Selecting database technologies for mere mortals

Choosing the best tool for the job

Kenny Gorman
Chief Technologist for Data at Rackspace
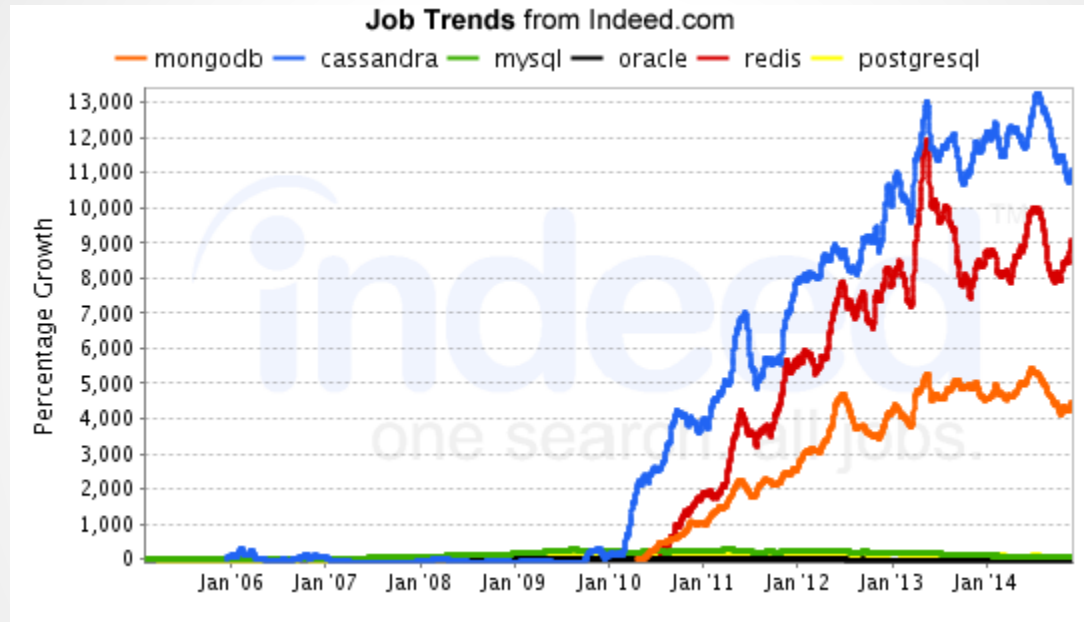Co-Founder, ObjectRocket

