



Günümüzde Postgres

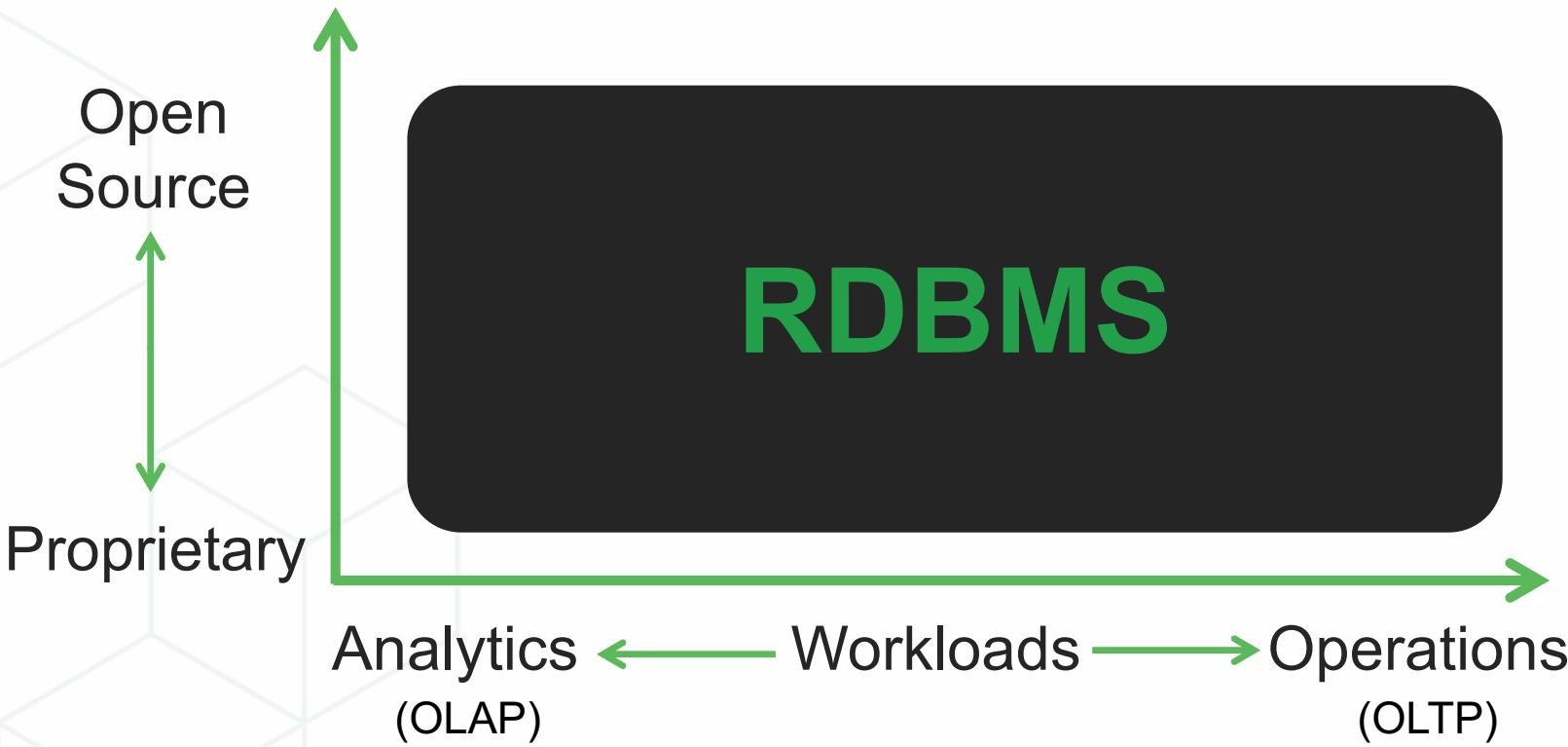
Modern, Ölçeklenebilir Uygulamalar

Utku Azman | Citus Data | PGDay İstanbul

| @citusdata | citusdata.com

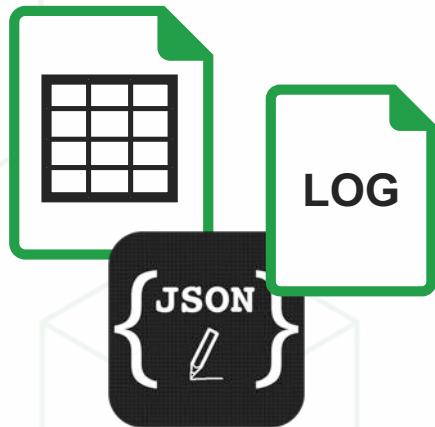


Veritabanlarının basit olduğu zamanlar (2000'lere kadar)

