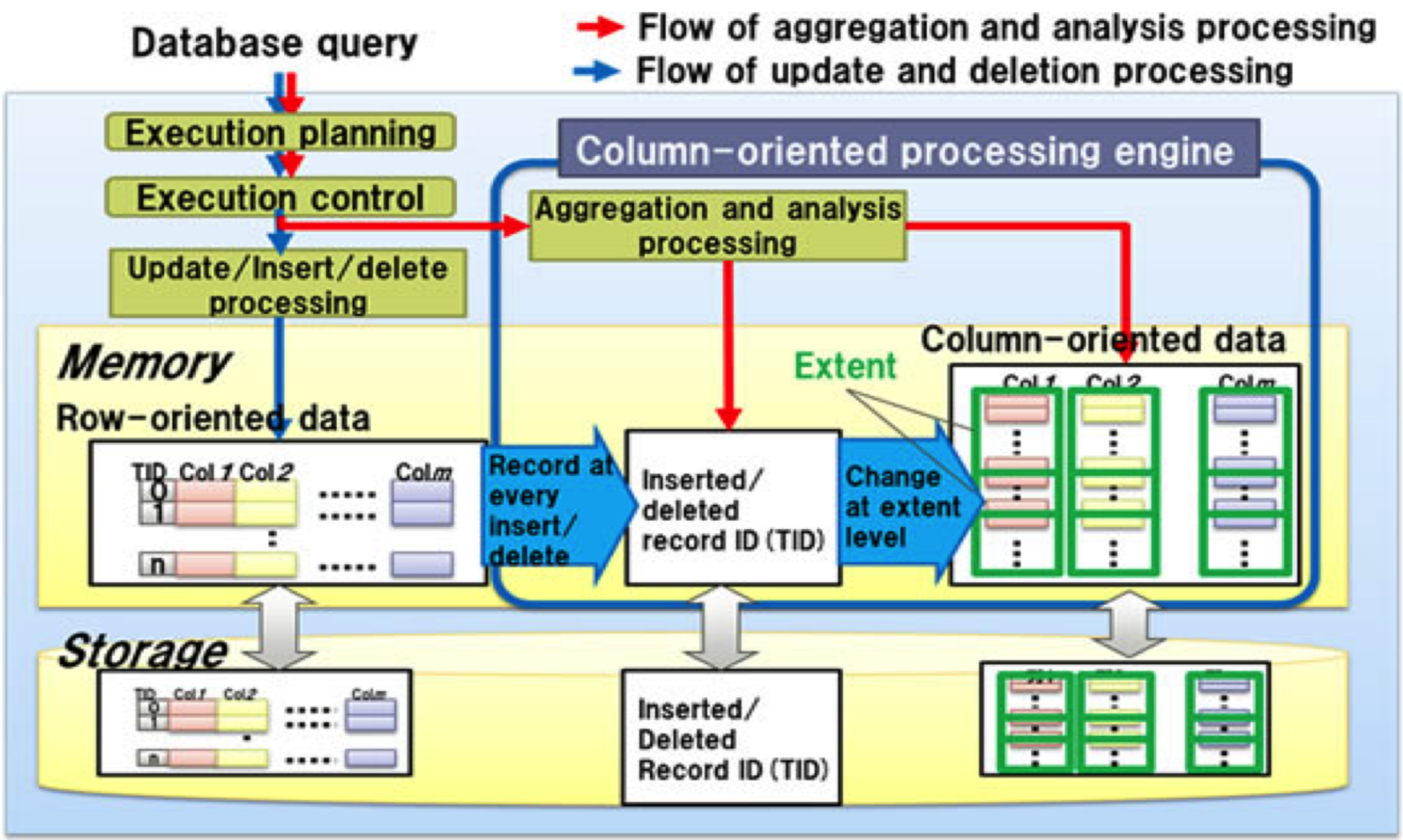
Источник: http://www.pgcon.org/2016/schedule/events/913.en.html

Analysis of Query Plans – What's missing?

- Ideal join strategy is a merge join, which we don't know how to do in parallel: Q2, Q13, Q15
- Need parallel hash table build: Q3, Q5, Q7, Q8, Q21
- Need parallel bitmap index scan: Q4, Q5, Q6, Q7, Q14, Q15, Q20
- Need parallel subquery handling: Q2, Q15, Q16, Q22
- Need gather merge: Q17
- Not sure: Q3, Q9, Q10, Q11, Q18
- Already looks ideal: Q1, Q12, Q19

Источник: http://www.pgcon.org/2016/schedule/events/913.en.html

In-memory column storage by Fujitsu



Источник: http://www.pgcon.org/2016/schedule/events/919.en.html

Snapshot too old

- Чтобы не накапливался большой bloat, нужно обрывать транзакции, которые видят старые данные.
- Параметр old_snapshot_threshold в интервале 1 мин. – 60 дн.
- Обрываем транзакцию только тогда, когда она действительно пытается обратиться к изменившимся данным.
- Осторожно! Значительный overhead на многоядерных системах.
- В будущем, можно применить этот механизм для standby.

