



NoSQL Databases

in the Cloud

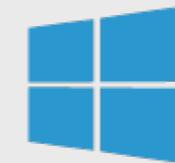


Inés Sombra

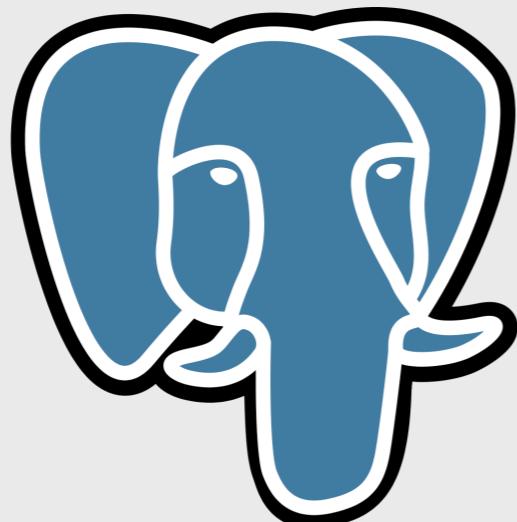
INES@ENGINEYARD.COM

@RANDOMMOOD

Engine Yard™



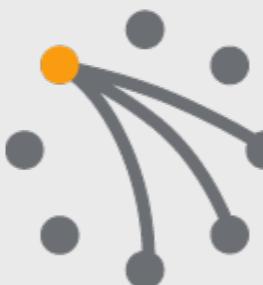
Windows® Azure™

 terremark®

MySQL®



Java™

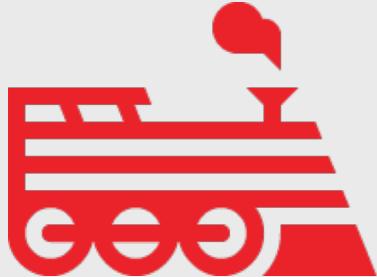
The PHP logo is the word "php" in a bold, black, sans-serif font inside a light blue oval.

riak

The node.js logo consists of the word "node" in a dark grey sans-serif font with a green hexagon icon, followed by ".js" in a green hexagon icon.The redis logo shows a stack of three red rectangular blocks with white symbols on them, and the word "redis" in a dark grey sans-serif font below it.

mongoDB

Today's Agenda



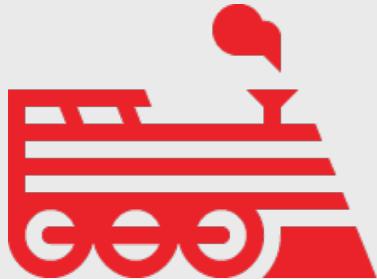
- NOSQL DATABASES
- MONGODB
- RIAK
- DEVELOPMENT IN THE CLOUD

A photograph of two young boys jumping into a large body of water from a wooden dock. The boy on the left is wearing red swim trunks and has his arms outstretched. The boy on the right is wearing white swim trunks with a black floral pattern and is also jumping. In the background, there is a shoreline with many houses built on a hillside. A jet ski is visible in the water. The sky is blue with some clouds.

