

MongoDB ve Diğer Veritabanlarında Sharding



www.spp42.com

Who the f**k is talking?

- Emir Karaburçak
 - emir.karaburcak@spp42.com
 - @kinchil
- SPP42 de Yazılım Geliştirme Uzmanı
- Python, Django, Java, JBoss Seam, Play
- MongoDB, PostgreSQL, Hibernate

Konular

- Sharding Nedir?
 - Problem
 - Çözüm
 - Shard
 - Veriyi Bölmek
 - Veriyi Dağıtmak
 - Chunk
 - Cluster
 - Shard Key

Veritabanları

- MongoDB
- CouchDB
- Neo4j
- MySQL

Problem

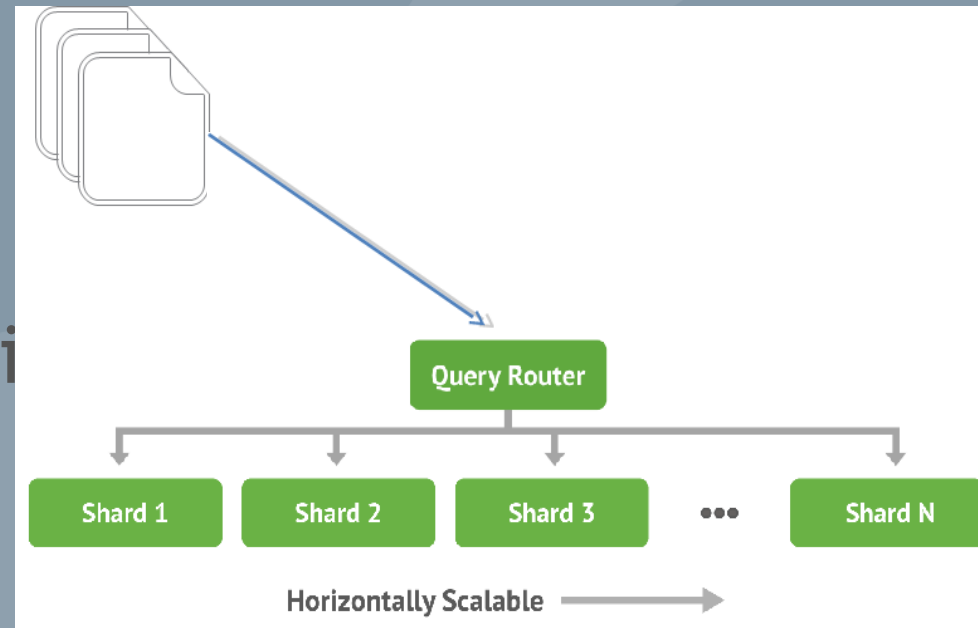
- Flickr - 10milyar fotoğraf
- Instagram - Android uygulamasında ilk 12 saatte 1milyon yeni kullanıcı

Çözüm

- Dikey Ölçeklendirme
- Yatay Ölçeklendirme

Sharding Nedir?

- Basit olarak büyük bir collectionı birkaç sunucu (cluster) arasında bölmek
- Sharding > Partitioning
- Her işlem otomatik



Shard nedir?

- Bir clusterdaki verinin bir alt kümesinden sorumlu 1 veya daha fazla sunucu
- Eğer 1den fazla sunucu varsa, hepsi aynı veriye sahiptir (replica set)

Veriyi Bölmek

- Matematiksel olarak; ['a', 'h')
- 'a' dahil olmak üzere 'a' dan başlayarak 'h' dahil olmamak üzere 'h' ye kadar
- Belirli bir aralığa 'chunk' denir

Veriyi Dağıtmak

Birinci Yöntem

- 1 shard 1 aralıktan sorumludur



- Sorunlu bir yöntem

Veriyi Dağıtmak

İkinci Yöntem

- 1 shard 1 veya daha fazla aralıktan sorumludur



- MongoDB'nin kullandığı yöntem

Chunklar Nasıl Yaratılır?

