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: April 2023
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

Code contributors to Postgres by company (PostgreSQL.org + PG 15 Release Notes) without personal or freelancers

Sarah Conway, 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

PostgreSQL

- 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

- Redis: 66.6%
- 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: 32.2%
- IBM DB2: 23.3%

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
- Good news
  - Open Invention Network
  - Unified Patents
  - Project Jengo at Cloudflare
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


3. Competitive Challenges

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

Other Solutions

- Relational (e.g., MariaDB)
- Embedded (SQLite)
- Document (MongoDB)
- Columnar (ClickHouse)
- 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

https://raw.githubusercontent.com/daamien/artwork/master/inkscape/PostgreSQL_timeline/timeline_postgresql.png
https://wiki.postgresql.org/wiki/PostgreSQL_derived_databases
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
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

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
Conclusion

https://momjian.us/presentations

https://www.flickr.com/photos/91451979@N00/