Future Postgres Challenges

BRUCE MOMJIAN

EDB

This presentation explores possible challenges to Postgres's success in the coming years.

Last updated: November 2021

Outline

1. Current status
2. Project challenges
3. Competitive challenges
4. Technical challenges

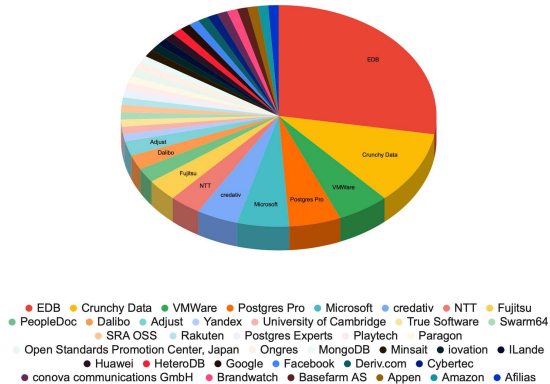*https://www.flickr.com/photos/snikologiannis/*

Consistent Development

- 35+ years of development
- 25+ years of annual major releases
- ~180 features per major release
- Quarterly minor releases

Healthy Community Structure

- BSD license guarantees software will be available forever, including for proprietary use.
- Development and leadership is diversified geographically, culturally, and is multi-company.
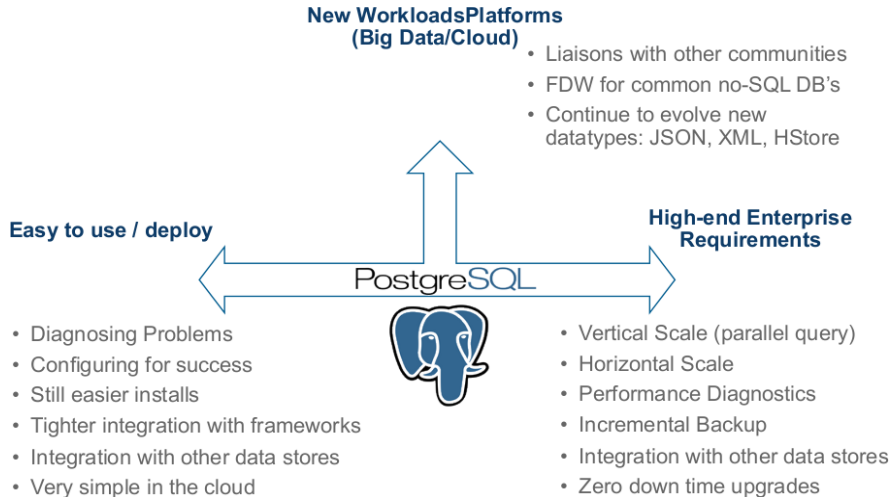
Strong Diversified Assistance

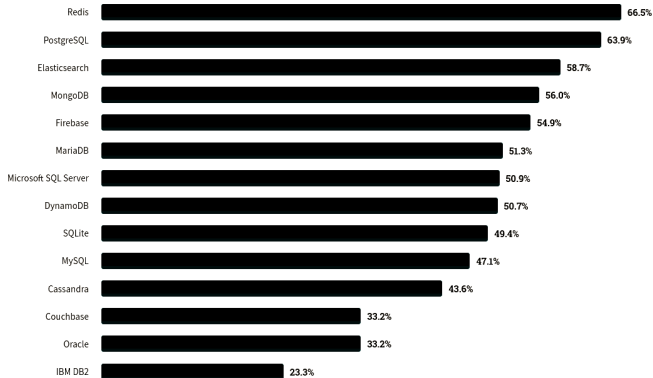Postgres contributors (PostgreSQL.org + PG 13 Release Notes) w.o. personal or freelancers

● EDB ● Crunchy Data ● VMWare ● Postgres Pro ● Microsoft ● credativ ● NTT ● Fujitsu
● PeopleDoc ● Dalibo ● Adjust ● Yandex ● University of Cambridge ● True Software ● Swarm64
● SRA OSS ● Rakuten ● Postgres Experts ● Playtech ● Paragon
● Open Standards Promotion Center, Japan ● Ongres ● MongoDB ● Minsait ● iovation ● ILande
● Huawei ● HeteroDB ● Google ● Facebook ● Deriv.com ● Cybertec
● conova communications GmbH ● Brandwatch ● Basefarm AS ● Appen ● Amazon ● Afilias

Marc Linster, EDB

Innovative Features

**New WorkloadsPlatforms
(Big Data/Cloud)**

- Liaisons with other communities
- FDW for common no-SQL DB's
- Continue to evolve new
  datatypes: JSON, XML, HStore

Easy to use / deploy

**High-end Enterprise
Requirements**

PostgreSQL

- Diagnosing Problems
- Configuring for success
- Still easier installs
- Tighter integration with frameworks
- Integration with other data stores
- Very simple in the cloud