- Aralıkları belirlemek için bir key gereklidir.
- Bu key in adı "shard key"dir
- Shard key her bir alan veya alanlardan oluşabilir.
- Her chunk 200mb dır, yeni bir shard yaratılması için 1 shardın diğerlerinden +9 chunk a sahip olmalıdır

- Her chunkın aralığı distinct olmalıdır
- Herhangi 2 chunk kesişen aralığa sahip olamaz
- Her chunk bir sonraki chunkin aralığını tatmin etmelidir
- null < numbers < strings < objects < arrays < binary data < ObjectIds < booleans < dates < regular expressions

Balancer

- Chunkları bir sharddan diğerine taşır
- Otomatik balancing
- Veriyi eşit olarak dağıtmakla ve olabildiğince az veri taşımakla yükümlüdür

mongos

- Kullanıcı ve cluster arasında köprü
- Tüm read/write lar mongos a gider
- Özellikle belirtilmediği takdirde shardlara direkt ulaşılmaz
- Sharda direkt ulaşmak için query de shard key kullanılır(targeted query)
- Eğer query de shard key yoksa query tüm shardlara gönderilir(spewed query)

Config Server

- Config Serverlar, özel mongod lardır
- Clusterların açıklayıcı bilgilerini tutar
- Veri taşıma için tüm config serverların ayakta olması gereklidir

Özet Olarak Cluster

- Veri depolama için; Shard a
- İstek yönelendirme için; mongos a
- Durum bilgileri için; Config Server a

İhtiyaç duyar

Shard Key

- Kötü shard key = kötü sharding
- Shard Key belirlerken en önemli nokta kardinalite
- Eğer bir shard key in N kadar değeri varsa, en fazla N kadar chunk ve N kadar shard olabilir

Shard Key hakkında Önemli Noktalar

- Devamlı artan bir shard key iyi bir key değildir
- Shard Key belirlerken kardinalite en önemli unsurdur
- Rastgele değerlere sahip bir shard key iyi bir key değildir
- Coarsely ascending key + search key

Hatalı Shard Key

- 2010 Foursquare olayı
- 17 saatlik downtime
- 3M kullanıcı, 200M checkin, günde 18K yeni checkin

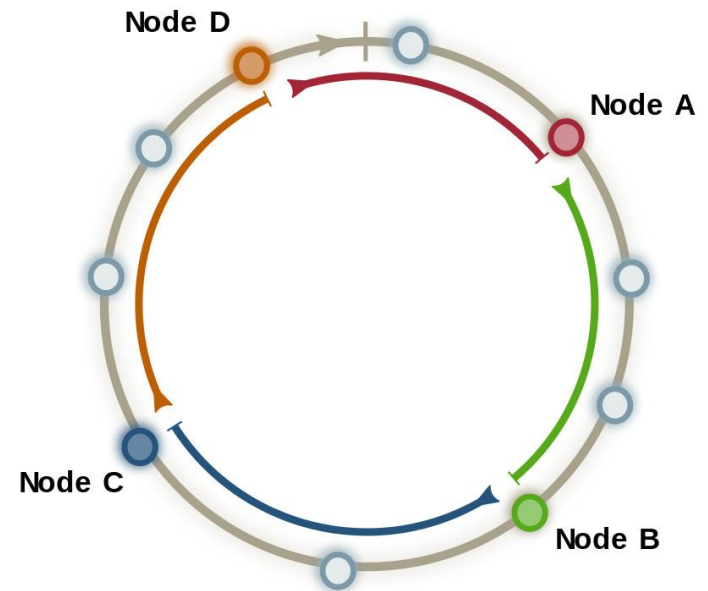
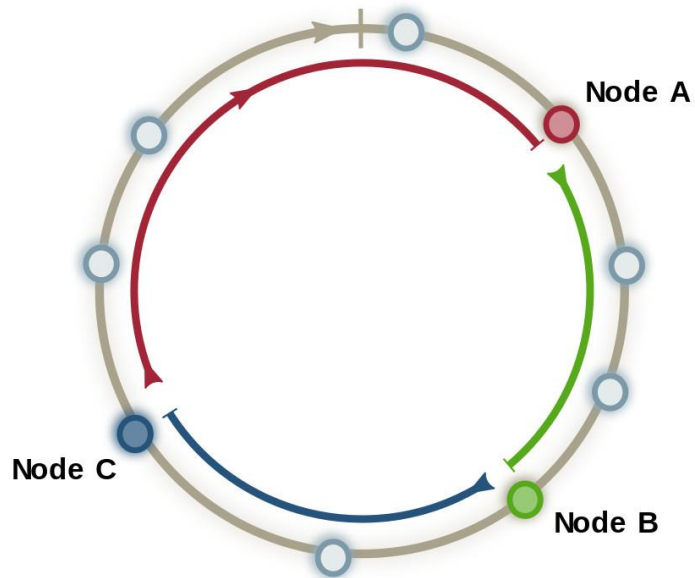
CouchDB