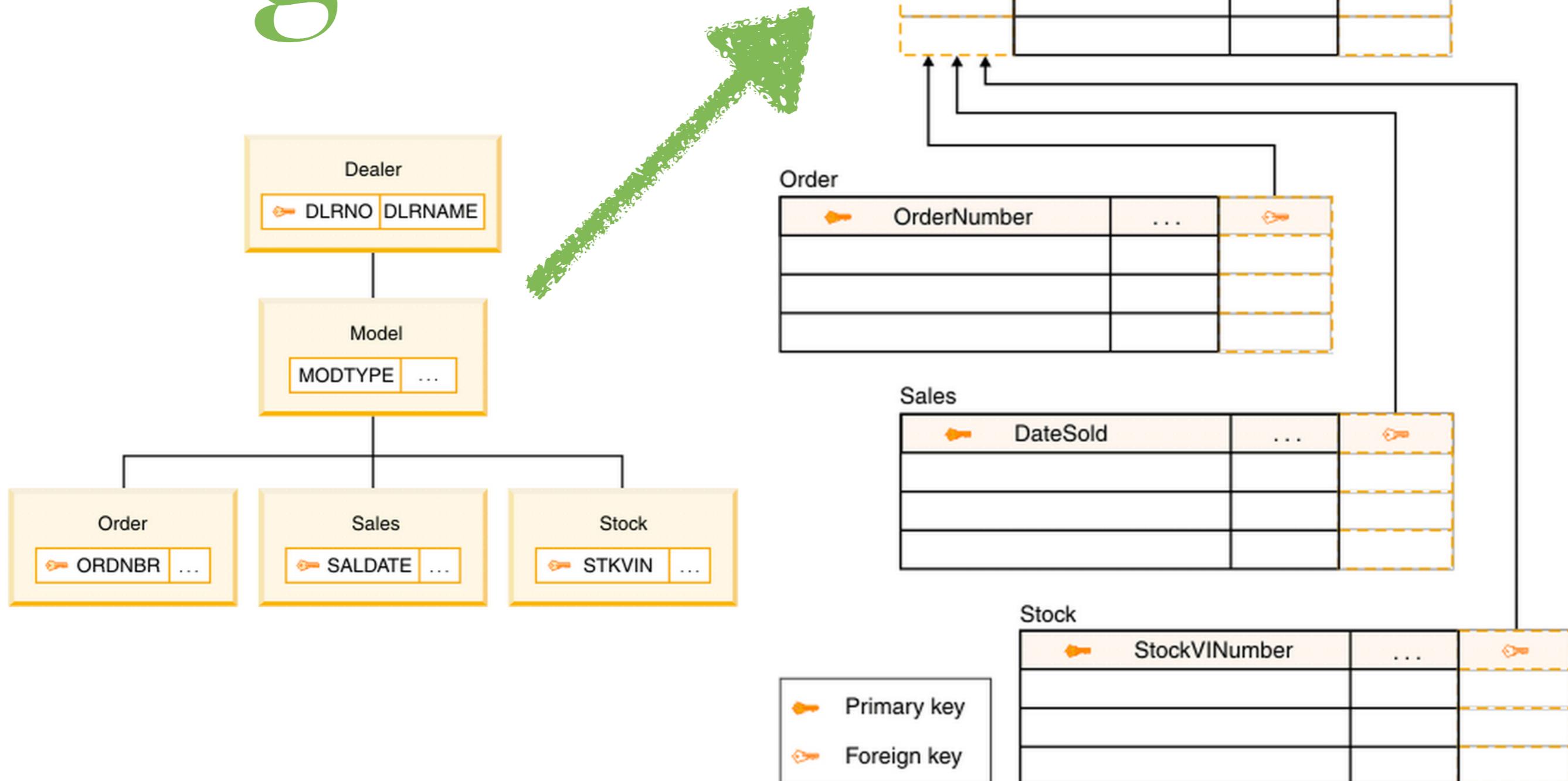
Let's
Jump In



NoSQL Intro & Tenets

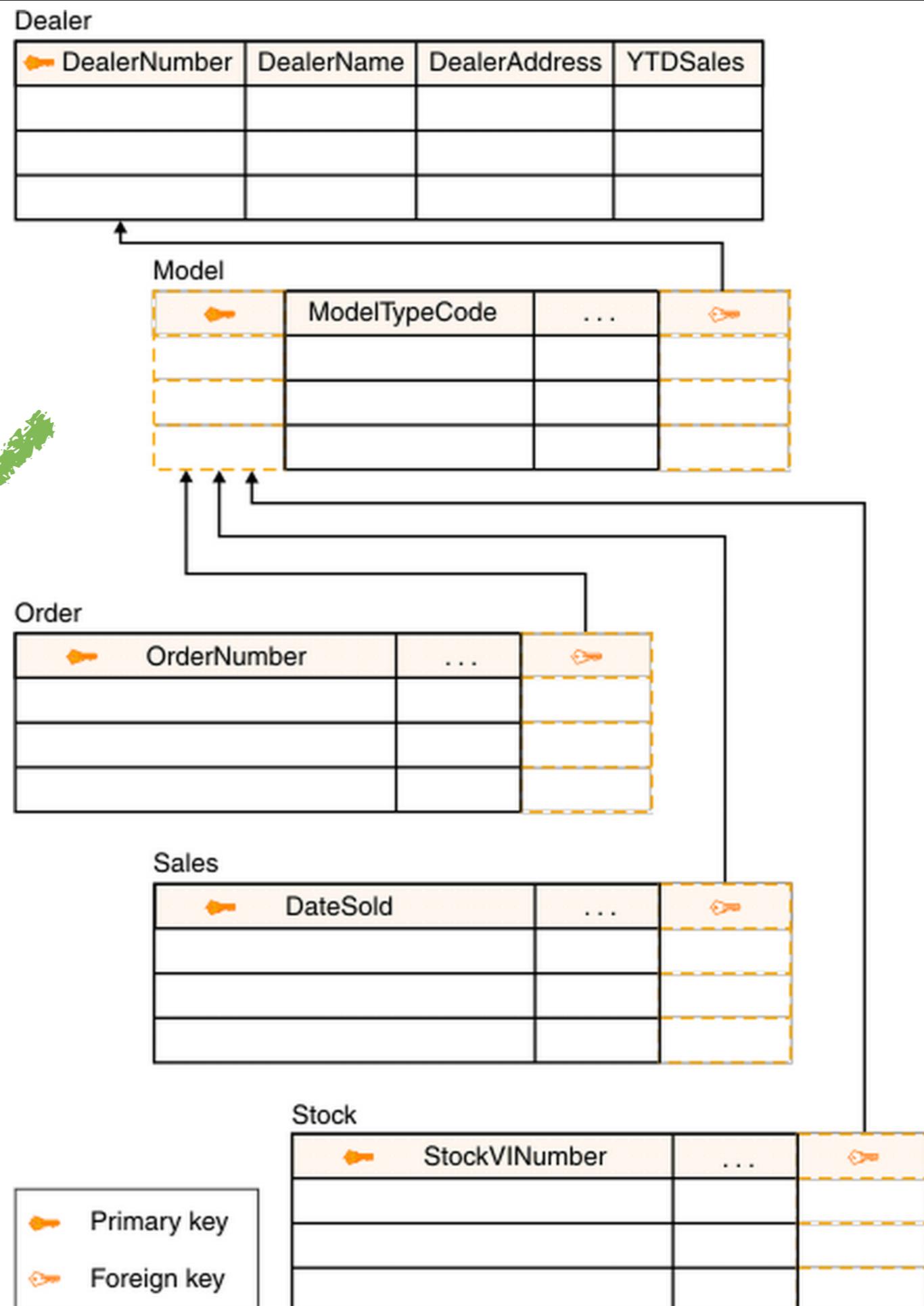


Relational Origin



NoSQL Origin

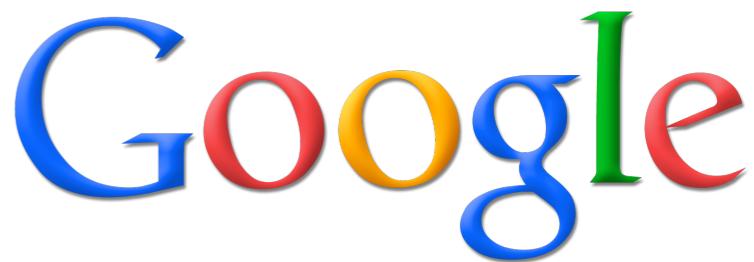
```
Cars
{
  "_id" : 1234,
  "dealership": 3423,
  "year": 2013,
  "make": "10gen",
  "model": "Mongos",
  "vin": 3928056,
  "mechanicNotes": "Runs great!" },
{
  "_id" : 54321,
  "dealership": 3423,
  "year": 1985,
  "make": "DeLorean",
  "model": "DMC-12",
  "vin": 8056309,
  "mechanicNotes": "Great Scott!" }
```