Ускорение простых запросов

Оптимизации:

- Push Down Scan key
- Save Expression Initialization for targetlist and qual
- Save Scan slot
- Save Executor State
- Save Expression Context

Это даёт:

- 25% ускорение INSERT
- 20% ускорение SELECT
- 20% ускорение UPDATE

Источник: http://www.pgcon.org/2016/schedule/events/936.en.html

Конференция - 2

Faster PostgreSQL: Improved Writes **Amit Kapila** Track **Performance**	**Audit Logging for PostgreSQL** **David Steele** Track **DBA**	**MLS PostgreSQL** *Implementing Multi-level Security in PostgreSQL with RLS and SELinux* **Joe Conway** Track **Hacking**
WAL Reduction *Lose WAL size, gain performance* **Rahila Syed** Track **Performance**	**Building a BaaS using PostgreSQL** *From schema to REST API instantly* **Lakshmi Narasimhan Parthasarathy** Track **Applications**	**Beyond EXPLAIN: Query Optimization From Theory To Code** **Ryoji Kawamichi** , **Yuto Hayamizu** Track **Hacking**
PostgreSQL's buffer manager – Problems & Improvements **Andres Freund** Track **Hacking**	**The PCI Compliant Database.** *... yours probably is not.* **Christophe Pettus** Track **Applications**	**How I Learned to Stop Worrying and Love Failover** *Automated HA on a container stack* **Josh Berkus** Track **Scaling Out**