- Lounge
 - dumbproxy
basit requestler (get/put)
 - smartproxy
CouchDB requestleri
(mapping/reducing)

CouchDB Lounge

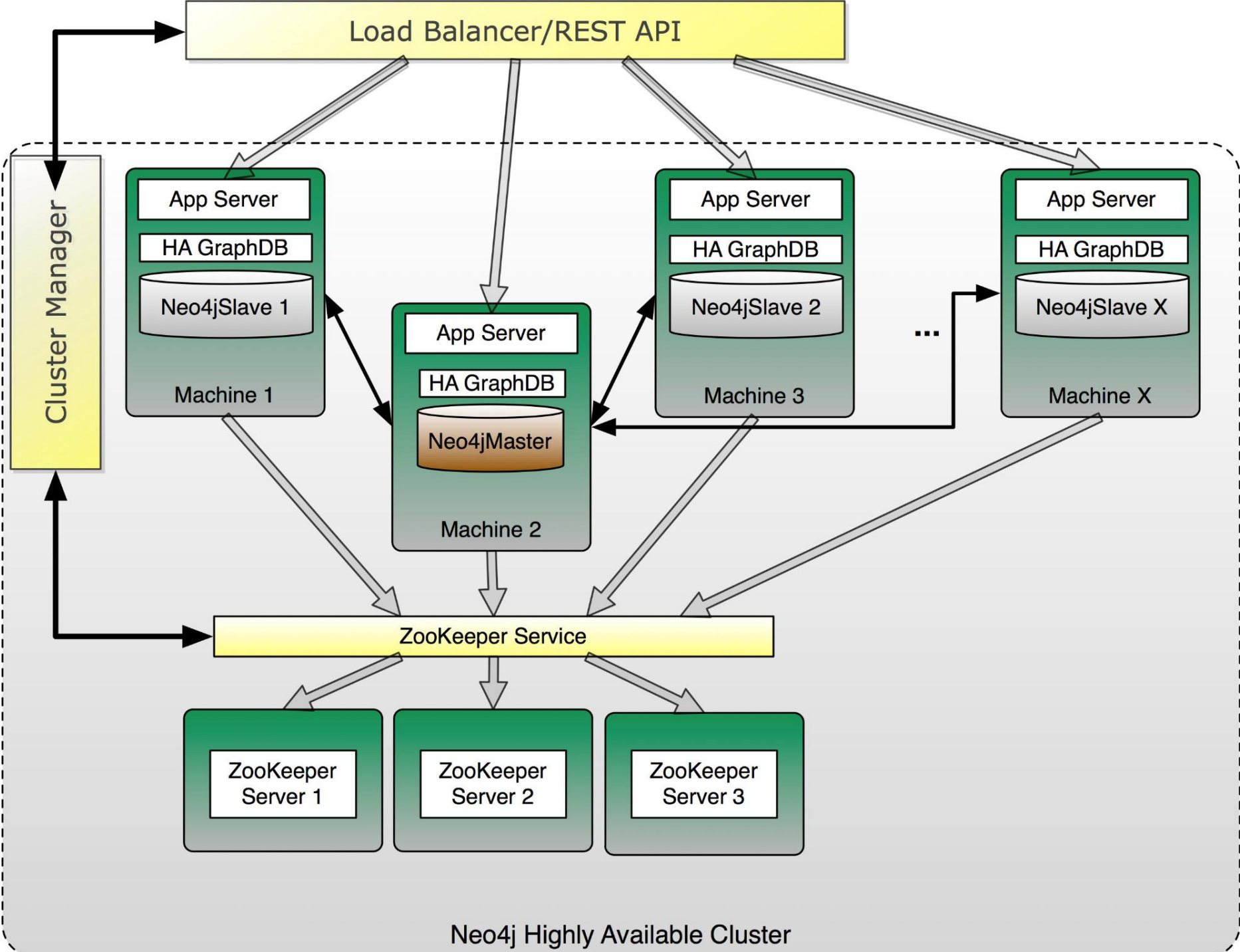
- Hashed DocID
Sharding için kullanılan key
- Keyspace
Her node üstünde hashed key için ayrılmış alan