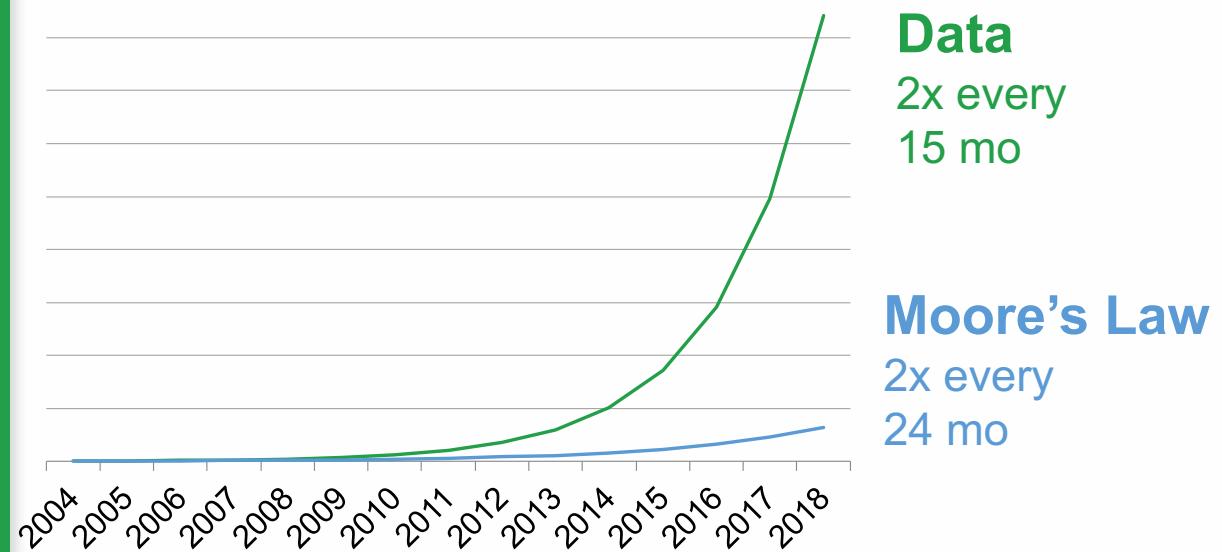


Bugün yapısal veritabanlarını zorlayan iki gelişme...

1 Data with less structure



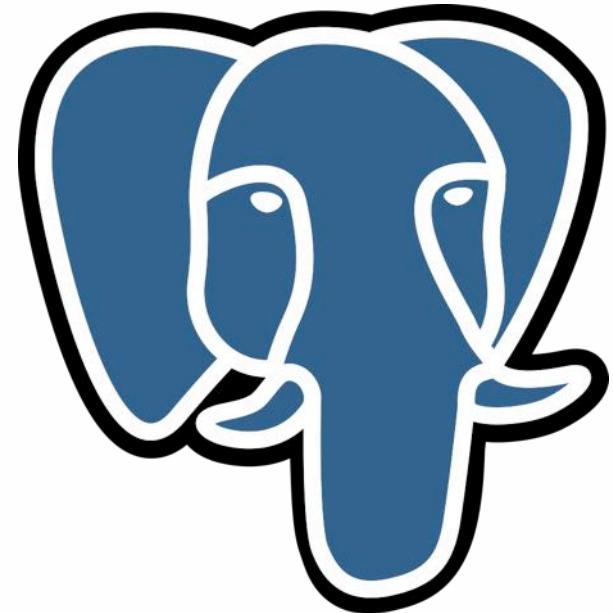
2 Data Growth >> Silicon Growth...