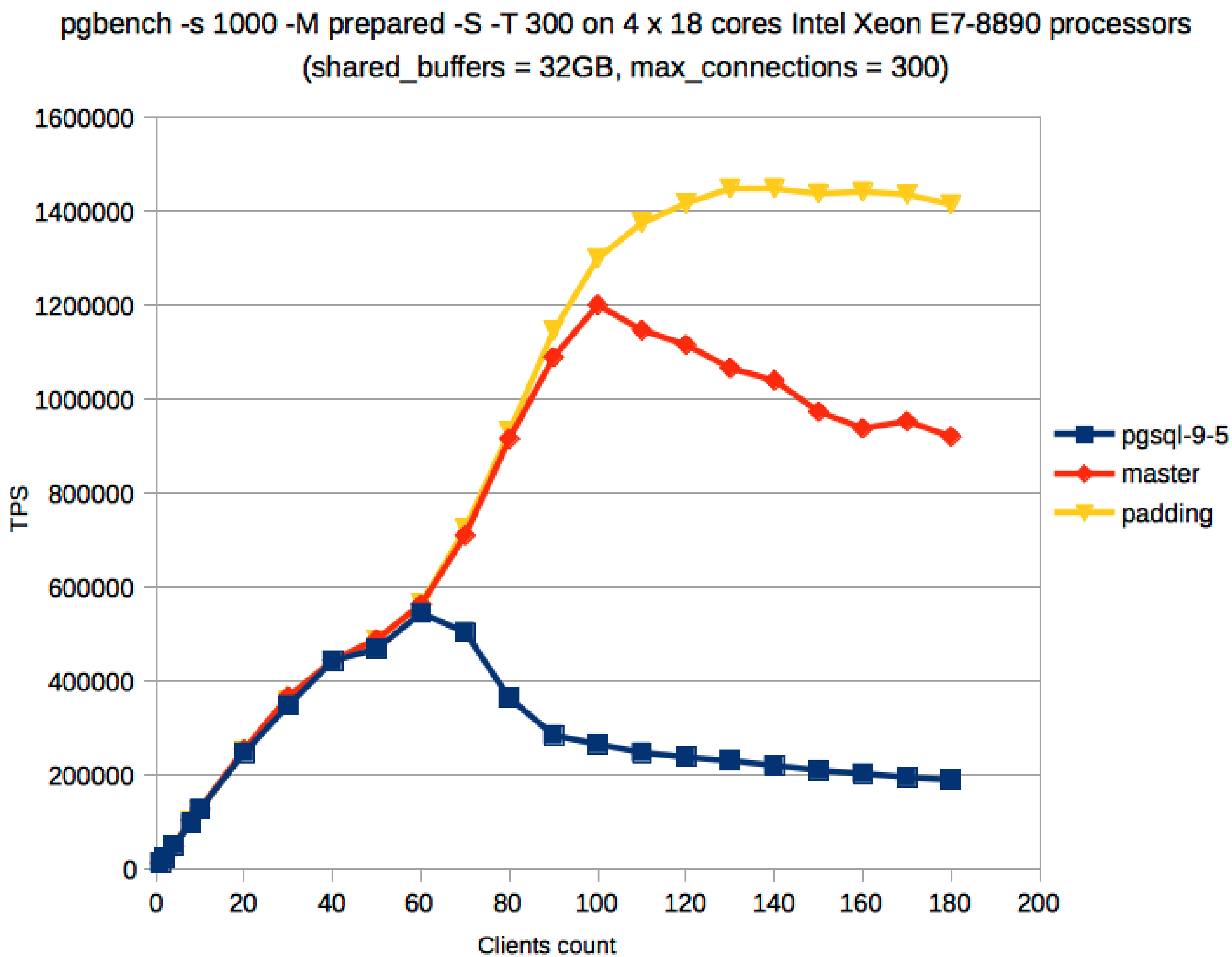
Конференция - 2

Queues in PostgreSQL *Can elephants queue?* **Thomas Munro** Track **Applications**	**Non-volatile Memory Logging** **Takashi Horikawa** Track **Performance**	**PoWA 3: towards a computer assisted DBA** **Ronan Dunklau** Track **DBA**
Overlooked Data Types *A study of useful lesser-known data types.* **Peter van Hardenberg** Track **Applications**	**A Challenge of Huge Billing System Migration** **Tatsuro Yamada** Track **Case Studies**	**Not Easy Being Green** *The Long Tail of Buildfarm Failures* **Noah Misch** Track **Hacking**
Benchmarking Databases *Performance Evaluation* **Jan Wieck** Track **Performance**	**Lessons in Building a Distributed Query Planner** *Why Some SQL Queries are Hard to Scale* **Ozgun Erdogan** Track **Scaling Out**	**Managing OS Provided PostgreSQL Packages** **Keith Fiske** Track **DBA**

Closing sessions
prizes, auctions, fun, games

Dan Langille
Track **Social**

Multicore scalability



«Improved reads and writes»

- Group commit – 30 на большой power-8 машине: 30% ускорение для 64 клиентов и 133% ускорение для 256 клиентов на read-write pgbench тесте.
- Увеличение clog buffers с 32 до 128: 90% ускорение на unlogged таблицах для 256 клиентов на read-write pgbench тесте.
- Параметры checkpointer_flush_after, bgwriter_flush_after и backend_flush_after предотрвращают накопление слишком большой очереди на запись в кэше ОС.

Проблемы buffer manager'а

- Медленная хэш-таблица. Нужно применять lock-free hash-table или prefix tree.

- Алгоритм вытеснения не масштабируется. Нужно менять на другой.

- Checkpoint'ы нужно специально «размазывать», чтобы выдержала дисковая подсპистама.

User Unconference

https://wiki.postgresql.org/wiki/Pgcon2016userunconference

User Unconference

Time	DMS 1160	DMS 1110	DMS 1130
10am	Session Pitches and Scheduling - Coffee and snacks		
10:45AM	PostgreSQL Project Code of Conduct discussion		
12:00 Noon	Lunch!		
1:00PM	Multi-Master Demo	User Complaints	Postgres on FreeBSD
2:15PM	Bug Tracker Discussion	Replaying SQL	PGCAC Annual Meeting PostgreSQL Trademark Q&A
3:30PM	PgInfra & Web Team	Figures & Pics in Docs	Running Postgres in Containers

https://wiki.postgresql.org/wiki/Pgcon2016userunconference

Code of Conduct

PostgreSQL Community Code of Conduct
Draft May 20, 2016 `

Introduction
The PostgreSQL Project prides itself on the quality of our code and our work, and the technical and professional accomplishments of our community. We expect everyone who participates to conduct themselves in a professional manner, acting with common courtesy and in the common interest, with respect for all users and developers.

To that end, we have established this Code of Conduct for community interaction and participation in the project's work and the community at large.

Inclusivity and Appropriate Conduct
The PostgreSQL Project is open to particip
regardless of their level of experience with
development and contributions from all us

We encourage thoughtful, constructive dis
directions for development among our con
related technology, projects and infrastruc