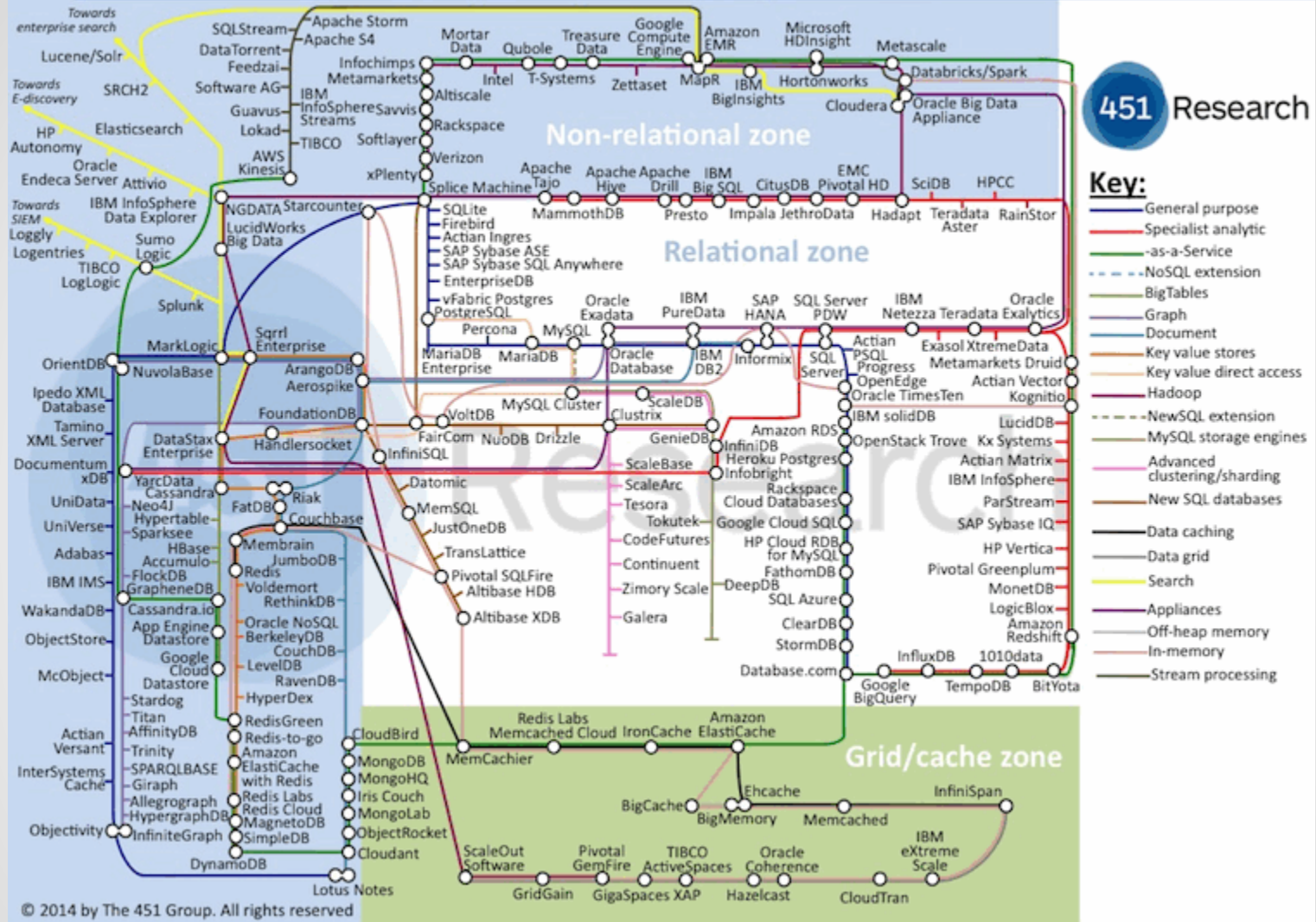@rackspace @kennygorman

In the olden days

One choice, or a few choices

These days

oh the choices
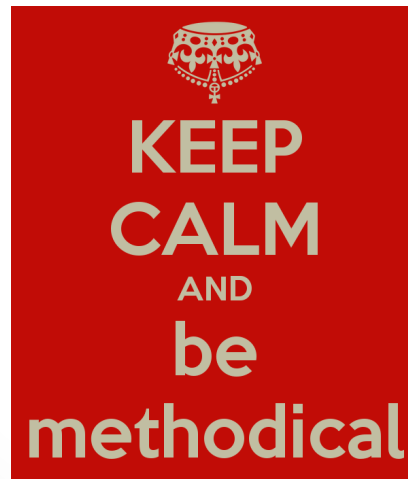
Job Trends from Indeed.com

— mongodb — cassandra — mysql — oracle — redis — postgresql

Percentage Growth

Data Platforms Landscape Map – February 2014

© 2014 by The 451 Group. All rights reserved

How do I make the correct choice?

Caveat: Sometimes it's obvious.

So forget the rest of this presentation, but please don't leave just yet.

Take small bites

- Think in terms of:
  - Projects
  - Workloads
  - Products
  - Teams

- Otherwise you will make yourself crazy and wish you went into gardening or became a welder.

Breaking it down

- Splitting the problem into manageable chunks

- Finding data stores for each chunk

  - Knowing the workload
  - Knowing the capabilities of each data storage mechanism

Knowing the workloads

- Understand the top N patterns
  - Example: index range scan

- You can use a variety of data stores

- Understand your consistency patterns
  - Example: count() in PG

Knowing the options

- MongoDB: document database
- Cassandra: Easy HQ, k/v columnar
- Hadoop: Analytics, batch insights
- Redis: in memory, fast, simple, k/v
- Oracle: for people with cash
- PostgreSQL: RDBMS, open, no drama
- MySQL: RDBMS, drama
- ...

Knowing your strengths

- Developer chops

- DevOps chops

- Still have to run a business; this isn't CS102 class here people

Growing over time

- Day one, with an empty DB it's easy

- As you grow, you will sleep less

- Sysbench FTW

- Slow rolling; go slow and safe!

The Matrix

- No not that matrix

- Getting empirical
  - Subjective
  - Qualitative
  - Record of your decisions

rackspace®
the #1 managed cloud company

ObjectRocket
by Rackspace

The Matrix

	B	C	D	E	F	G
			MongoDB		PostgreSQL	
	Attribute	**Weight**				
	Transactions	5	0	0	5	25
	Slave Reads	3	5	15	5	15
	Built in Sharding	2	5	10	0	0
	MVCC	2	0	0	5	10
	Atomic Operators	4	4	16	5	20
				41		70

Shutterfly Case Study

Shutterfly Case Study

- Understand the use case
- Existing system, migration required
- Data patterns
  - Materialized paths
    - Ensure good plans for hierarchical structures
  - XML payloads
    - Making updates win

Shutterfly Case Study

- Sit with developers
- Understand patterns
- Test patterns
- No guessing, no hype!
- I said, no hype!

http://www.mongodb.com/presentations/sharing-lifes-joy-using-mongodb-shutterfly-case-study

Contact

@kennygorman

@rackspace

kenny.gorman@rackspace.com

http://developer.rackspace.com/databases/

ObjectRocket
by Rackspace