Postgres'in geçmişinden... Bu zorluklar yeni değil

1 SQL or not? (1995)

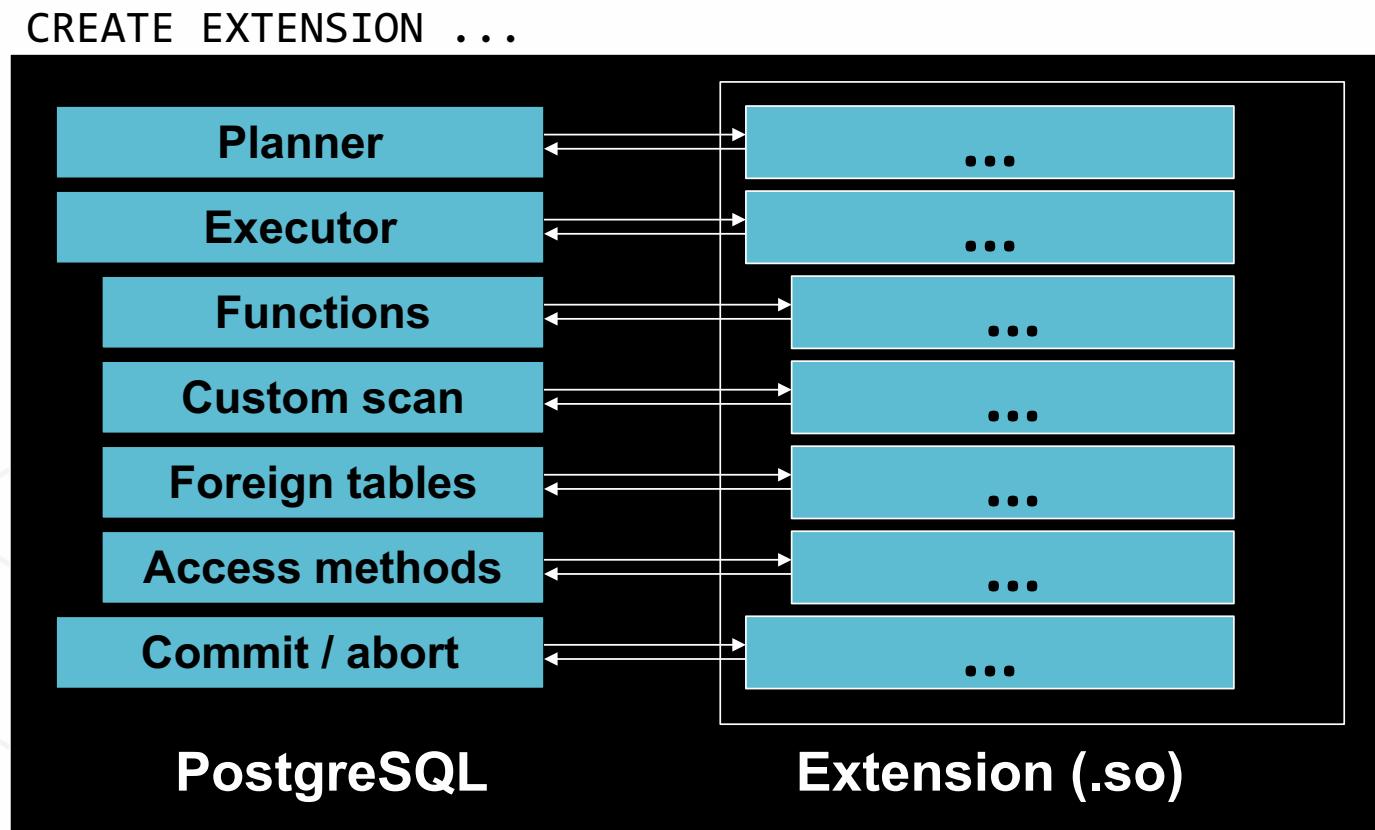
- Post-Ingres
- Started life as object store
- Added SQL API in 1995



2 Scaling out to handle data growth (2005)

- For analytics only: MPPs
- So many forks! AsterData, Netezza, ParAccel (Redshift), Greenplum

Bugün ise Postgres bu zorluklara karşı çok daha güçlü - SQL Extension APIs (2011)





RDBMS ile günümüz ihtiyaçlarına cevap verme

- 1 Yapısal ve yapısal olmayan veri
 - 2 Ölçeklenme—işleme & performans
- 



Dosya sisteminden başlama

- (+) Any data, any structure
- (+) 'Infinitely' scalable storage
- (+) Write fast

- (-) Pay cost at query time
- (-) Batch vs. real-time
- (-) Indexes (Append only FS)



Tek bir erişim şeklini destekleme

Semi-structured (JSON)

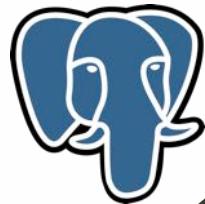
(+) Simple: Put & Get

(+) Scalable

(-) No expressiveness for analytics

(-) No JOINS, data duplication

(-) Enforce structure at app layer



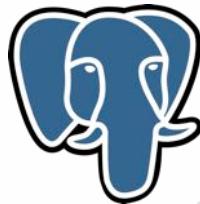
1

Yapısal ve Yapısal olmayan veri

Veritabanını JSON data tipiyle genişletme

Table "public.events"

Column	Type	Sample Data
user_id	<i>bigint</i>	09288
created_at	<i>timestamp</i>	2018-03-08 00:57:12.6936+00
payload	jsonb	{ "person":{ "name":"Emma", "emails": {"west":"emma@sf.com", "east":"emma@nyc.com"}, "address":{ "street_number":12, "street_name":"15th Street", "city":"San Francisco", "state":"CA", "zip":94107 }, "phone_numbers":["6501234567","6500123456"] }, "country":"USA", "event_type":"push_notification", "device_type":"mobile" }



