Future Postgres Challenges

BRUCE MOMJIAN

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

https://momjian.us/presentations

Creative Commons Attribution License

Last updated: March 2022
Outline

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

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

Marc Linster, EDB
Innovative Features

**New Workloads/Platforms**
(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**

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

**High-end Enterprise Requirements**

- 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

2. Project Challenges

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
- Unified Patents
Identity Challenges

- Domain name
- Website
- Trademark
Cloud Vendors 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


3. Competitive Challenges

https://www.flickr.com/photos/oui-ennui/
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

4. Technical Challenges

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

https://wiki.postgresql.org/wiki/Transparent_Data_Encryption
- Allows data storage larger than possible on a single server
- Allows write scaling
- Enables large CPU and memory scaling
- Development is in progress
Obsolete Toolchain

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