



ElasticSearch分布式特性

rockybean@yourdream

2015.02.10

Elastic

CAP

CAP

Consistency

Availability

Partition
Tolerance

Durability

Scalability

ES Quick Start



Node 0

`cluster.name:local`
`node.name:node0`

ES Quick Start

Node 0

Node 1

Node 2

New Nodes Join!

Zen Discovery

MultiCast

UniCast

Node Role

Master

Data

Client



Node 0

Term

Index

A Collection Of Documents

Shard

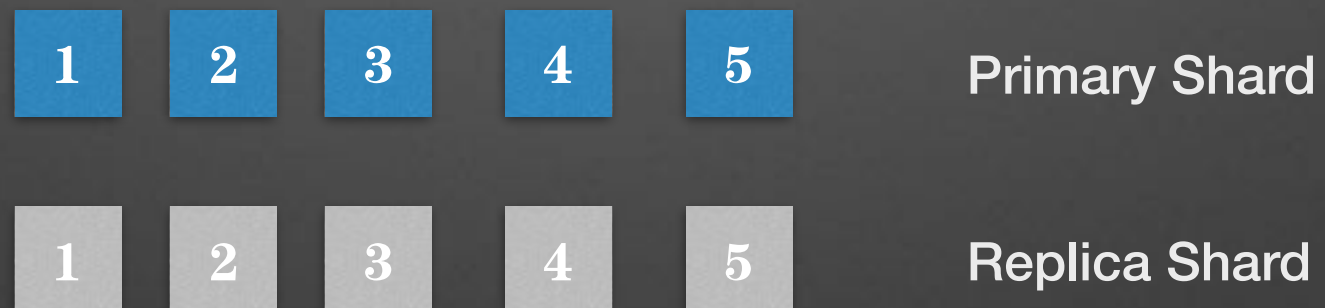
Unit of Scale

Shard

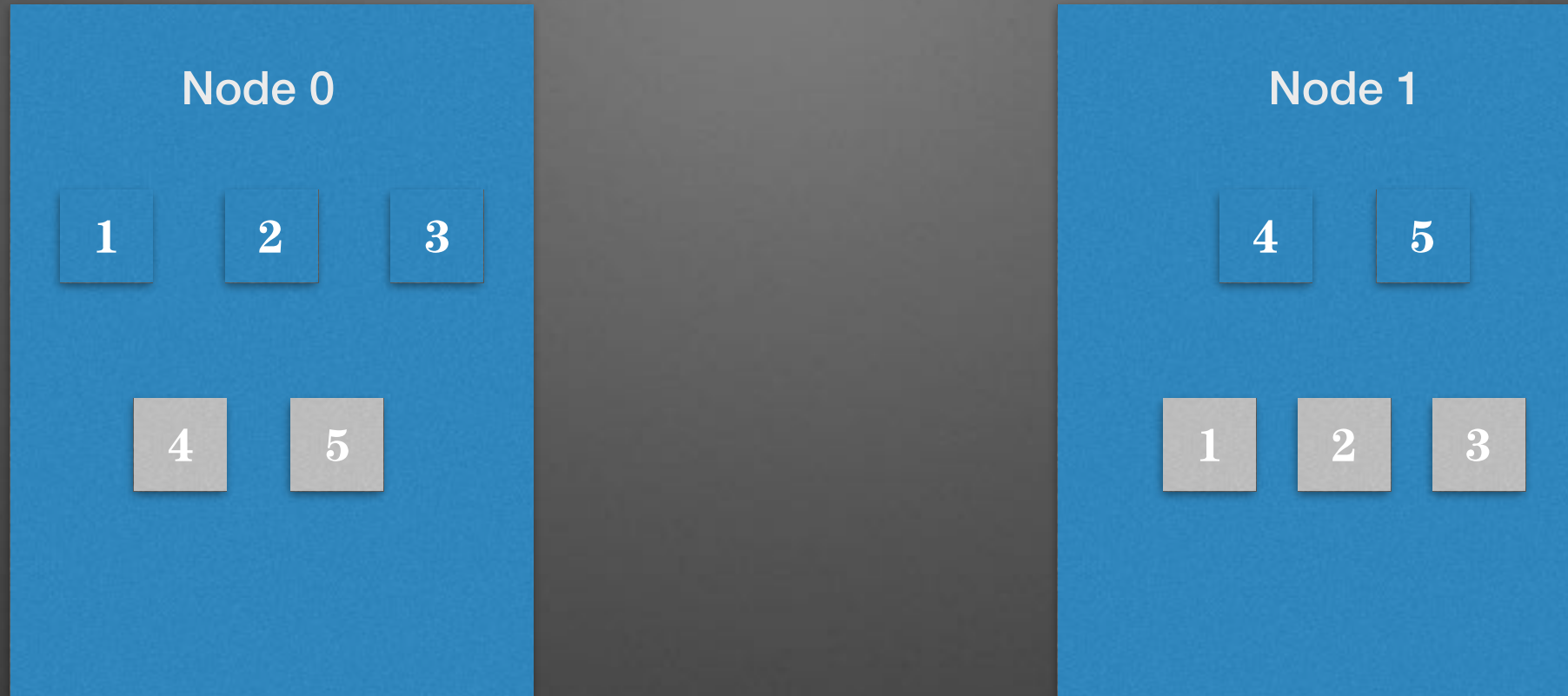