2

Ölçeklenme – İşleme ve performans

Index ve diğer temel performans araçları

B-tree indexes

GIN & GiST indexes

Secondary indexes

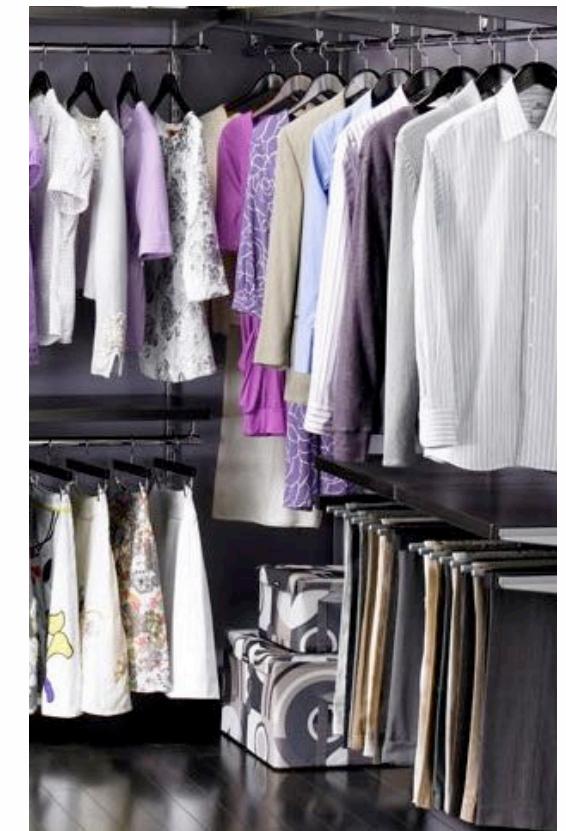
Full text search

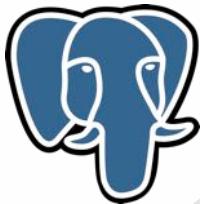
Index-only scans

Fitting indexes into memory

+

Not to forget: Parallel queries, MVCC, and many more.

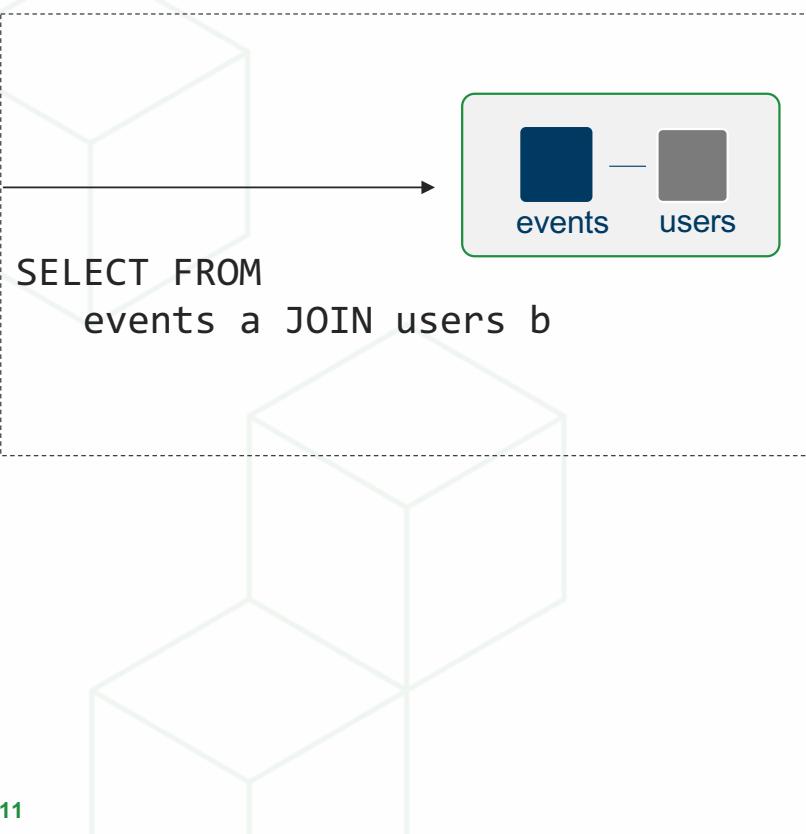




2

Ölçeklenme – İşleme ve performans

İşlem (ve join)leri birden fazla PostgreSQL instances'a gönderme



SELECT FROM (a JOIN b)

SELECT FROM (a JOIN b)

SELECT FROM (a JOIN b)

SELECT FROM (a JOIN b)

Data Node 1

Events_101 [blue square] —> Users_101
Events_103 [blue square] —> Users_103
...
[green square]

Data Node 2

Events_102 [blue square] —> Users_102
Events_104 [blue square] —> Users_104
...
[green square]

Data Node N

[blue square] —> [green square]
[blue square] —> [green square]
[blue square] —> [green square]
[green square]

Postgres'e yatay ölçeklenme için eklenti: Citus



This repository

Search

Pull requests

Issues

Marketplace

Explore



citusdata / citus

Watch

133

Unstar

2,541

Fork

175

Code

Issues 483

Pull requests 16

Projects 0

Wiki

Insights

Settings

Scalable PostgreSQL for multi-tenant and real-time workloads <https://www.citusdata.com>

Edit

database

citus

multi-tenant

postgresql

scale

sharding

sql

distributed-database

postgres

Manage topics

1,734 commits

88 branches

36 releases

29 contributors

AGPL-3.0

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download



onderkalaci Merge pull request #2031 from citusdata/fix_immediate_shut_down_issue ...