Nosql Databases

PRAGMATIC BEGINNINGS
ADDRESS RELATIONAL LIMITATIONS

NO UNIFIED MODELS FOR
QUERYING & DATA



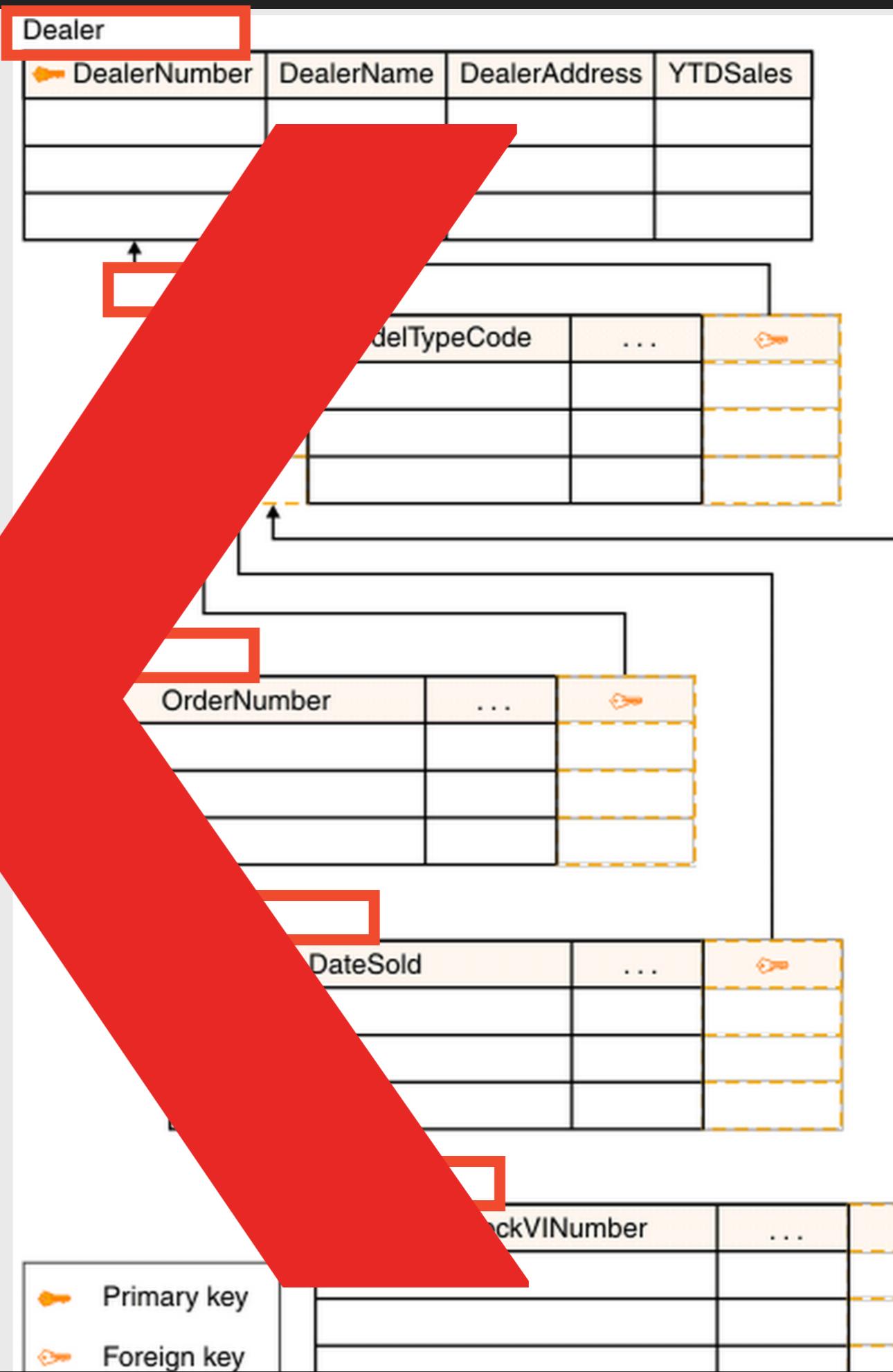
Relationshipal Concepts

ENTITIES

ATTRIBUTES

RELATIONS

CONSTRAINTS



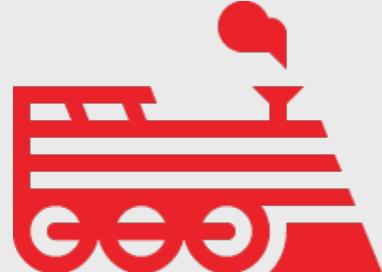
