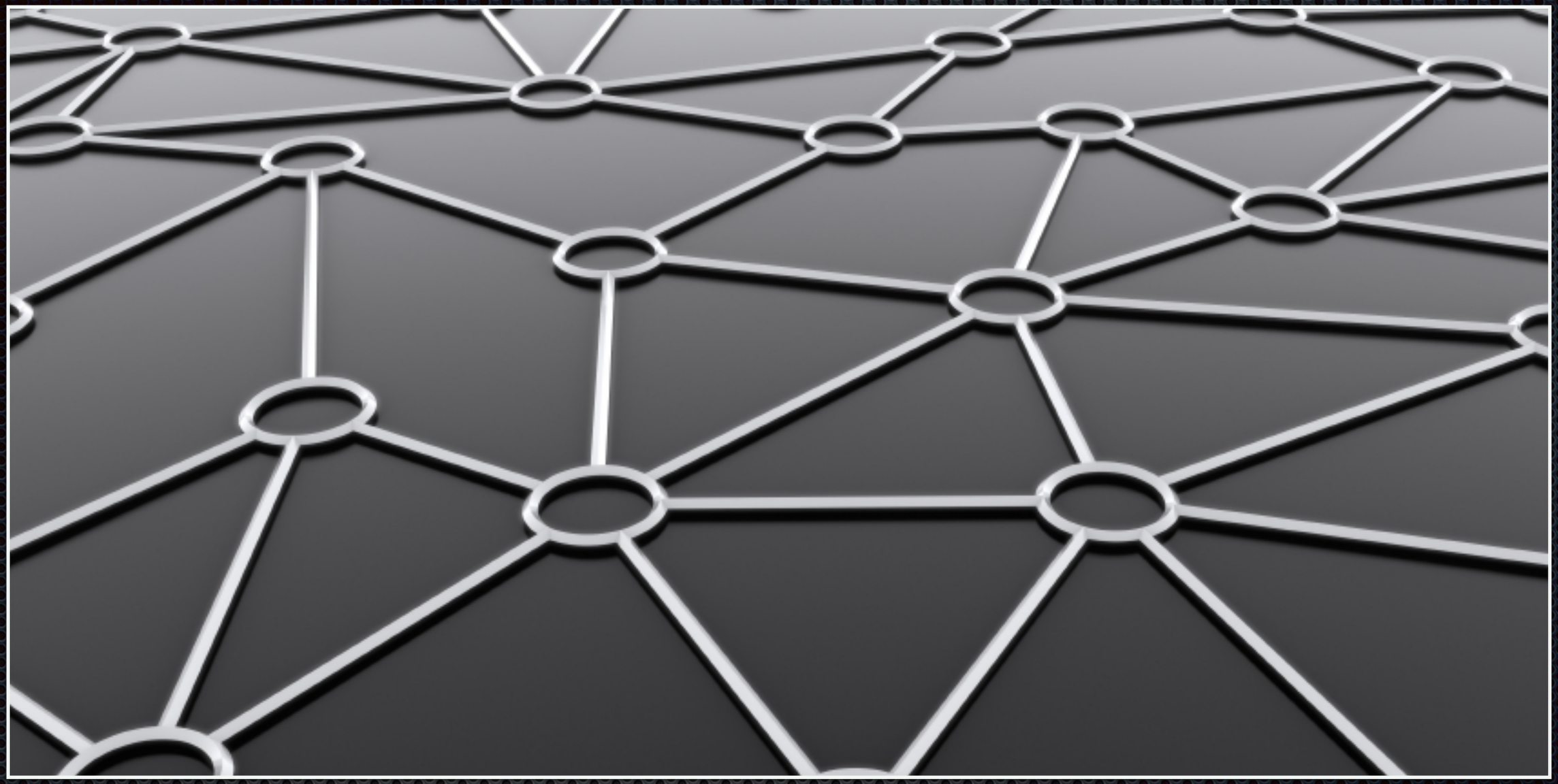
Conflict Free Replicated Data-types in Eventually Consistent Systems

Joel Jacobson, Basho Technologies

@joeljacobson
#NoSQLrs
London
2013



Riak

Distributed, Masterless, Key/Value Database
+ Extras

Key	Value
Key	Value
Key	Value
Key	Value
Key	Value
Key	Value

Buckets, Keys and Values

Simple operations;
GET, PUT, DELETE

Distributed & Scalable

Fault Tolerance

- Any node can serve R/W requests
- Data is replicated
- Fallback Node
- Hinted Handoff