Latest commit e7b28dd 3 days ago



config

Add citus_version(), analogous to PG's version()

5 months ago



src

Handle failures during I/O

3 days ago



.codecov.yml

Remove obsolete lines

5 months ago



.editorconfig

Set tab size for GitHub display

a year ago



.gitattributes

Add ruleutils file for PostgreSQL 11

5 months ago

PostgreSQL: Dinamik ve global ekosistem

Sample PostgreSQL Extensions

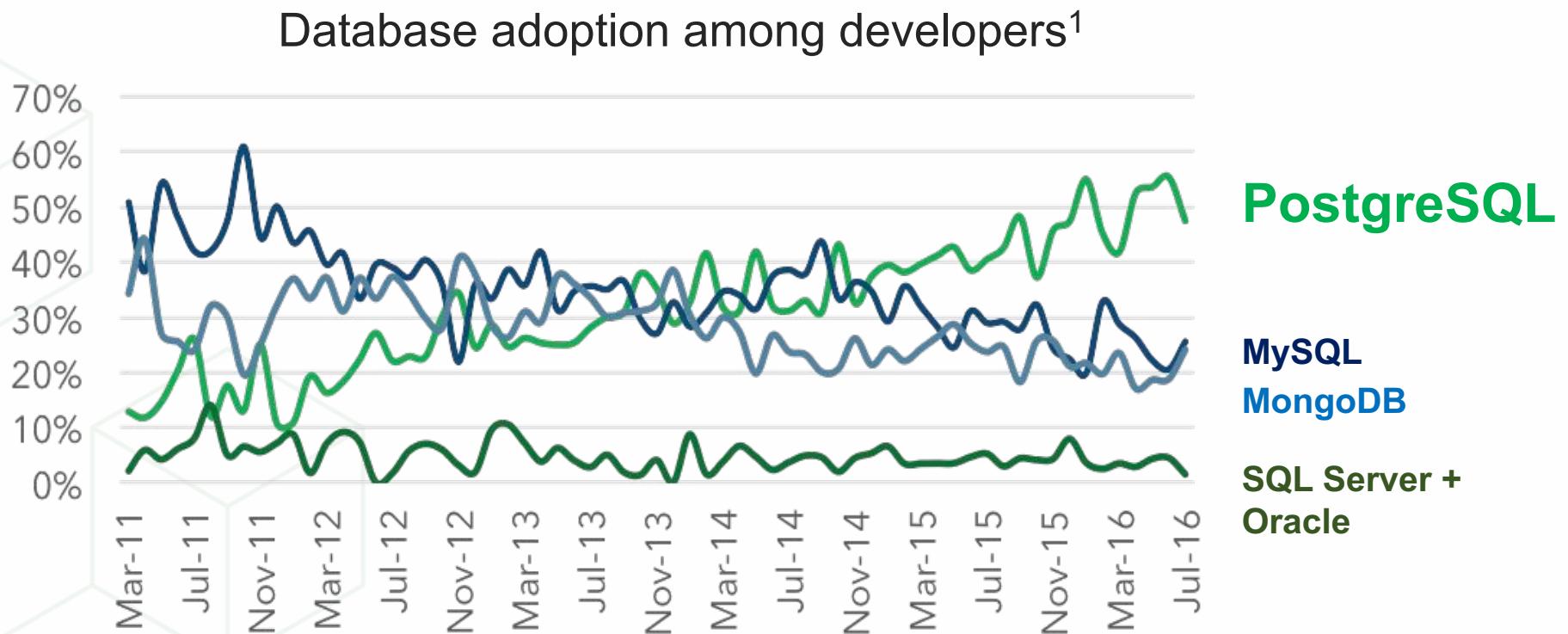
- citus
- pgcrypto
- pg_cron
- pg_partman
- postgresql-HLL
- cstore_fdw
- unaccent
- cube
- jdbc_fdw
- pg_trgm
- PostGIS
- ...

- pg_buffercache
- pg_prewarm
- btree_gin
- btree_gist
- postgis_topology
- pg_stat_statements
- postgresql-unit
- plpgsql
- plv8
- pg_telemetry
- foreign data wrappers
- ...

Integrations



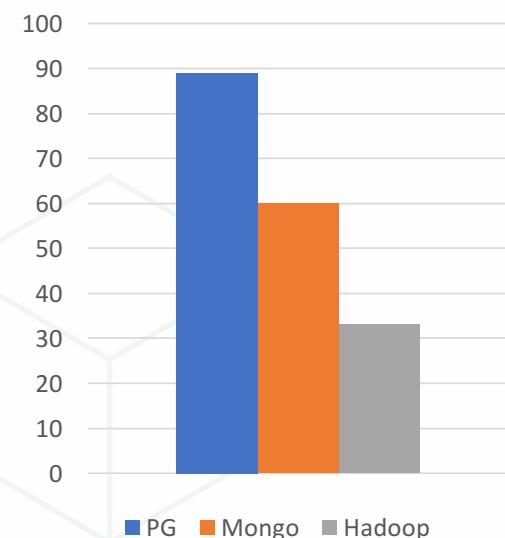
PostgreSQL'in artan momentum



Source: % database job postings that mention each specific technology, across 20K+ job posts on Hacker News, <https://news.ycombinator.com>

