Elastic Search

Seminar Presentation

Mohammed Nisham

July 15, 2016

College of Engineering, Trivandrum

Table of Contents

Introduction

What is elastic search

- Built on top of Apache Lucene
- Uses RESTful APIs
- Facebook, Github
- Ranked first in Search engine Data Bases

History

- Shay Banon
- Compass, library for Java
- Not Scalable
- Rebuilt in distributed approach using RESTful API

Features

Document Oriented

- JSON object
- Index
- Type - mapping
- Id
- All elements are optional
- Auto detection of field types

Sharding

- Distributed document store
- Complete Lucene search engine
- Hash functions for shard routing
- Immutability by segments
- Near real timing by In-memory Buffer
- Crash recovery with Translog

Sharding

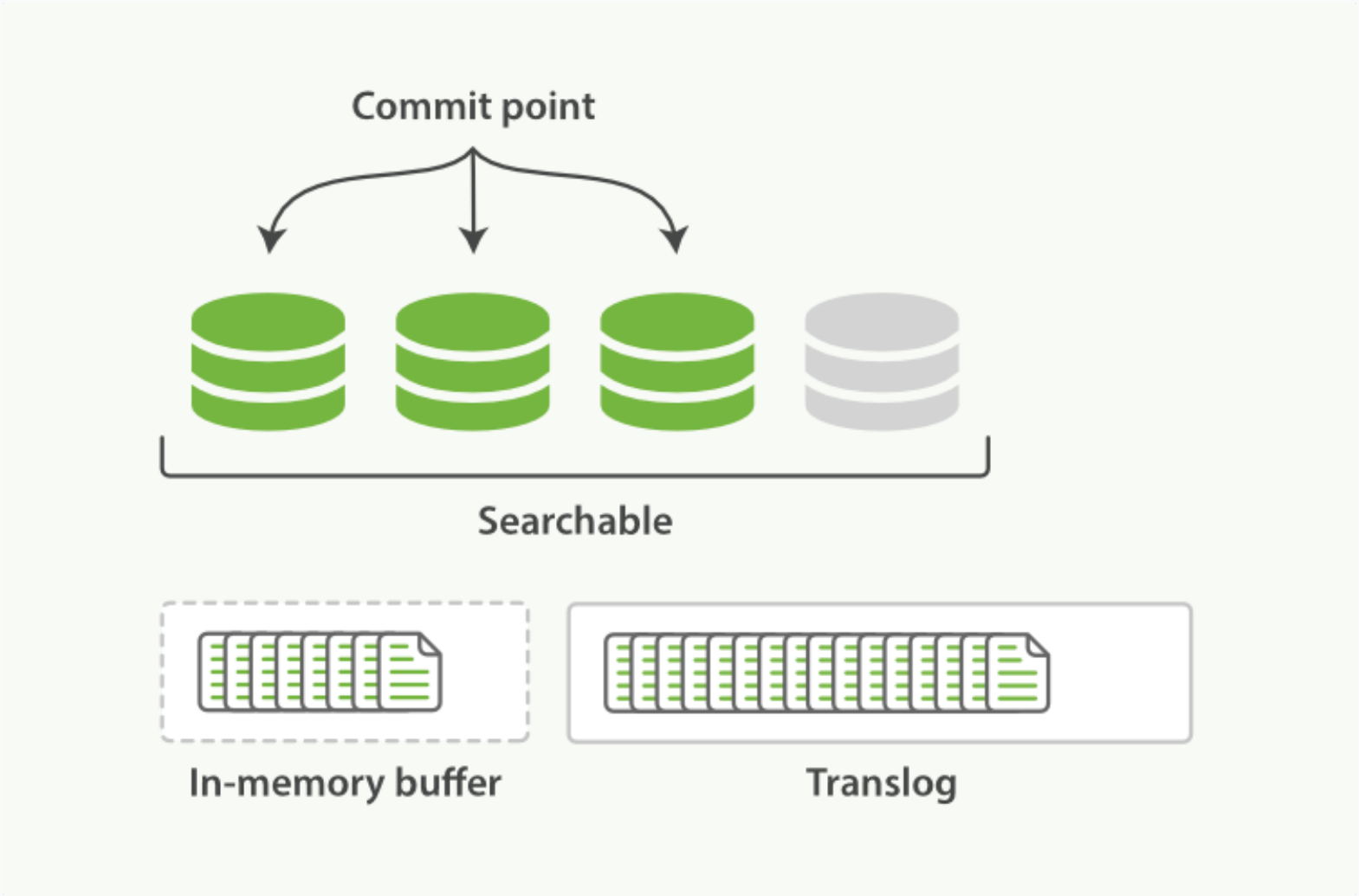

Figure 1: Inside a shard

Clustering

- Node
- Primary and Replica shards
- Failure recovery
- Horizontal scaling
- Cluster master
- Completely autonomous