Dynamo Systems

- Distributed
- High Availability
- Masterless
- Eventually Consistent

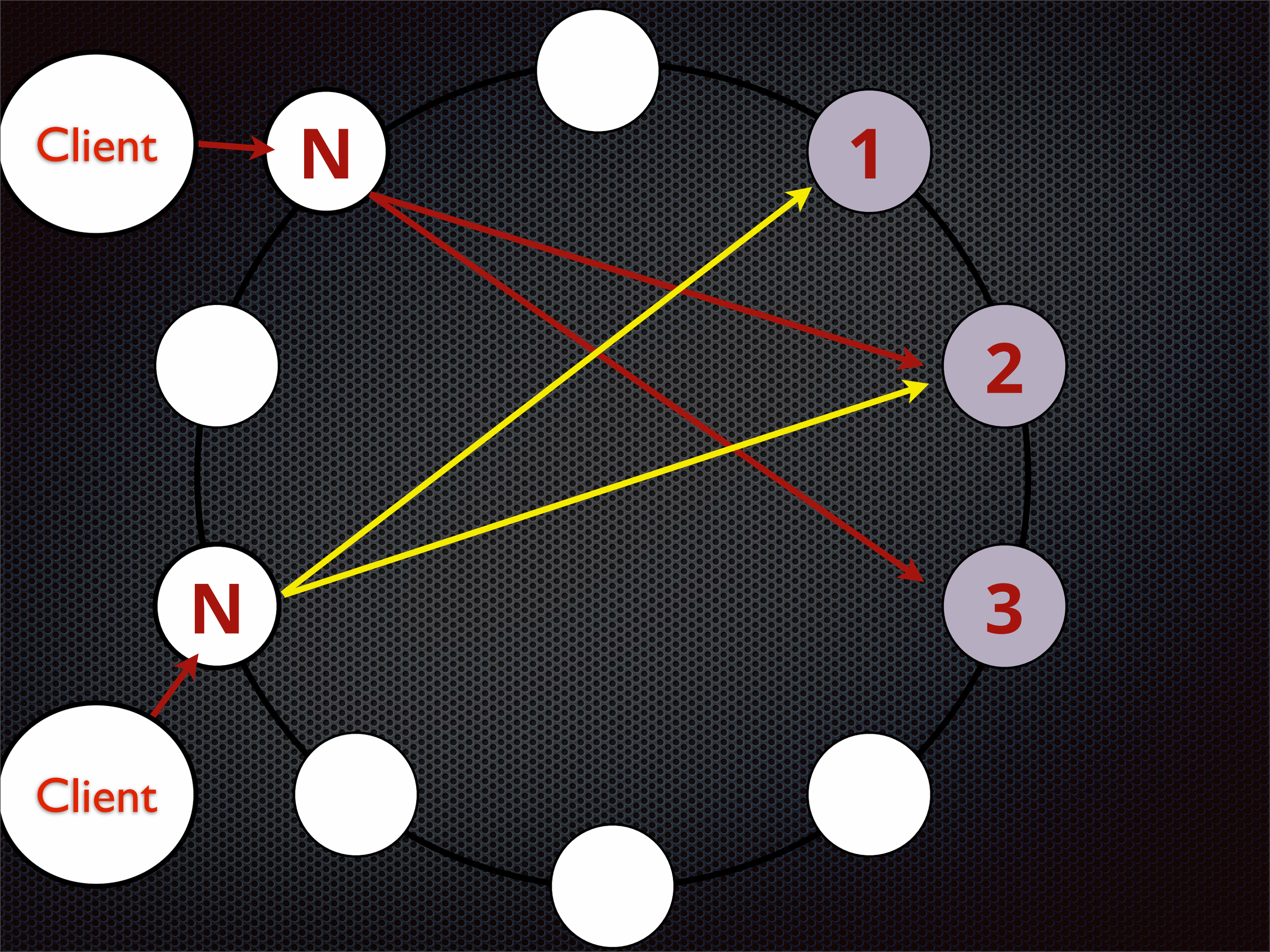
Client

A

1

2

3


Client
N
1
2
3
N
Client

Conflicts

Conflict Resolution

- Last-Write Wins (LWW)
- Vector Clocks (Siblings)

Sibling Resolution

Conflict-Free Replicated Data Types

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A comprehensive study of Convergent and Commutative Replicated Data Types

Marc Shapiro, INRIA & LIP6, Paris, France
Nuno Preguiça, CITI, Universidade Nova de Lisboa, Portugal
Carlos Baquero, Universidade do Minho, Portugal

Marek Zawirski, INRIA & UPMC, Paris, France

- 13 Jan 2011

http://hal.upmc.fr/docs/00/55/55/88/PDF/techreport.pdf

Primitives

- Counters
- Sets
- Maps
- Registers
- Booleans

Counters

```
{
  'type': 'g-
counter',
  'e': {
    'a': 1,
    'b': 5,
    'c': 2
  }
}
```

```
{
  'type': 'pn-counter',
  'p': {
    'a': 10,
    'b': 2
  },
  'n': {
    'c': 5,
    'a': 1
  }
}
```

P=12, N=6, value = 6

The Problem of Absence

Sets

```
{
  'type': 'g-set',
  'e': ['a', 'b', 'c']
}
```

```
{
'type': 'or-set'
[ {a, 3}, {b, 2}, {c, 2}],          %% Version Vector
  [{ 'bob', [{a, 1}]},              %% Bob was added by
'a' at logical time 1 (a's first event)
  {'joel',  [{a, 3}]},              %% Joel was added by
'a' at logical time 3
  {'sean', [{a, 2}, {c, 1}]},
  {'matt', [{b, 1}]},
  {'stu', [{c, 2}, {b, 2}]}]
]}
```

VV, element->dots

Maps

```
{
  'type': 'map',
          gold=100, stone=10,
   weapons=[sword, knife, shotgun],
          armour=[helmet, shield]
}
```

```
Hey Riak with Map at Key K ->

[{increment gold by 10, decrement
stone by 4, add uzi to weapons and
remove helmet from armour}]
```

Use-Cases

- Advert clicks (G-Counter)
- Shopping cart (Modified OR-Set)
- Logged in users (P-N Counter)
- Maps can be composed of complex data

Common Questions

- What do they cost me?
- How big are they?
- What can’t they do?

Riak 2.0

- Riak Data Types
- Strong Consistency
- New Full-Text Search (Solr)
- Security

Useful links

http://hal.upmc.fr/docs/00/55/55/88/PDF/techreport.pdf

http://arxiv.org/pdf/1210.3368.pdf

https://gist.github.com/russelldb/f92f44bdfb619e089a4d

http://gsd.di.uminho.pt/members/cbm/ps/scadt3.pdf

http://arxiv.org/abs/1011.5808

Thank you

joel@basho.com