Personal attacks and comments on person

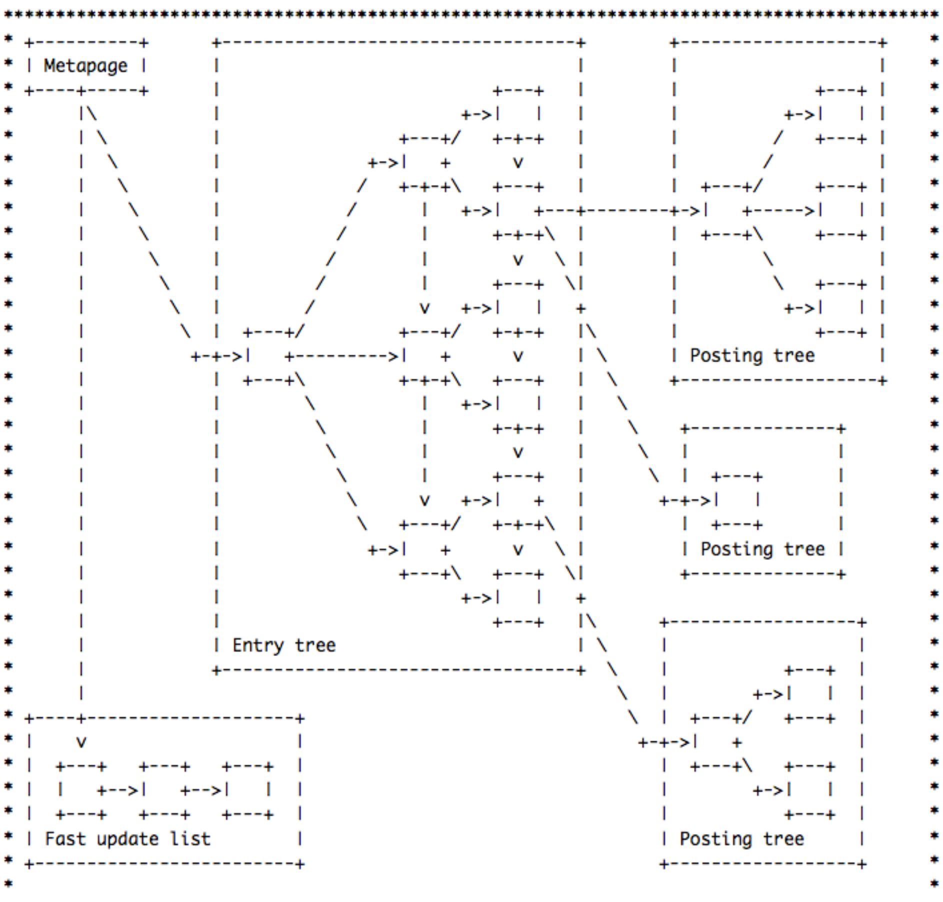
Personal characteristics include, but are no

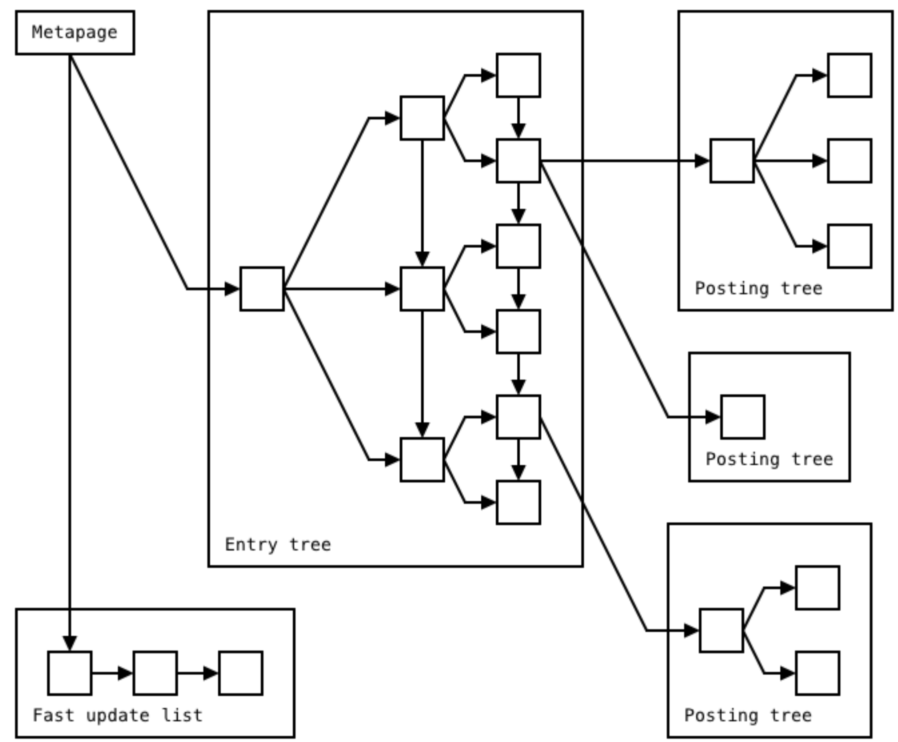
- age
- race
- ethnicity
- color
- gender