NoSQL Properties

CONSISTENCY

AVAILABILITY

PARTITION TOLERANCE

*What Matters To You When A
Partition Happens?*



Many Data Models

KEY/VALUE

DOCUMENT ORIENTED

COLUMN ORIENTED

GRAPH

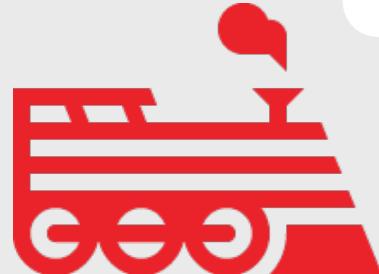
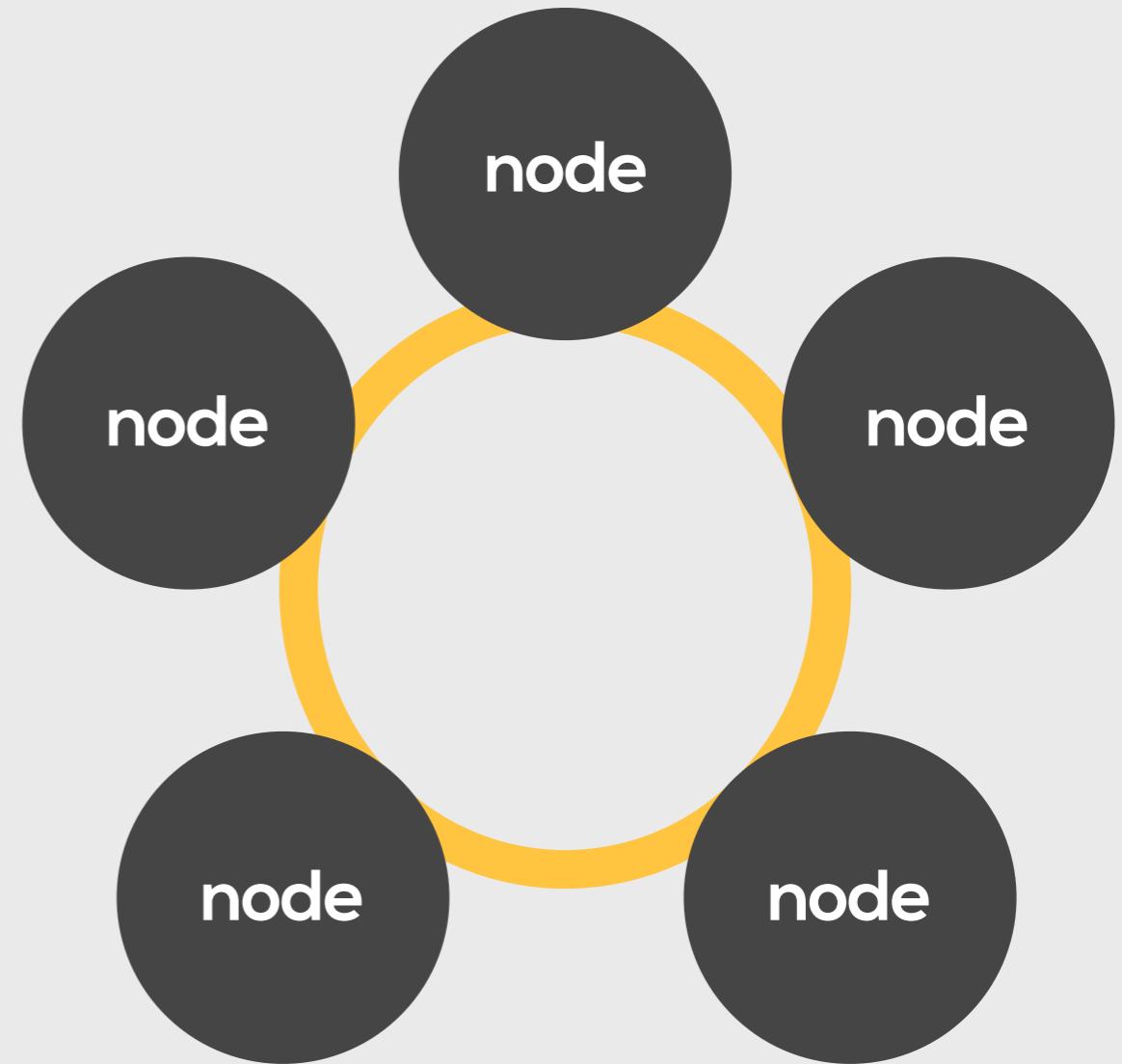