...ve büyüyen kullanıcı kitlesi

PostgreSQL popularity =
Hadoop + Mongo combined

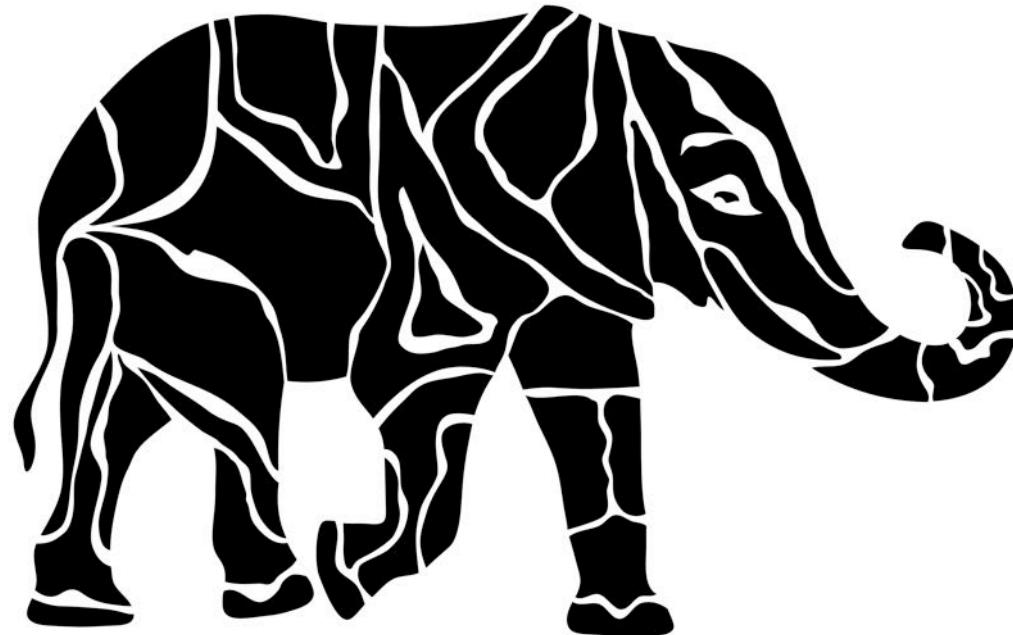


Source: Google Trends for the past 2 years

Winning Startups &