Node Ring



Neo4j

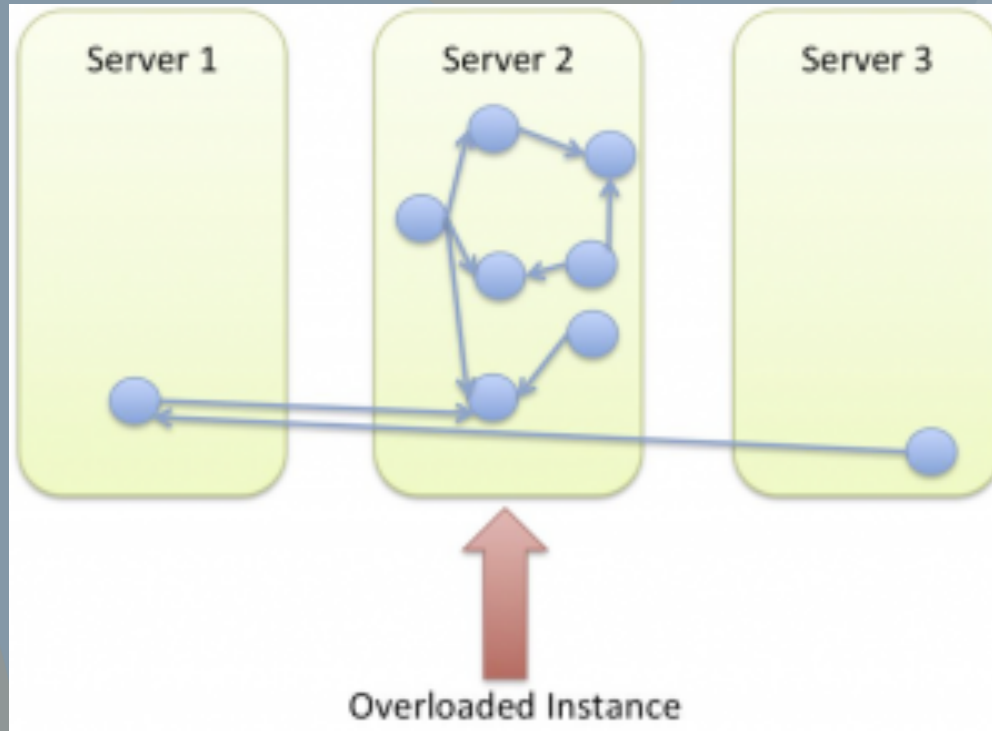
- Graph Veritabanlarında sharding kolay bir işlem değildir
- Neo4j High Availability
- Veri yerine işyükünü ölçeklendirme
- Apache ZooKeeper