Consistency Models

SINGLE MASTER



MULTI-MASTER

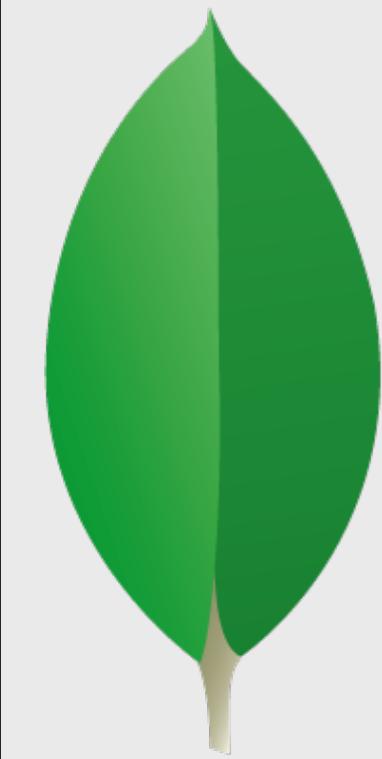




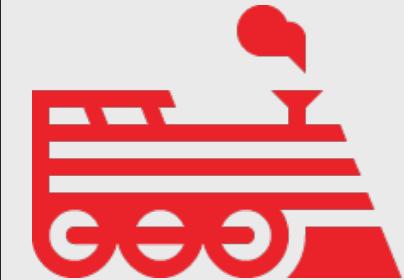
Developing & Running

NOSQL DBS IN THE CLOUD
(AND A LITTLE OPS)





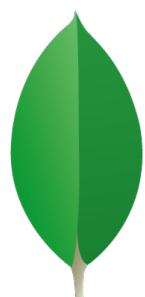
mongoDB



A Little Context

DOCUMENT ORIENTED DB
BINARY JSON

"SCHEMA-LESS"
FLEXIBLE SCHEMAS



mongoDB * *Trivia Q: Why 'Mongo'?*

Pro: Dreamy (For Devs)

