



DataStax Hands-On Workshop

Aaron Regis & Richard Henderson

12th Jan 2018

Agenda

| | |
|---|----------------------------------|
| 1 | Quick intro to DataStax and Team |
| 2 | Key Facts DataStax Enterprise |
| 3 | Hands-On Agenda today |

DataStax

DataStax

400+
Employees

Founded in April 2010

Santa Clara • San Francisco • Austin •
London • Paris • Berlin • Tokyo • Sydney

\$190M
Funding



(Series E – Sept. 2014)

500+
Customers



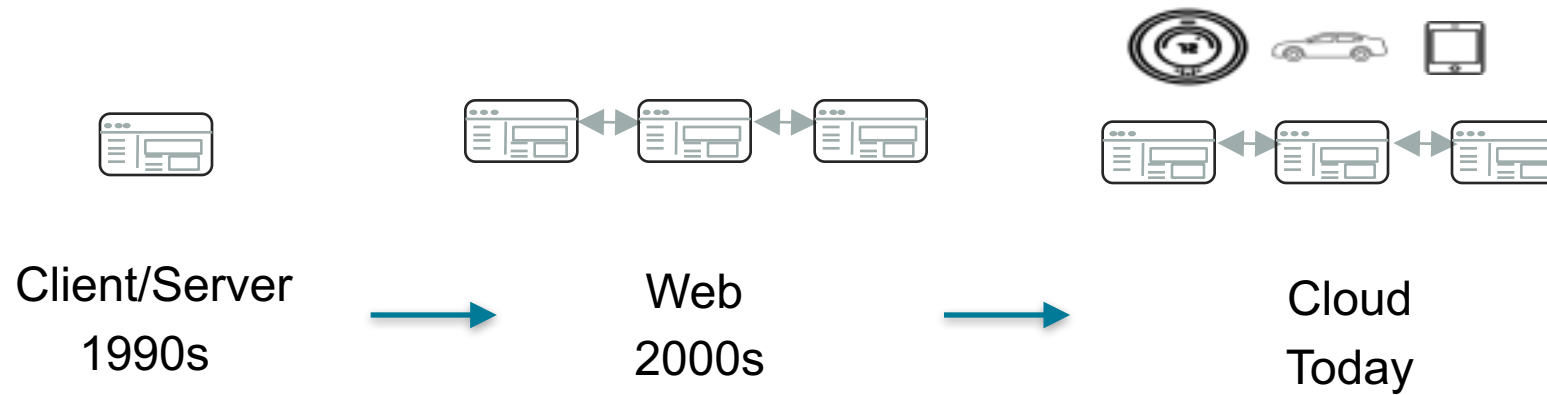
Forbes

2016 World's Best
100 Cloud Companies

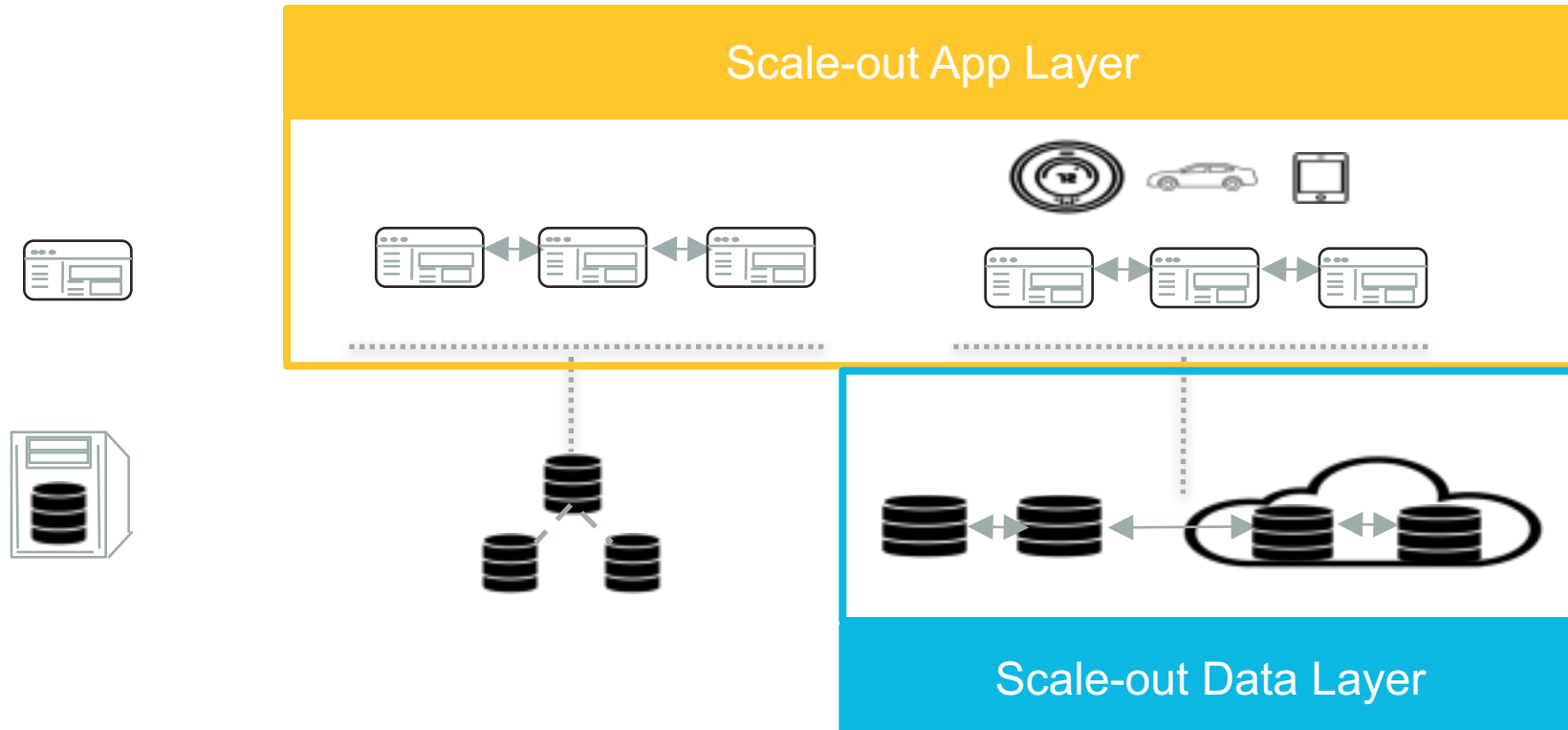
Gartner

Ranked #1 in multiple operational
database categories

Apps have changed



And what powers them



Cloud Application Characteristics



Contextual



Always-On



Real-Time



Distributed



Scalable

Powering Cloud Applications

Always-on

- Designed to handle any failure, no matter how catastrophic.