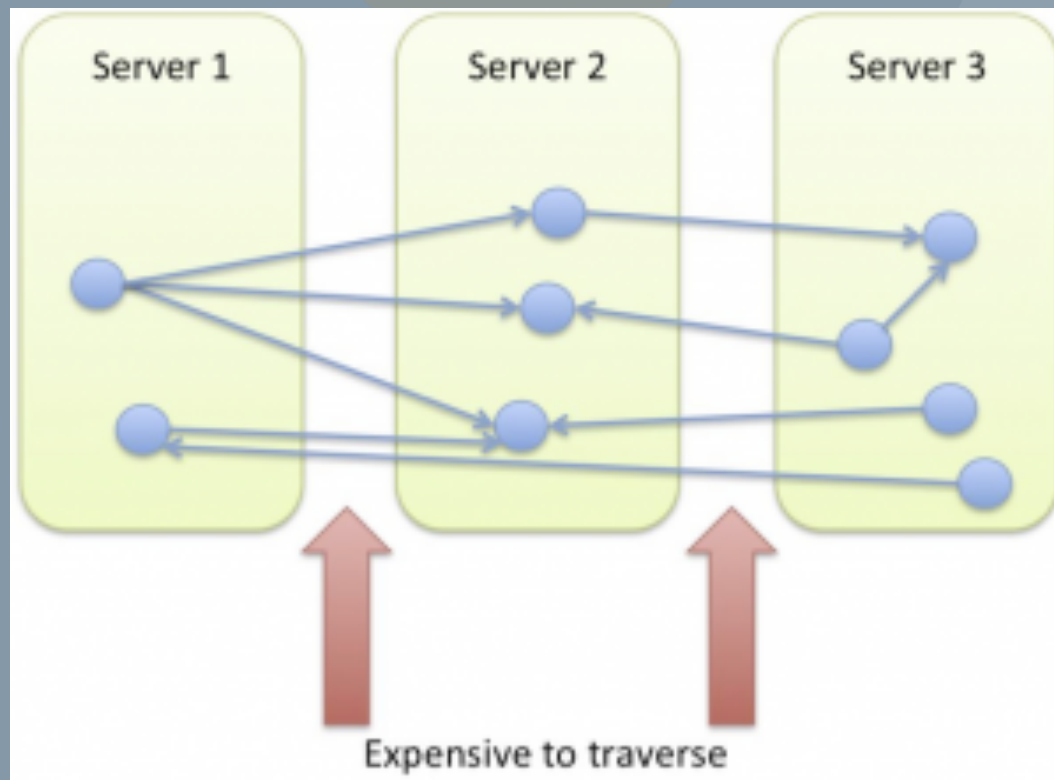


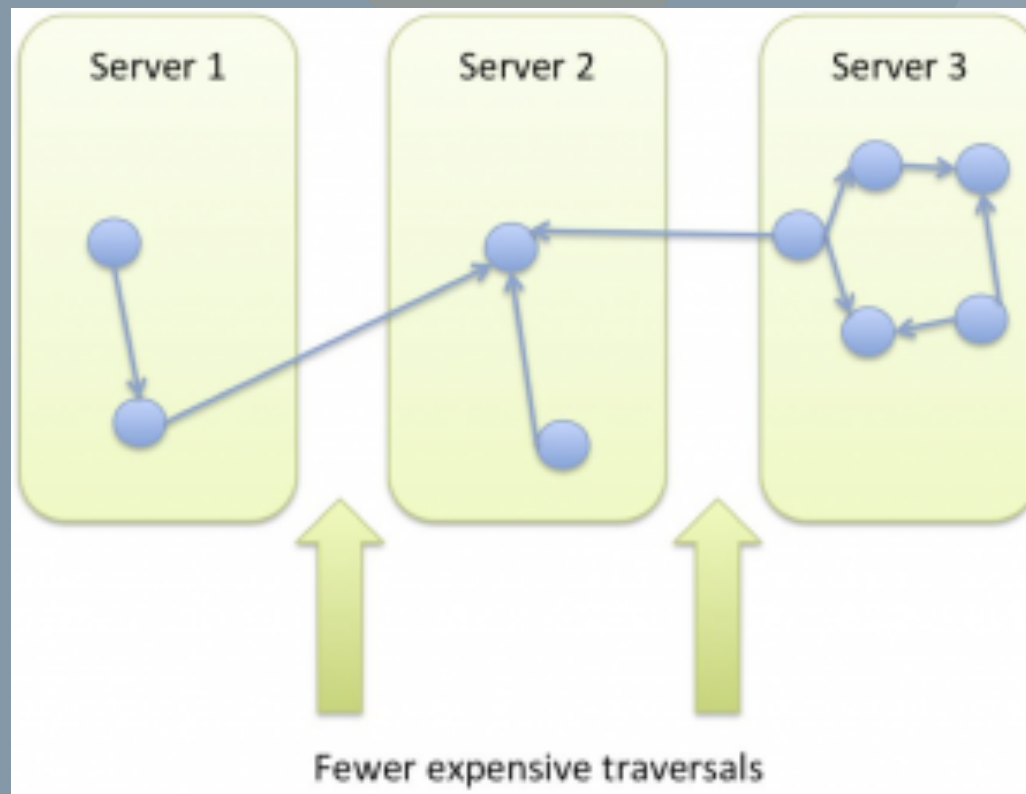
Sharding Neden Kolay Değil?

- Çok değişken bir yapıya sahip olması
- Dolaşım performansı vs. Fazla veri yüklemesi