DEVELOPER FRIENDLY

MAP/REDUCE

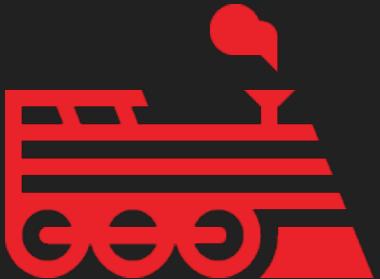
GEOSPATIAL DATA

AGGREGATION FRAMEWORK



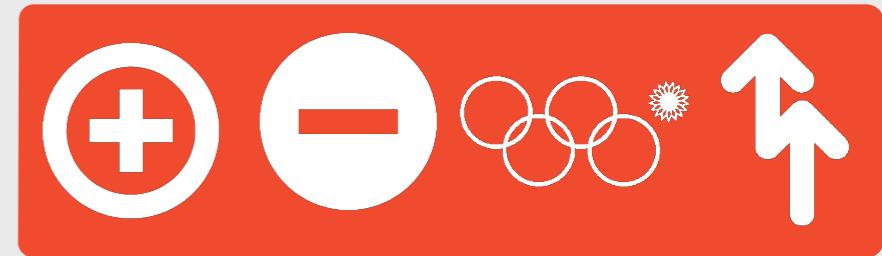
Detour

HOW DOES A GIVEN
DATABASE HANDLE
GROWTH
REDUCTION
NODE FAILURES
UPGRADES



(THINKING IN OPS)

Ops For Devs



REPLICA SETS

**AUTOMATIC
FAILOVER**

NODE RECOVERY

Secondary

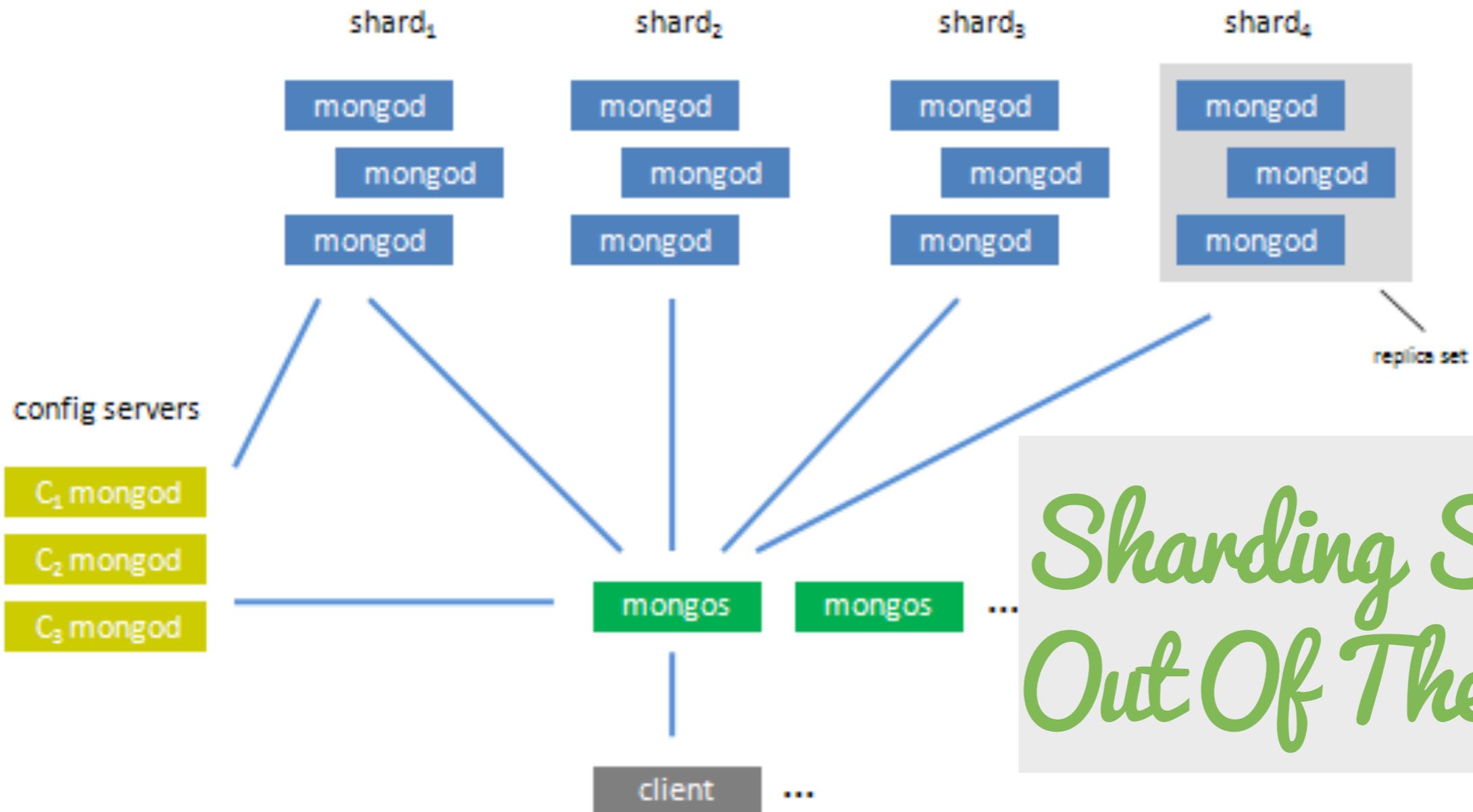
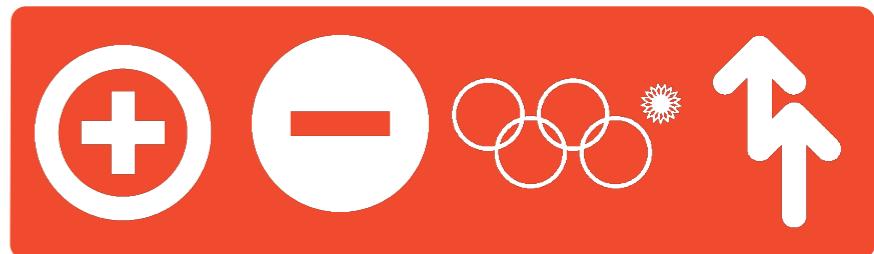
Master