Enterprises

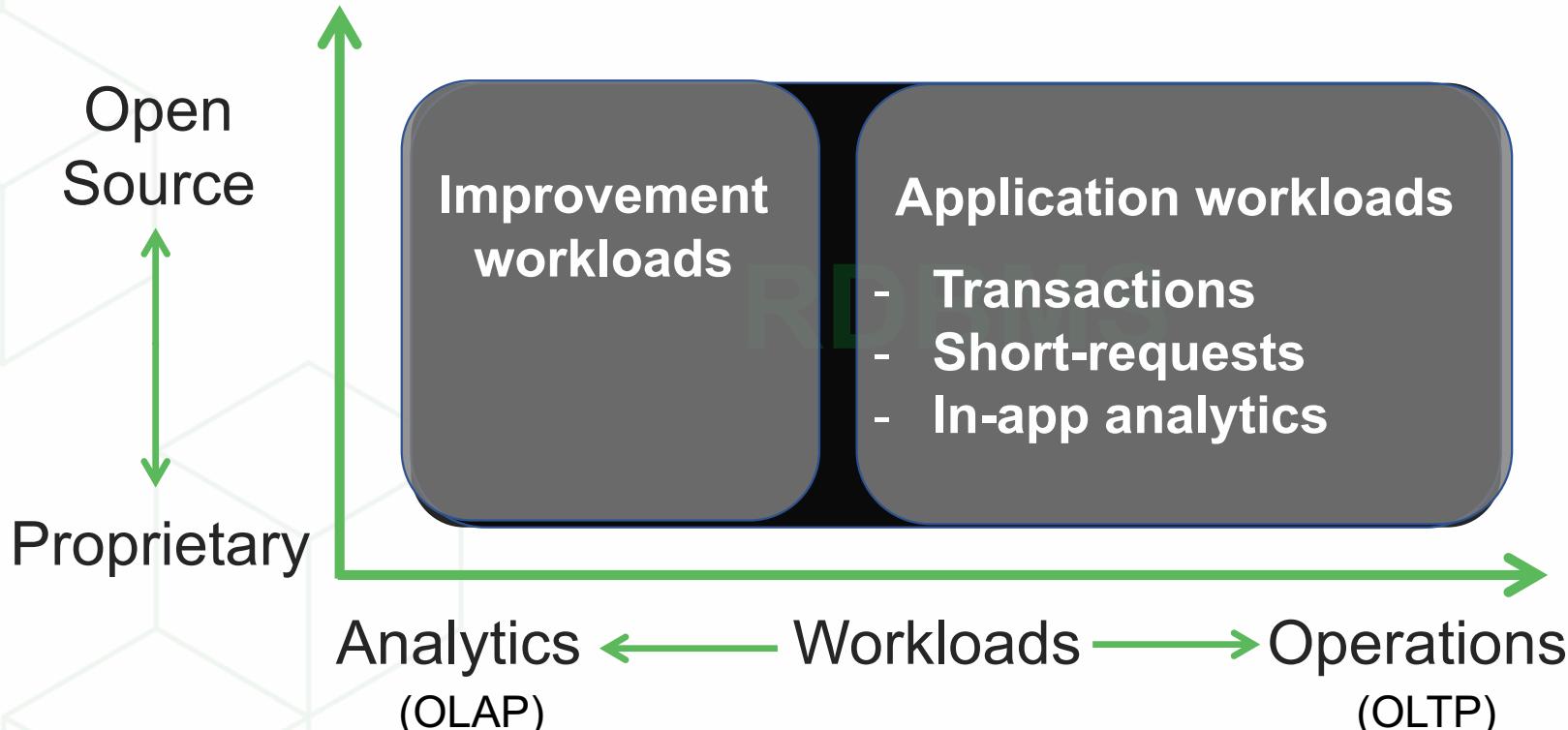


Postgres'in uzun soluklu yürüyüşü

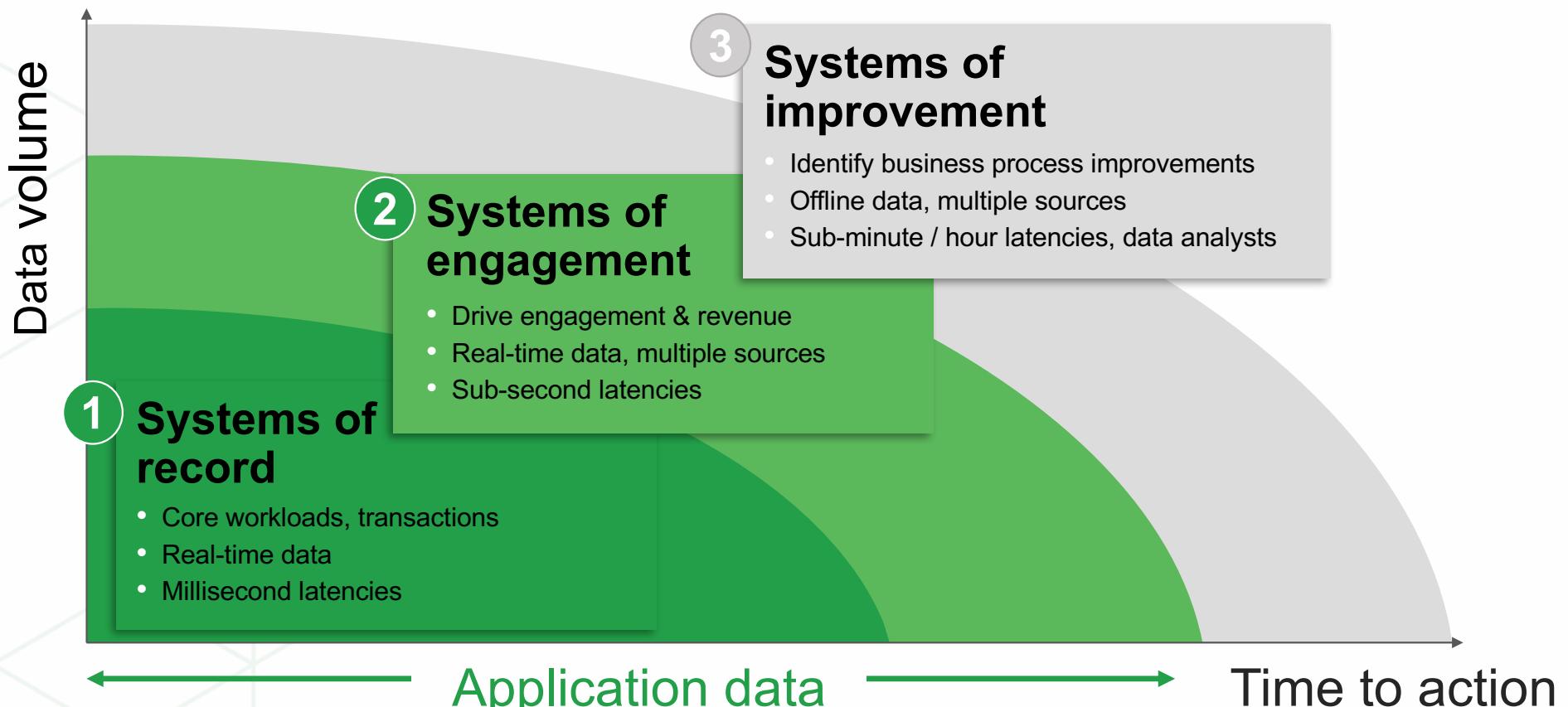


Günümüz uygulamalarında Postgres stack'in neresinde?