Dolaşım Performansı



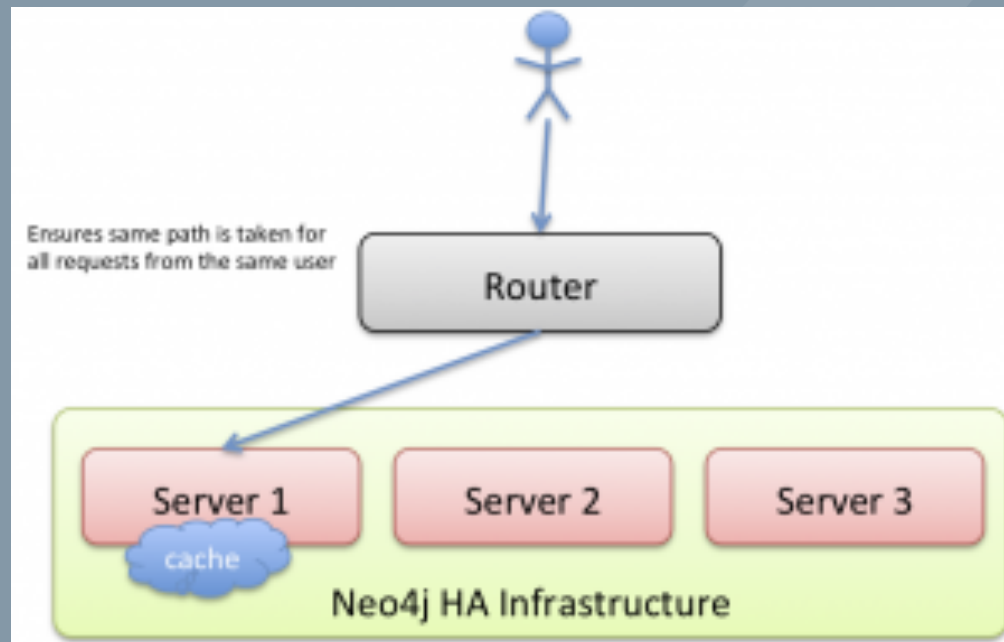




- Graphlar runtime da çok çabuk ve beklenmedik şekilde değişirler
- Insert time algorithm ve periyodik re-balancing