https://wiki.postgresql.org/wiki/Coc_qa_pgcon2016

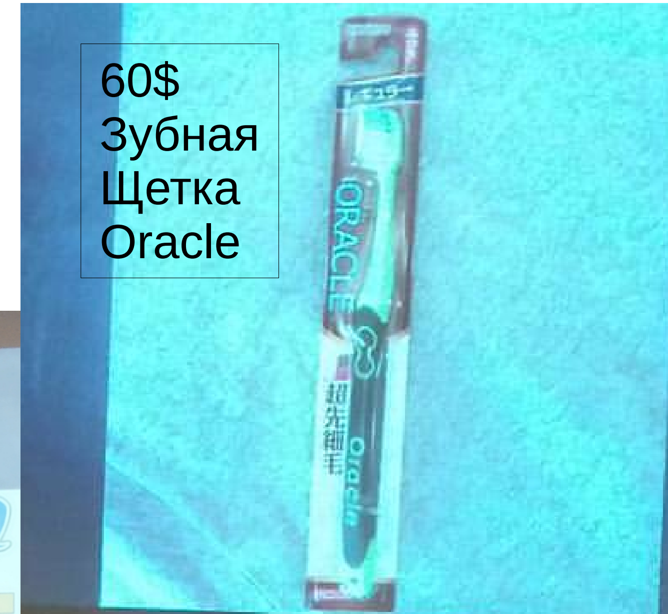
Figures and Pics in Documentation

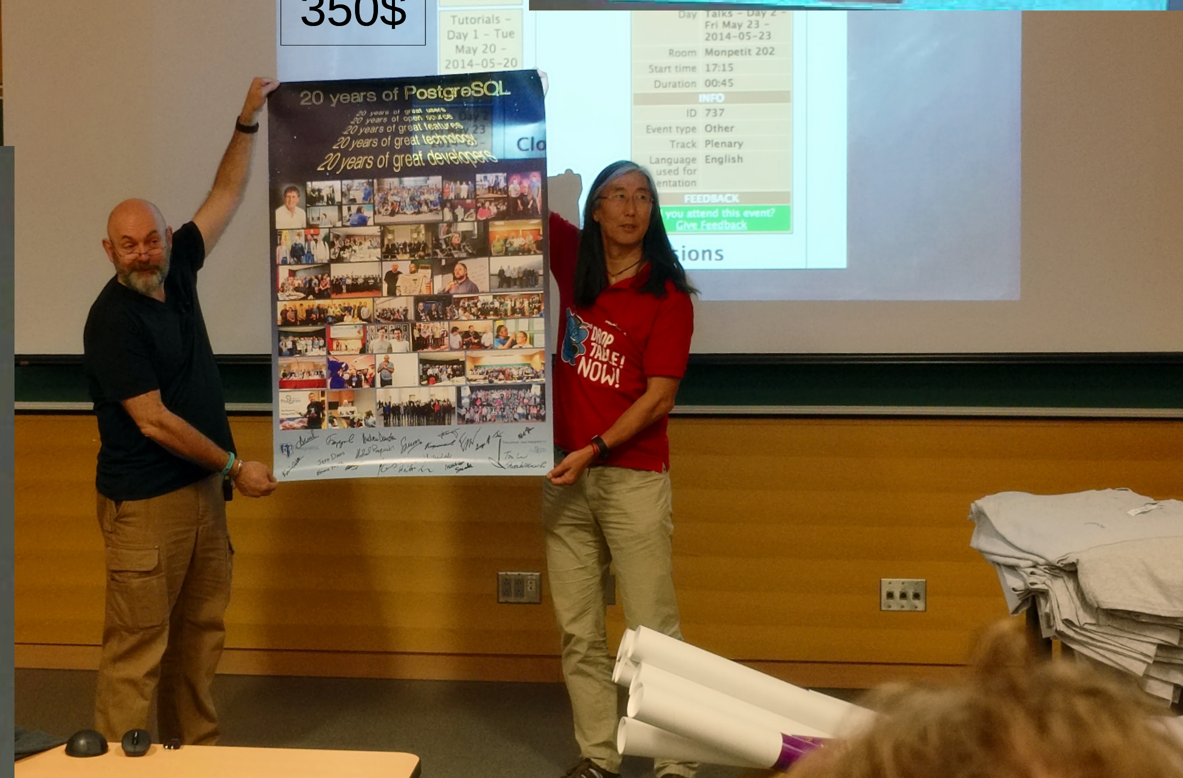
GIN diagram — ASCII