Secondary

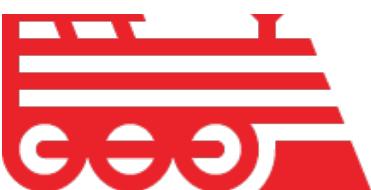


mongoDB

Ops For Devs



*Sharding Support
Out Of The Box!*



Pro: There's More

HANDY SHELL 'MONGO'
ALL THEM JAVASCRIPTS

VIBRANT COMMUNITY
TONS OF HOSTING OPTIONS &
RESOURCES



Cons



**COLLECTION LEVEL
LOCKING LIMITS DB SCOPE**

**SHARDING IS THE ONLY WAY
TO SCALE WRITES**

SOMETIMES NAIVE APPROACH



Common Mistakes



NOT USING A FULL (3) SET

USING RELATIONAL SCHEMAS

DISABLING THE JOURNAL

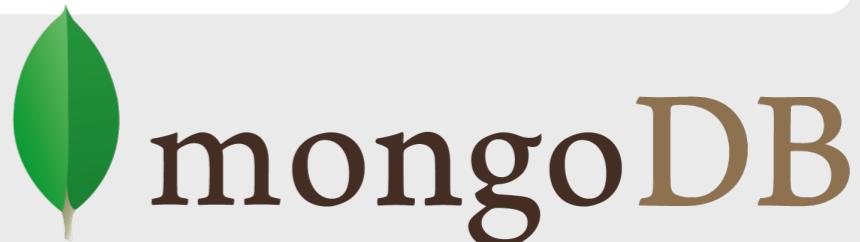


Best Use Case



When you need to
store **json data** & want
to richly query it.

* and your app does more reads than writes



General Cloud Tips

NOT LIKE PHYSICAL HARDWARE
CLOUD RESOURCES VIRTUALIZED

MONITORING MATTERS A LOT
CAPACITY PLANNING TOO



* *Trivia Q: What is a PaaS?*

Cloud Tips

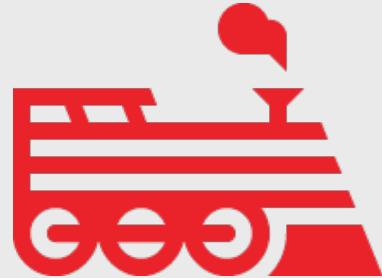


**KEEP WORKING SET IN MEMORY
DATA + INDEXES**

**SIZING YOUR CLUSTER
64-BIT ONLY - KEEP CURRENT**

**SCALE UP IF METRICS SHOW IT
MMS IS HANDY**





A Little Context

**KEY/VALUE DATABASE
VALUE AGNOSTIC**

**EXTREMELY FAULT-TOLERANT
ALL NODES EQUAL**



* *Trivia Q: Who Did Dynamo?*

Pro: Dreamy (For Ops)