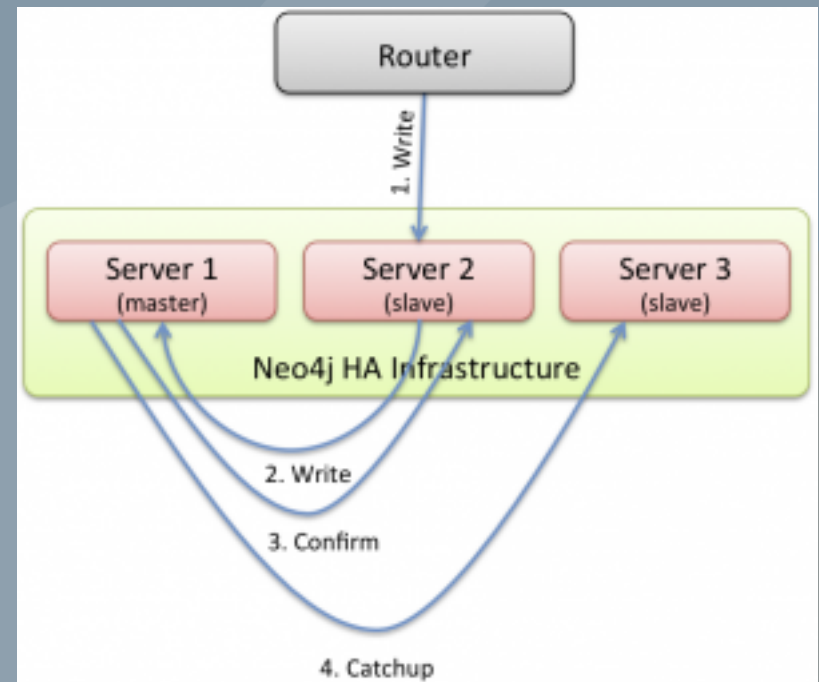
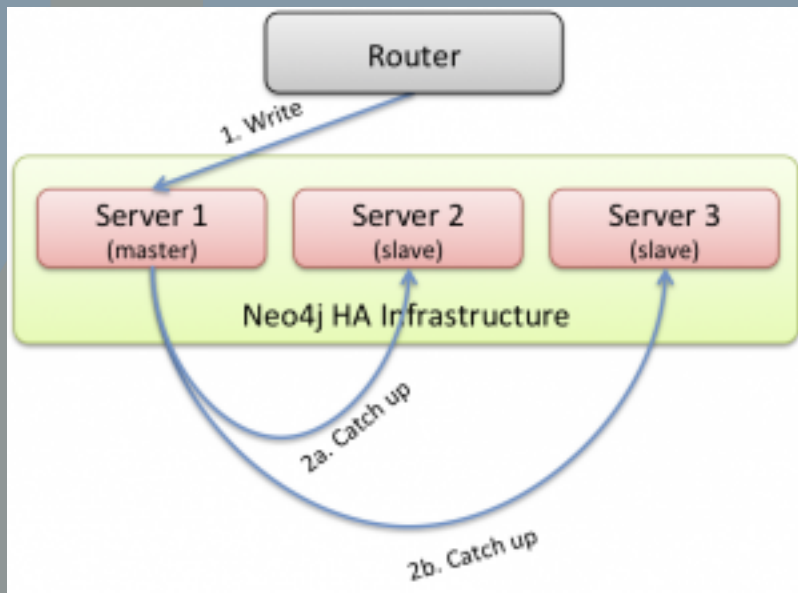
Cache Sharding

- Sharding olmayan sharding?
- İş yükünü ölçeklendirmek
- Warm Cache



İş Yükünü Ölçeklendirme

- Her sunucu aynı dataya sahiptir
- master + slave



MySQL

- MySQL Cluster kullanır
- MySQL Cluster = MySQL server + Ndb (Network Database) cluster

MySQL Cluster

- Hashed Primary Key
- Primary + Secondary Fragments