Effortless scale

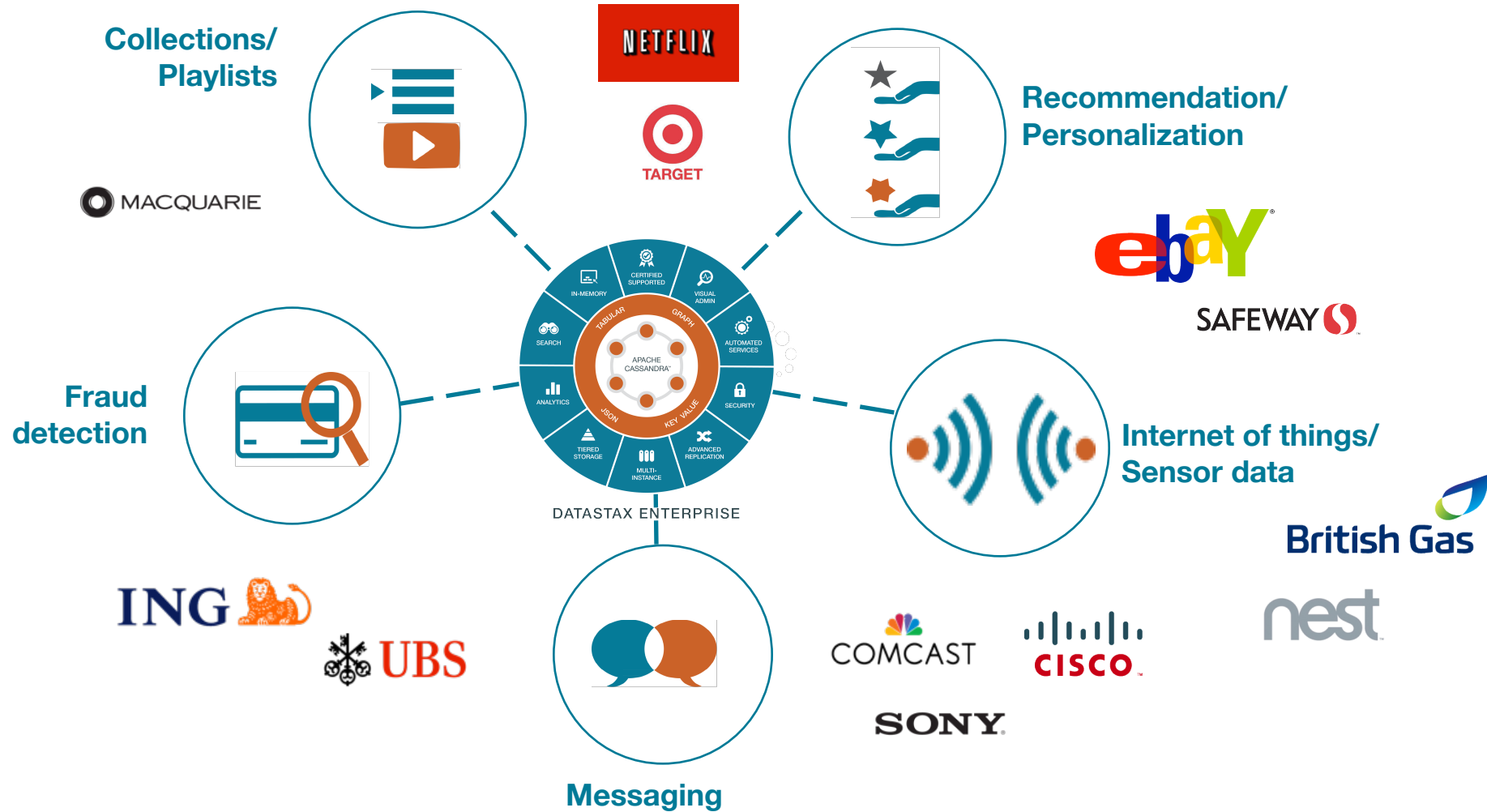
- Take advantage of every opportunity.
- Focus on what matters most to you.

Instant insight

- Built into your application to create actionable, modern experiences.