FAILURES ARE OK

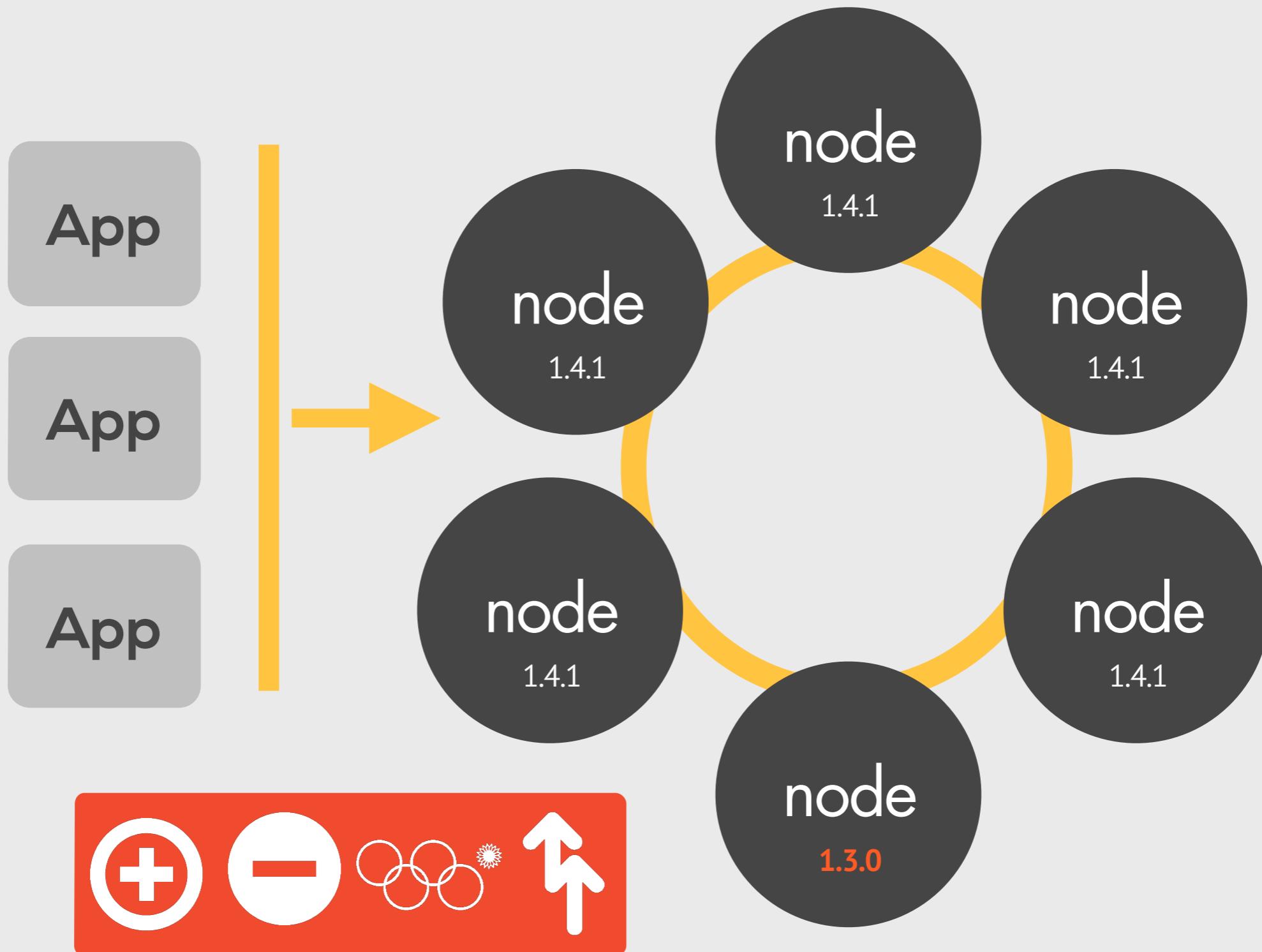
MAP/REDUCE

SECONDARY INDEXES

FULL TEXT SEARCH



Ops For Devs



Pro: Highly Available

ALWAYS RESPONDING
MULTI-DATACENTER REPL

STORAGE BACKENDS
BITCASK, LEVELDB, MEMORY



* Trivia Q: is MDC free?

riak

Cons



MINIMUM CLUSTER-SET IS 5 NODES

EVENTUAL CONSISTENCY TAKES SOME TIME TO GROK

**QUERYING STORY IS MEH
GETTING BETTER W FTS**



Common Mistakes



NOT USING 5 OR MORE NODES

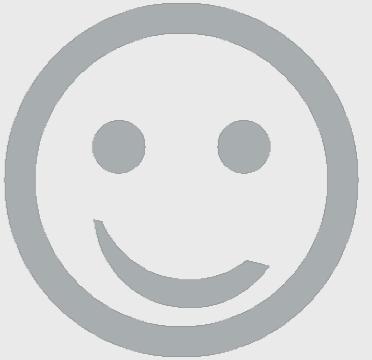
CHOOSING WRONG BACKEND

PROTOCOLS FOR DEV VS PROD

LEAVING DEFAULT RING SIZE



Best Use Case



When you need a
highly available DB

*and your app does not need a lot of introspection on values



Cloud Tips



CLUSTER SIZE MAY NEED MORE
BIGGER RINGS HELP

MIND YOUR DATA LOCATION
EBS VS EPHEMERAL DRIVES

CHECK OUT BASHO'S CLOUD DOCS
THEIR DOCS ARE GREAT





Let's Wrap Up

(+ RESOURCES & Q&A)



Another Advantage



A photograph of a woman with long dark hair, wearing a grey cardigan over a black top. She has a large, dark, fake mustache applied to her upper lip. She is looking directly at the camera with a neutral expression. The background shows an office environment with cubicles and other people working.

Questions?

THANK YOU!

GITHUB.COM/RANDOMMOOD/
GREATWIDEOOPEN2014