Clustering

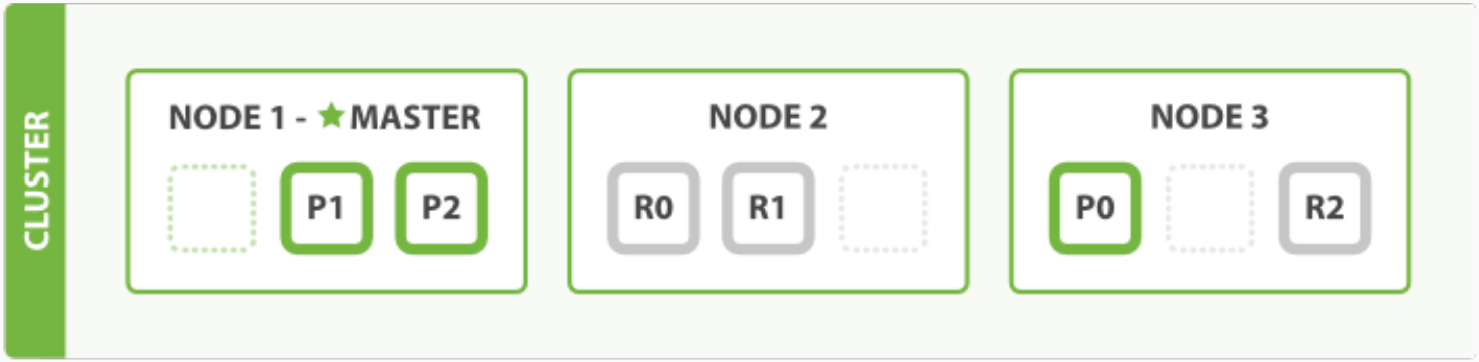

Figure 2: A sample cluster

Clustering

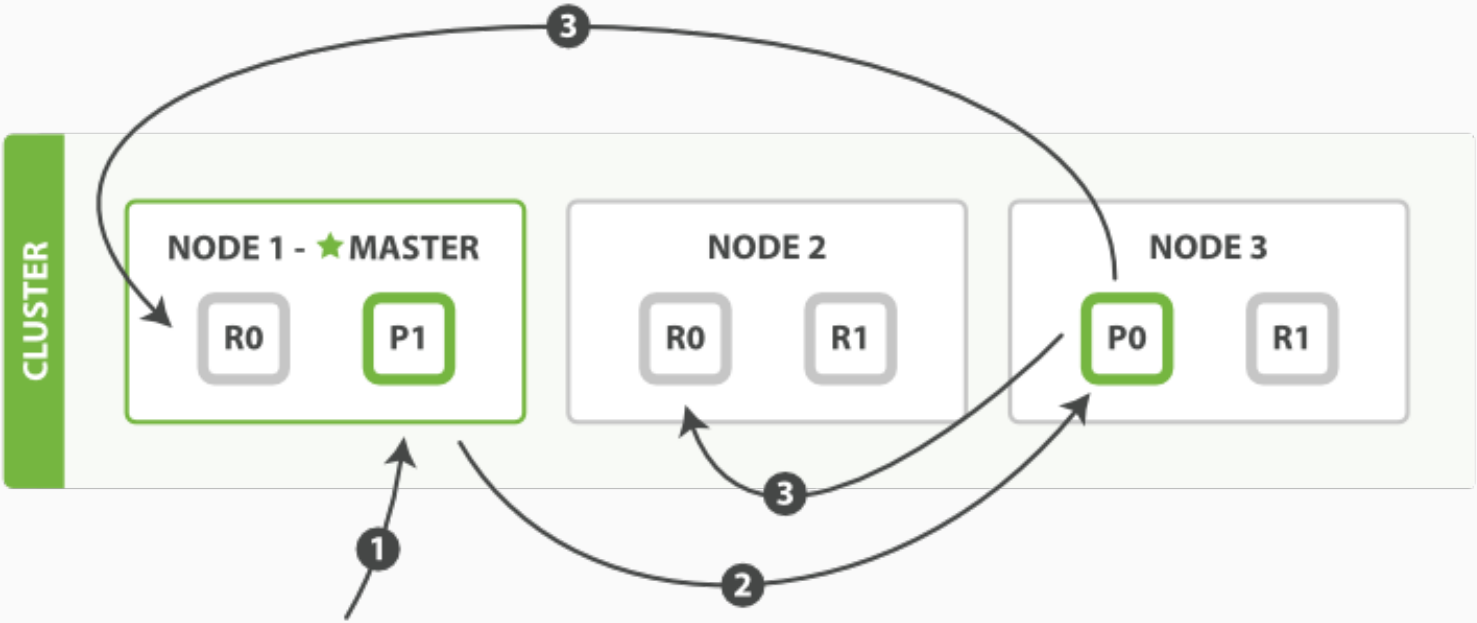

Figure 3: Request routing in a cluster

Concurrency

- Optimistic concurrency control
- Operations are asynchronous and concurrent
- Metafield version

Search

test

Conclusion

test

Questions?

References I


C. Gormley and Z. Tong.
Elasticsearch: The Definitive Guide.
O'Reilly Media.


P. Gupta and S. Nair.
Survey paper on elastic search.
International Journal of Science and Research (IJSR), January 2016.


https://www.elastic.co/products/elasticsearch.
Elasticsearch: Search and analyze data in real time.


O. Kononenko, O. Baysal, R. Holmes, and M. Godfrey.
Mining modern repositories with elasticsearch.
University of Waterloo, Waterloo, ON, Canada, 2014.