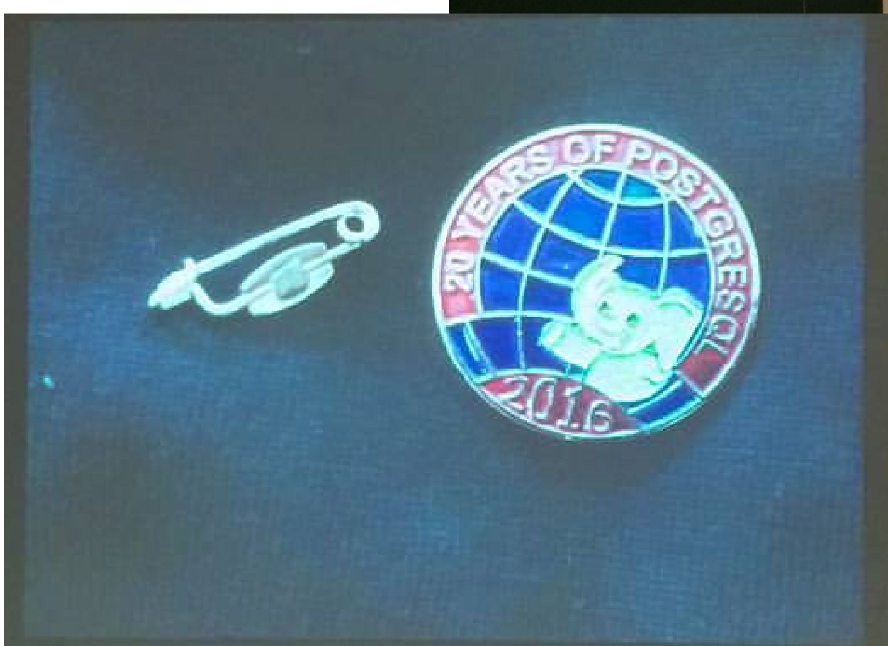




Charity Auction







NEW FEATURES of FTS

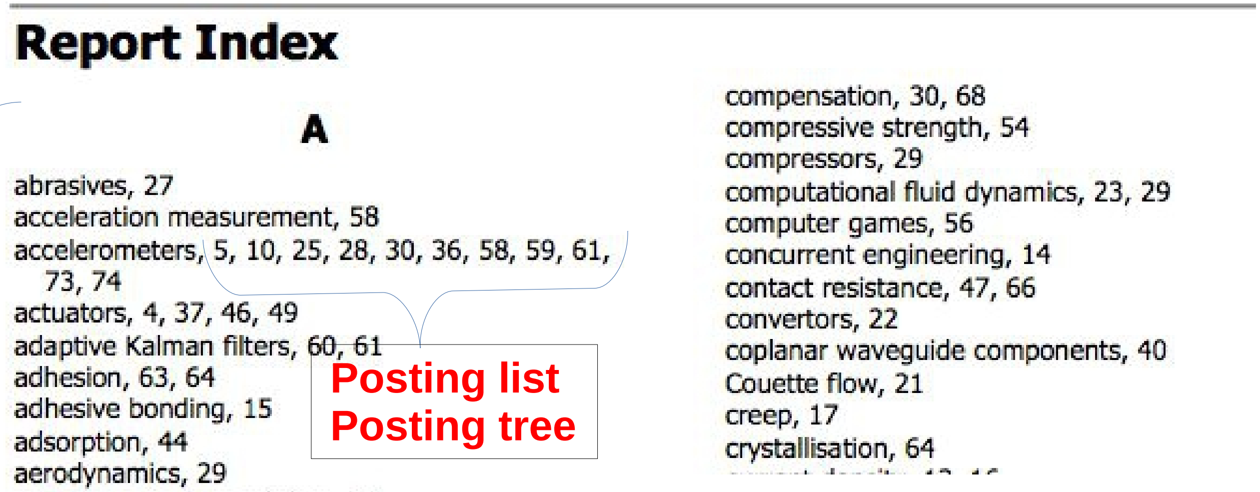
Oleg Bartunov, Teodor Sigaev
Postgres Professional

No positions in index !