Functional use cases



DataStax Enterprise

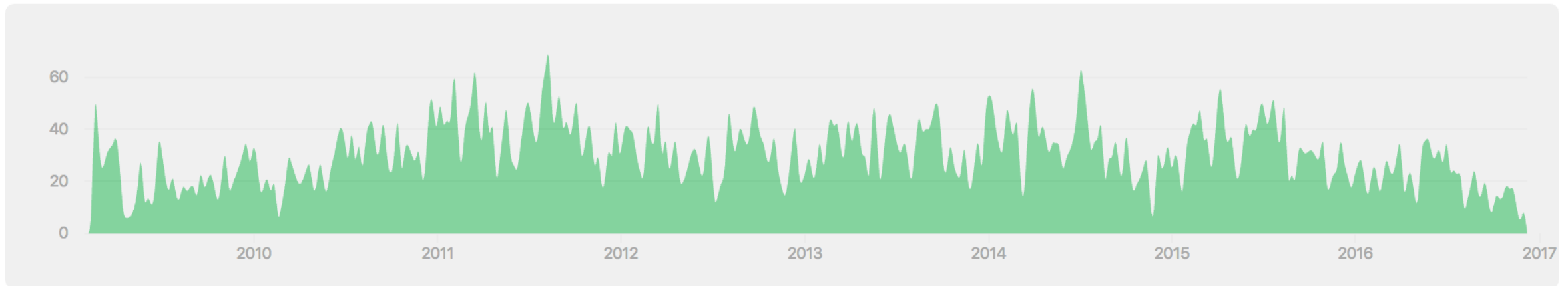
Key Facts

DataStax: The World Leader in Apache Cassandra™ Engineering

Mar 1, 2009 – Jan 8, 2017

Contributions: **Commits** ▼

Contributions to trunk, excluding merge commits



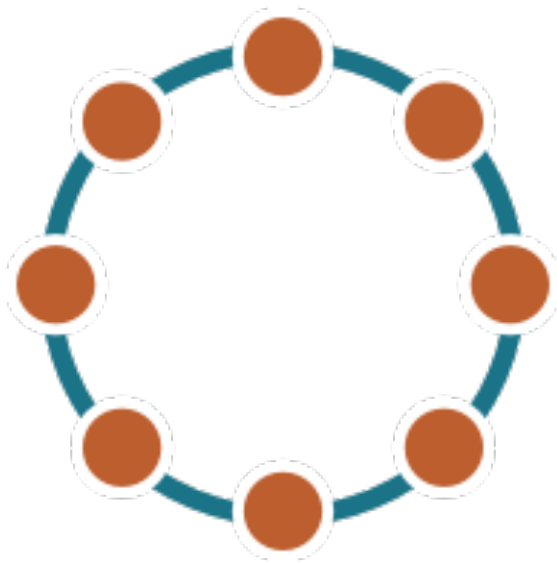
Contributors

110 Total
22 DataStax Employees (20%)

Contributions

10,423 Total
8,950 DataStax Employees (86%)

DataStax Enterprise (DSE)