- Vertical Scale (parallel query)
- Horizontal Scale
- Performance Diagnostics
- Incremental Backup
- Integration with other data stores
- Zero down time upgrades

Keith Alsheimer, EDB, 2013

Most Loved Relational Database in 2020

Loved Dreaded Wanted

% of developers who are developing with the language or technology and have
expressed interest in continuing to develop with it

Database	Percentage
Redis	66.5%
PostgreSQL	63.9%
Elasticsearch	58.7%
MongoDB	56.0%
Firebase	54.9%
MariaDB	51.3%
Microsoft SQL Server	50.9%
DynamoDB	50.7%
SQLite	49.4%
MySQL	47.1%
Cassandra	43.6%
Couchbase	33.2%
Oracle	33.2%
IBM DB2	23.3%

https://insights.stackoverflow.com/survey/2020#technology-most-loved-dreaded-and-wanted-databases-loved4

https://www.flickr.com/photos/croydonclicker/

Leadership Disruption

- Gimp was abandoned by its lead developers, later resurrected
- Red Hat took over CentOS, changed stability

Poor Reputation

- Security flaws
- Buggy releases
- Instability
- Poor performance
- Data corruption

Patent Attacks

- Developer with patents, Rambus
- Competitor with patents, Microsoft
- Patent trolls, Rothschild Patent Imaging LLC
- Open Invention Network

Identity Challenges

- Domain name
- Website
- Trademark

Cloud Vendor Starvation

- Cloud vendors use open source as upsell
- Already have infrastructure-as-a-service relationship with customers
- Company-controlled open source already impacted, changed licenses
- Red Hat challenged by cloud vendors

https://momjian.us/main/blogs/pgblog/2020.html#September_25_2020

https://momjian.us/main/blogs/pgblog/2020.html#September_28_2020

Other Solutions

- Relational (e.g., MySQL, MariaDB)
- Embedded (SQLite)
- Document (MongoDB)
- Columnar (Cassandra)
- Data warehouse (Hadoop)
- Full text search (Elasticsearch)
- Time series (InfluxDB)

The Rise of Forks

- Forks of Postgres go back to the early 1990's with Illustra
- Popular fork goals
  - cloud customization
  - horizontal scaling
  - data warehouse
- BSD split into FreeBSD, NetBSD, and OpenBSD
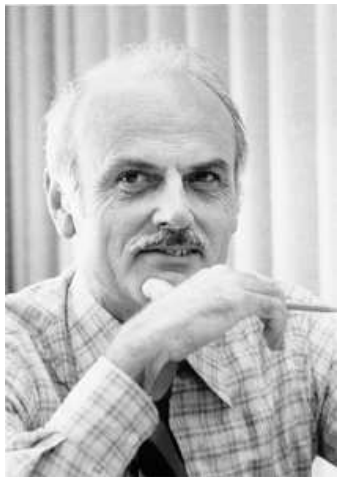- Egcs forked gcc, became popular, later became the new gcc
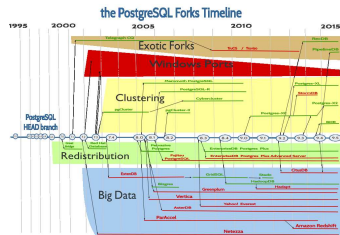
Decline of Relational

- Relational storage was proposed by E. F. Codd in 1970
- 50+ years still in use
- Very flexible
- Resisted challenges
  - XML databases
  - Object databases
  - NOSQL

*https://www.flickr.com/photos/afc16/*

Write Amplification

- Non-HOT updates can cause massive index updates
- Dead and old row version cleanup can become expensive for certain workloads
- Writes cause full page image and hint WAL writes
- Freezing of old transaction ids
- Incremental improvements
  - Are radical improvements needed?

Cluster File Encryption, TDE

- Newer versions of the PCI DSS specification make storage-only encryption less acceptable
- This is a check-box requirement for many financial deployments
- Development is in progress

Horizontal Scaling

- Allows data storage larger than possible on a single server
- Allows write scaling
- Enables large CPU and memory scaling
- Development is in progress
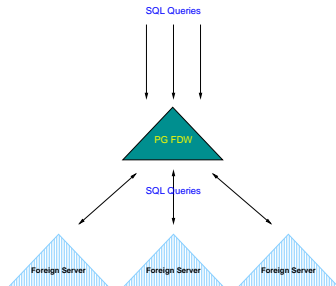
Difficulty replacing obsolete or abandoned:

- Programming languages
- Support libraries
- Testing frameworks

Drastic Technology Changes

- New language, architecture, or storage that are difficult for Postgres to adopt
- Technology changes have happened before
  - SSDs, added random_page_cost to tablespaces
  - virtual machines, containers, cloud

Conclusion

https://momjian.us/presentations