Inverted Index in PostgreSQL

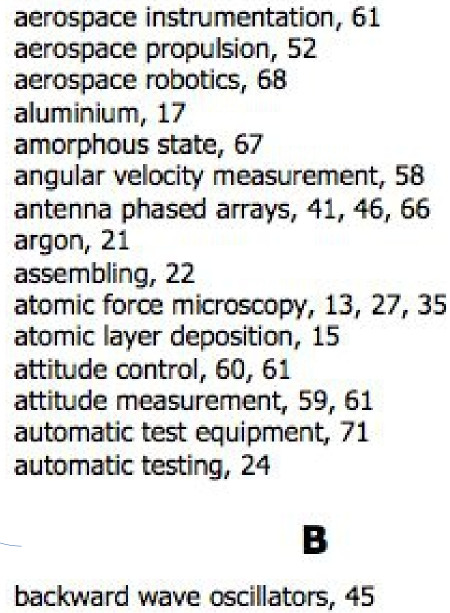
E
N
T
R
Y

T
R
E
E



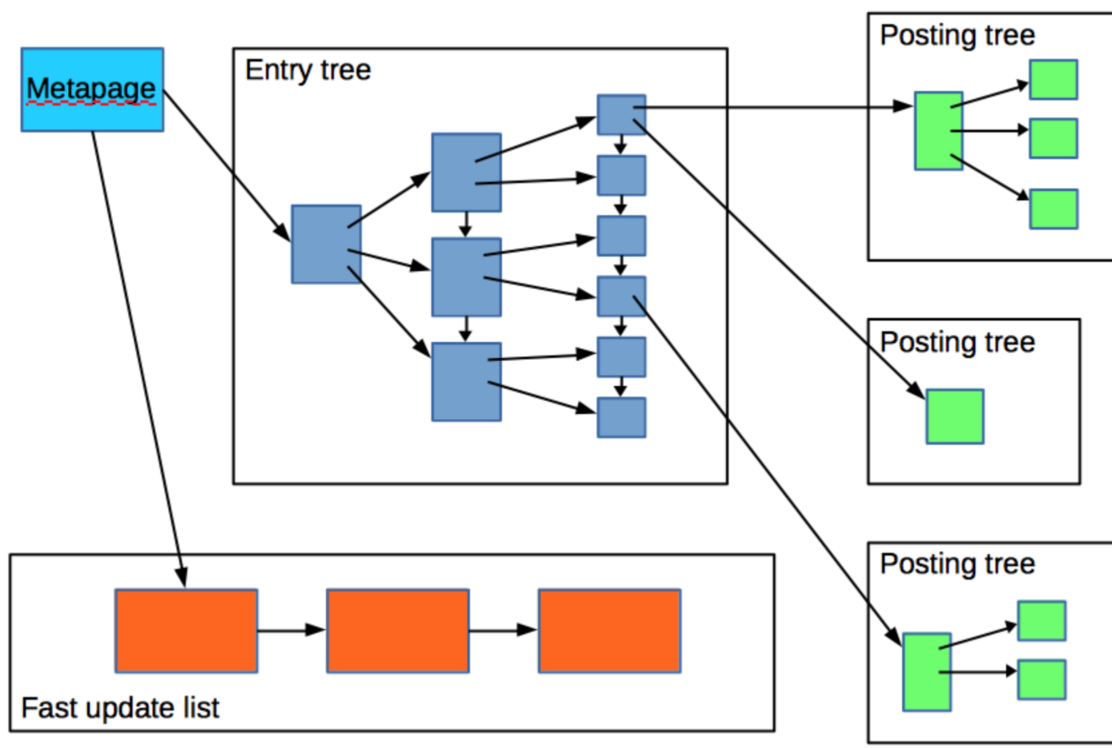
Report Index

A

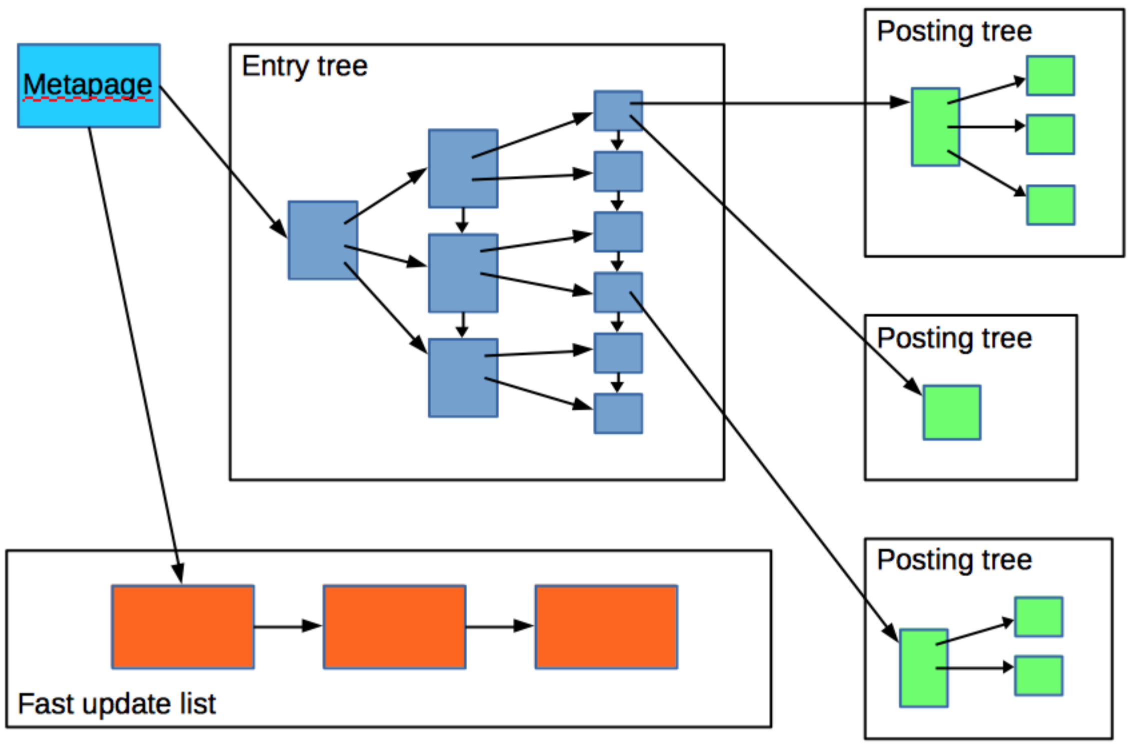


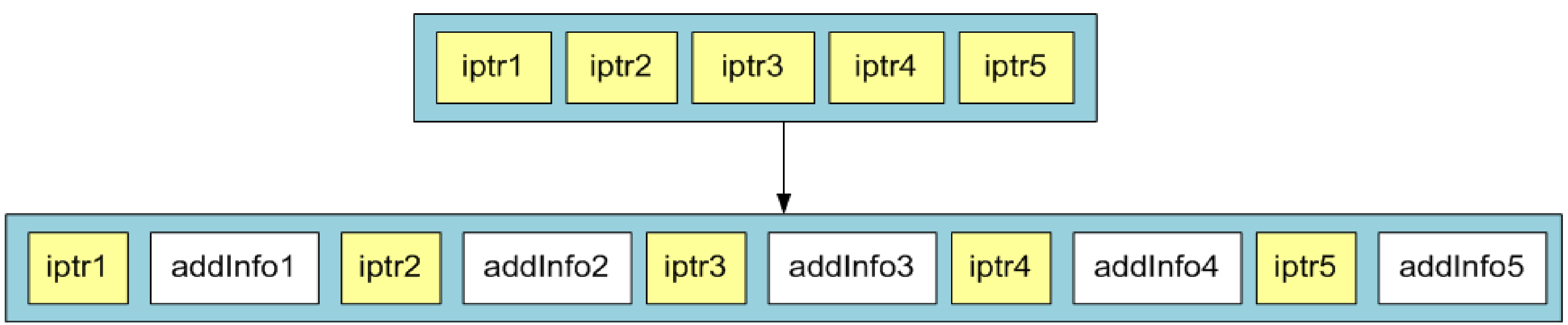
B

Posting list
Posting tree



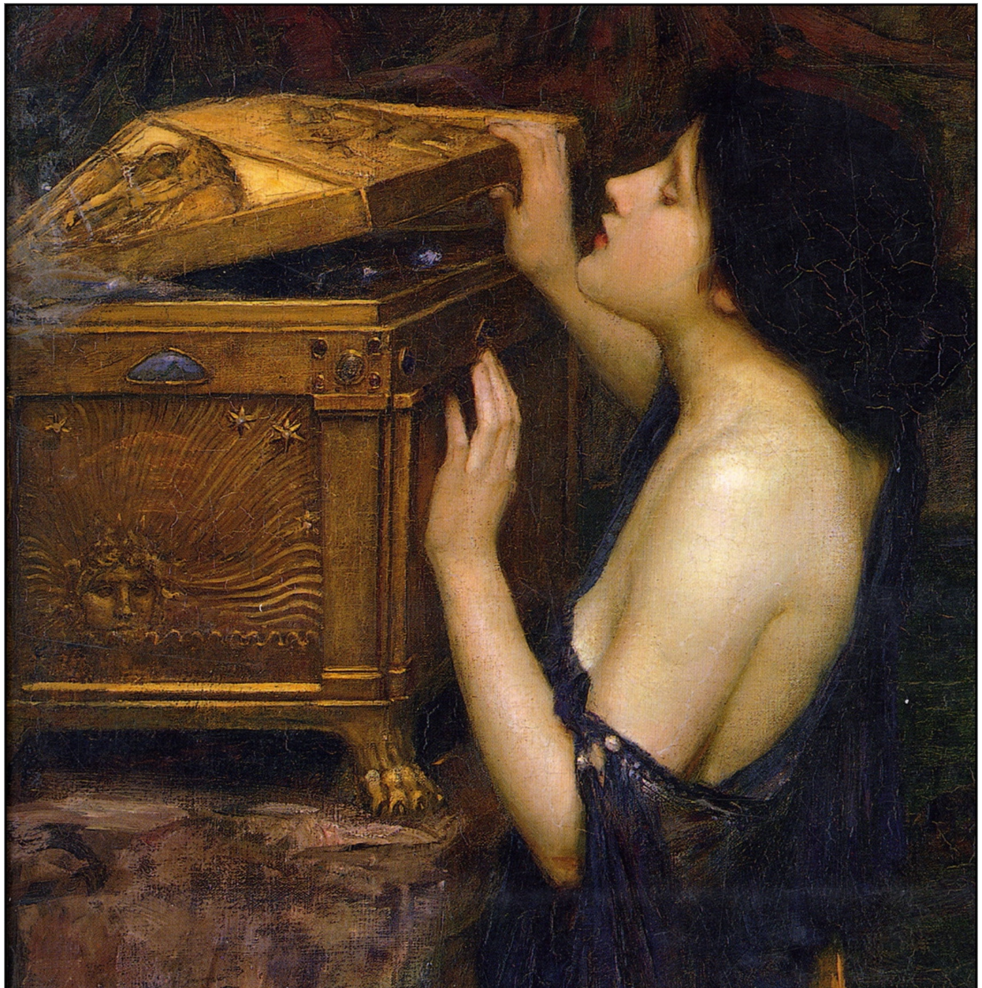
Improving GIN





9.6 opens «Pandora box»

Create access methods as extension ! Let's call it RUM



- Храним в addinfo
  - Координаты слов
    - Ускоряем поиск с ранжирование
    - Ускоряем phrase search
  - Timestamp
    - Ускоряем поиск с упорядочиванием по времени
  - Ветки query tree
    - Ускоряем "обратный" поиск Full Query Search

http://www.sai.msu.su/~megera/postgres/talks/pgcon-2016-fts.pdf

Postgres Professional на PGCon


Thanks !