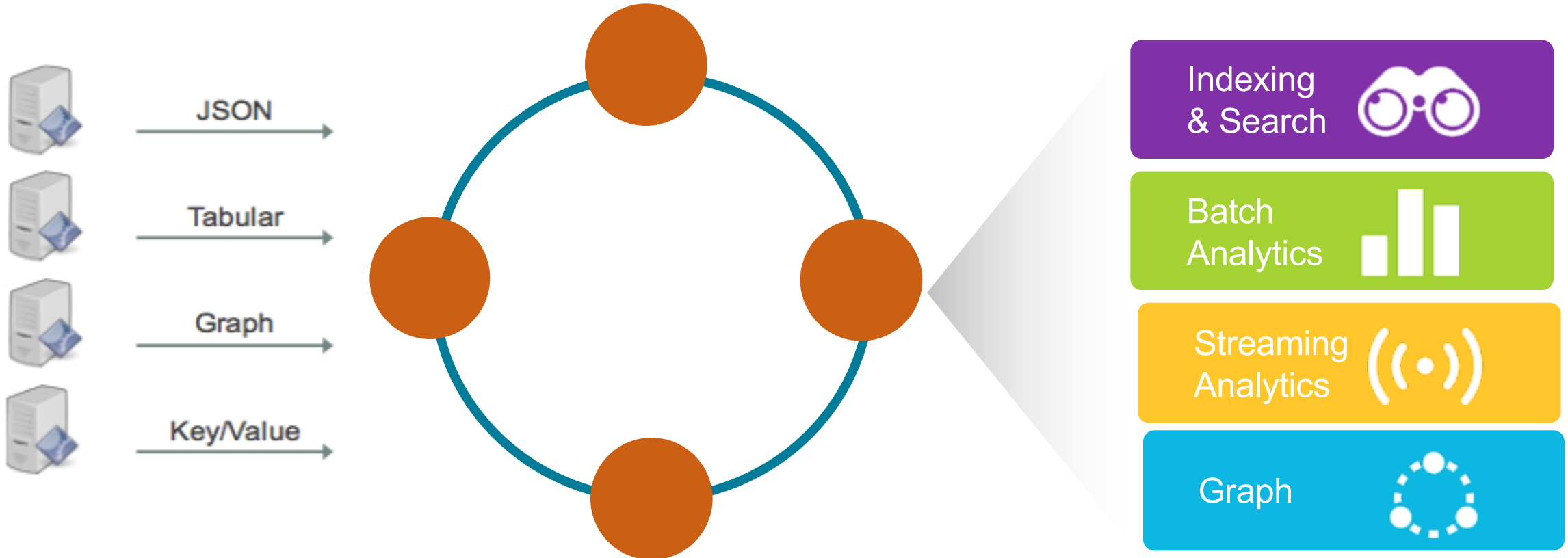


Based on the best distribution of Apache Cassandra™

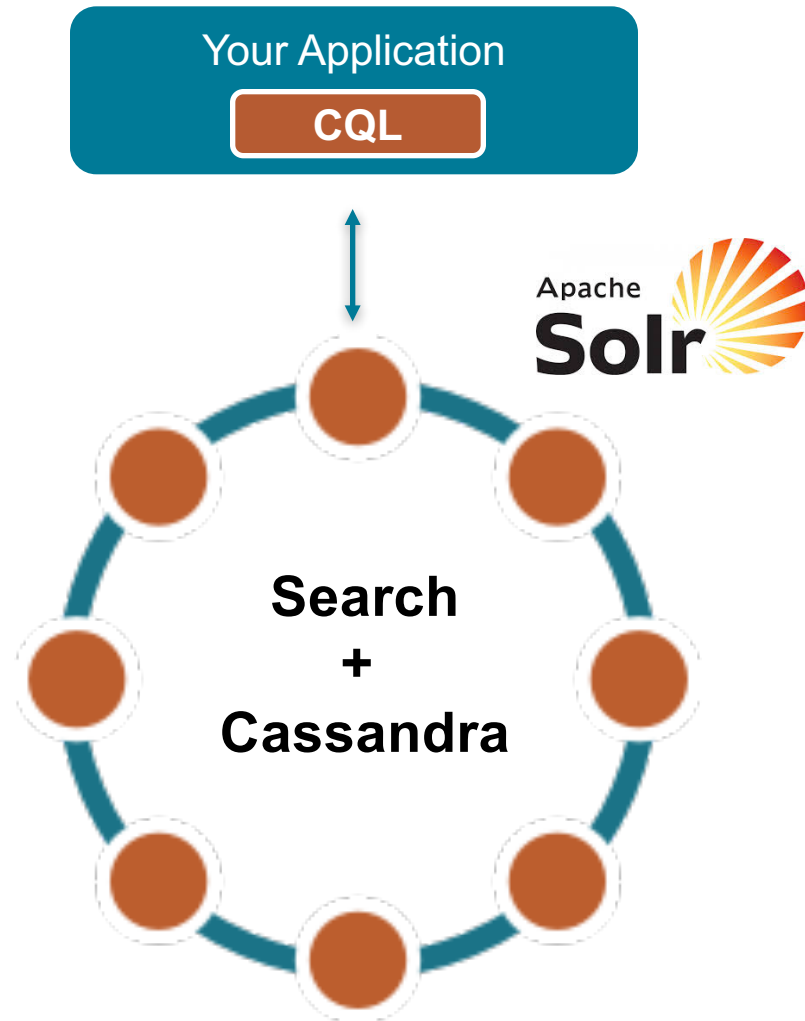
- Continuously Available
- Master Less
- Fully Distributed
- Shared-Nothing Architecture
- Build In Replication
- Linear Scalability
- Scale out