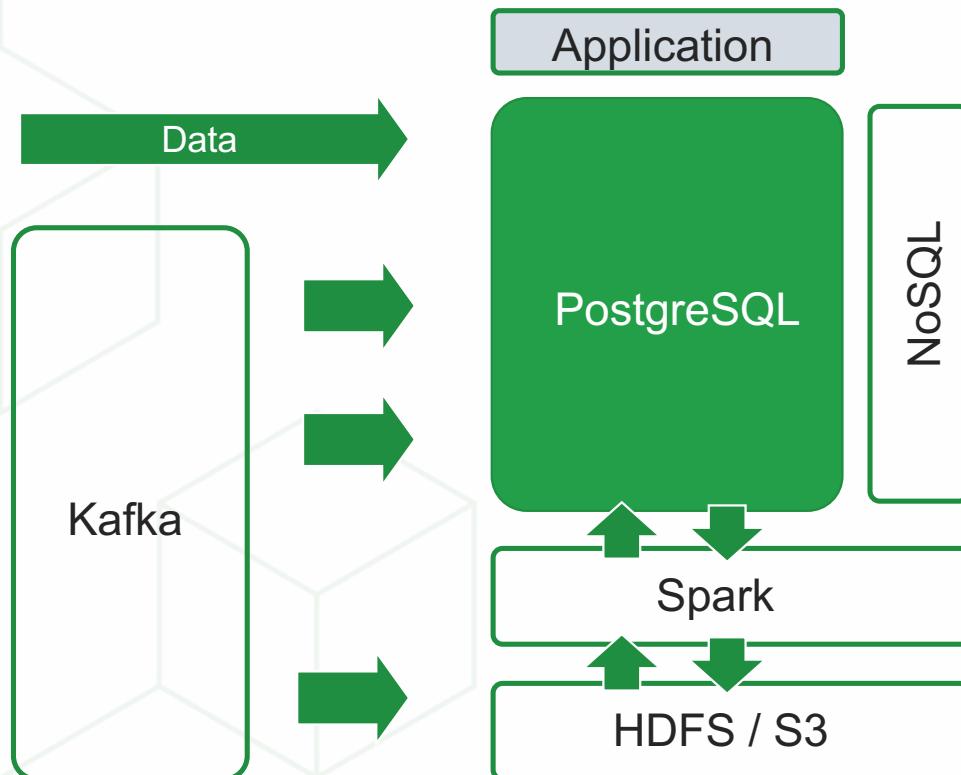
Modern iş yükleri de evrim geçiriyor...



Modern veritabanlarının desteklediği 3 tip uygulama



Altyapınızda PostgreSQL'in yeri



- Standalone database
 - Storage
 - Compute
- Persistence layer for Kafka
- Persistence layer for Spark
- Adjacent to NoSQL

Note: Standard PostgreSQL connectors for all tools (e.g. ODBC / JDBC, PostgreSQL language bindings) available for integrations.

Özetle: PostgreSQL





Rasty Turek @synopsi

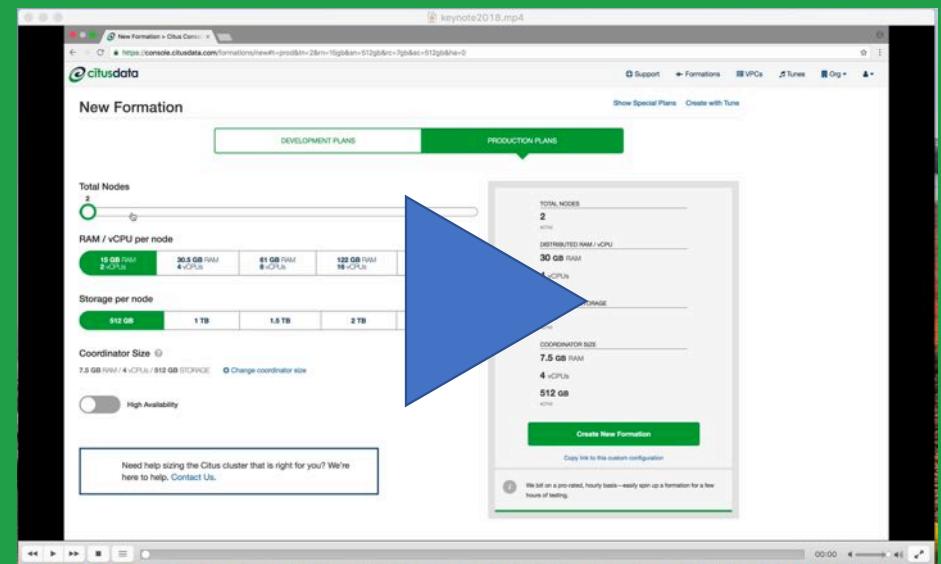
Dec 21, 2016

Today one of our PostgreSQL tables surpassed half of trillion rows [over 54TB of data].

And still most of our queries run under 600ms



Ölçeklenme – Video demo





Teşekkürler



@citusdata

github.com/citusdata/citus

citusdata.com/jobs