Integrated Multi-Model/Mixed Workload Platform



DSE Search



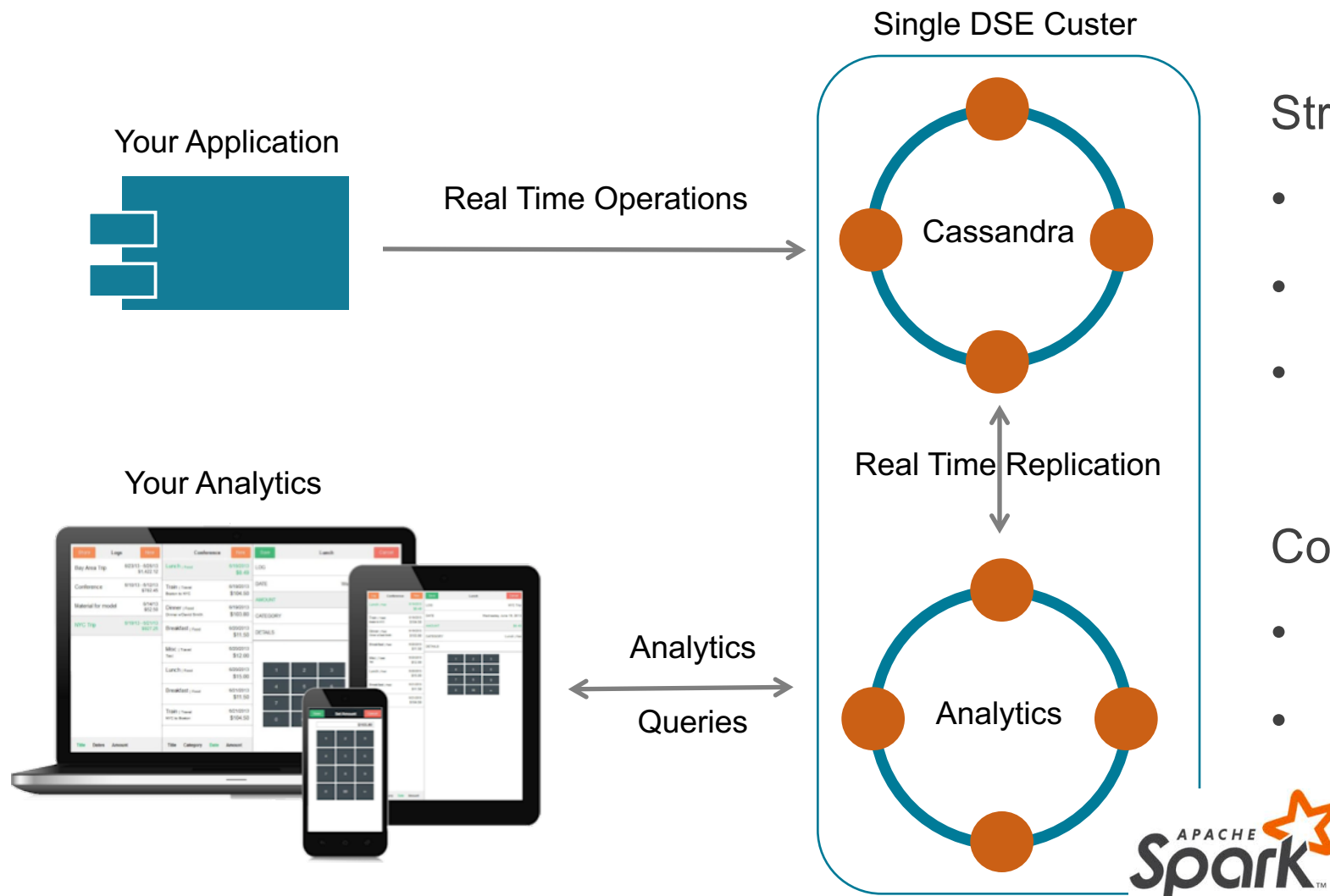
Live indexing engine with powerful search

- Automatic indexing on insert
- Higher ingestion throughput
- Distributed query optimization

Compared to self-managed:

- No separate search cluster to manage
- Probably less total hardware required
- No “Split Brain” data inconsistencies
- No ETL or synch to build and maintain
- No app level data management code

DSE Analytics



Streaming, ad-hoc, and batch

- High-performance
- Workload management
- SQL reporting

Compared to self-managed:

- No ETL
- True HA without Zookeeper

DSE Graph

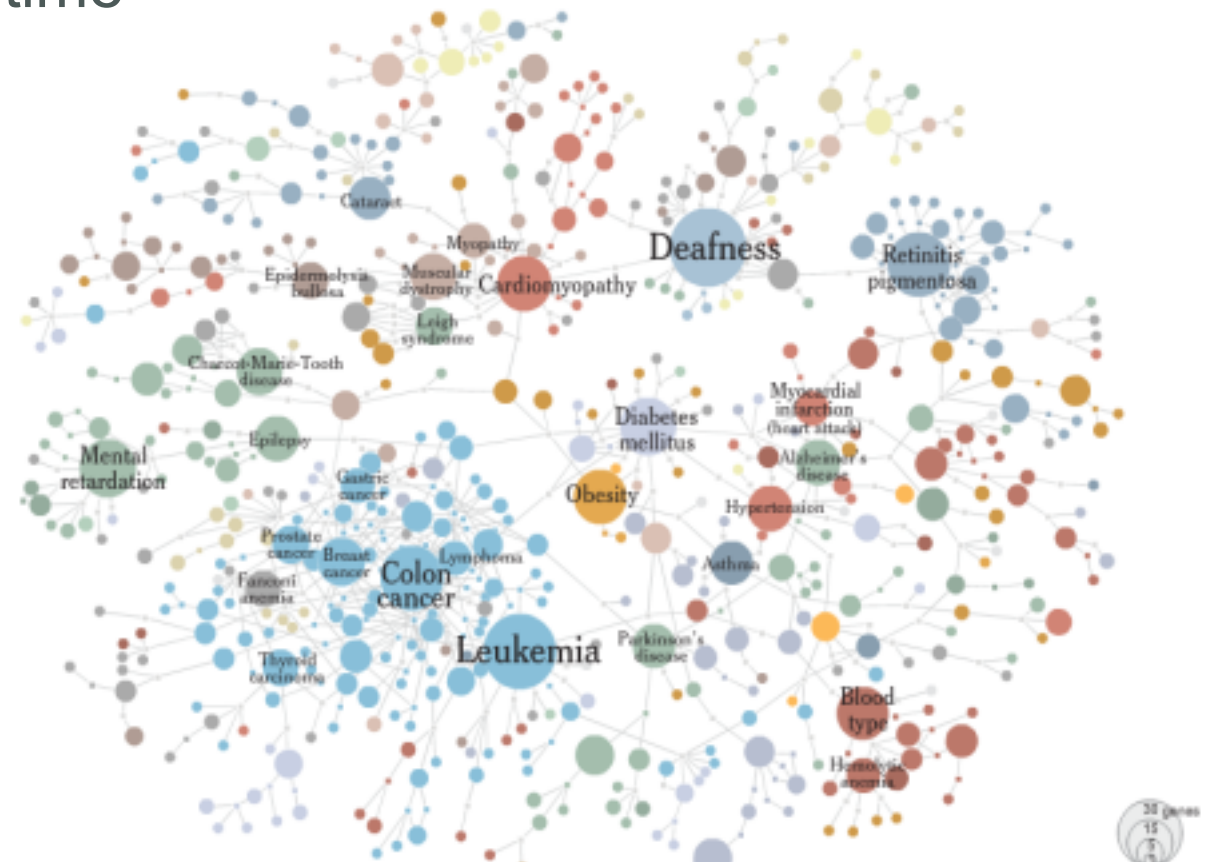
A scalable, distributed graph database that is optimized for storing, traversing and querying complex graph data in real time

- Value data between relationships
- DSE Analytics and Search integrated
- Perfect for use cases:

Customer360

Recommendations

Fraud Detection



Agenda

| | | |
|---|---|-------|
| 1 | Workshop: Apache Cassandra™ Data Structures and Replication | 10:15 |
| 2 | Workshop: DataModelling with DataStax Enterprise and Apache Cassandra™ | 12:00 |
| | BREAK | 13:00 |
| 3 | Workshop: DataStax Enterprise Search | 14:00 |
| 4 | Workshop: DataStax Enterprise Analytics with Apache Spark | 15:00 |
| 5 | Discussion, Questions and Answers | 16:00 |