Index Setting

`number_of_shards 5`

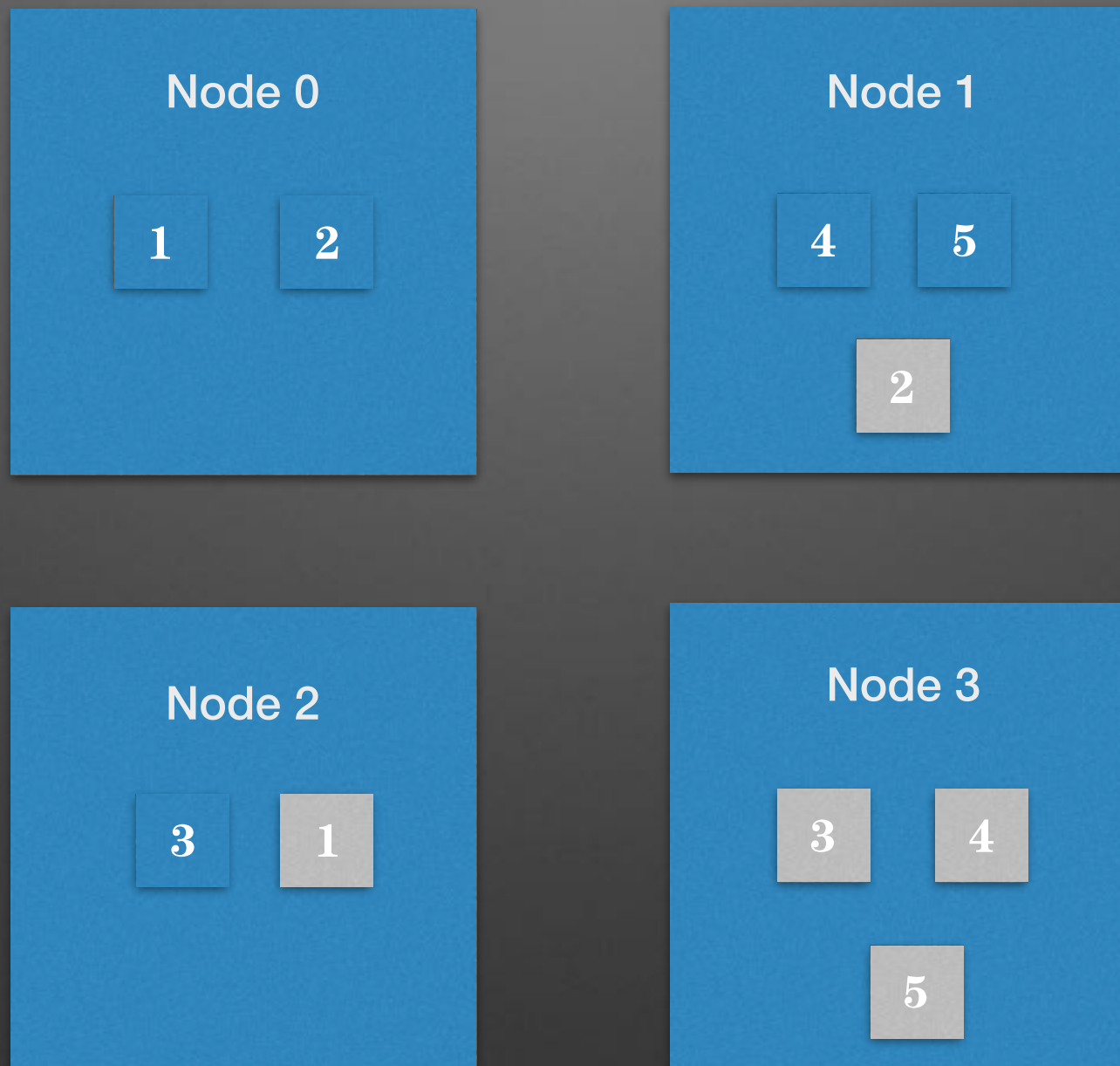
`number_of_replicas 1`



Shard Allocation Algorithm



Shard Allocation Algorithm



Shard Allocation Algorithm

`org.elasticsearch.cluster.routing.allocation allocator.EvenShardsCountAllocator`

Busy -> Not Busy

Based on Node , Not Index!

Write

Read

Write

doc



shard

$\text{shard} = \text{hash}(\text{routing}) \% \text{number_of_primary_shards}$

Write Durability

One

All

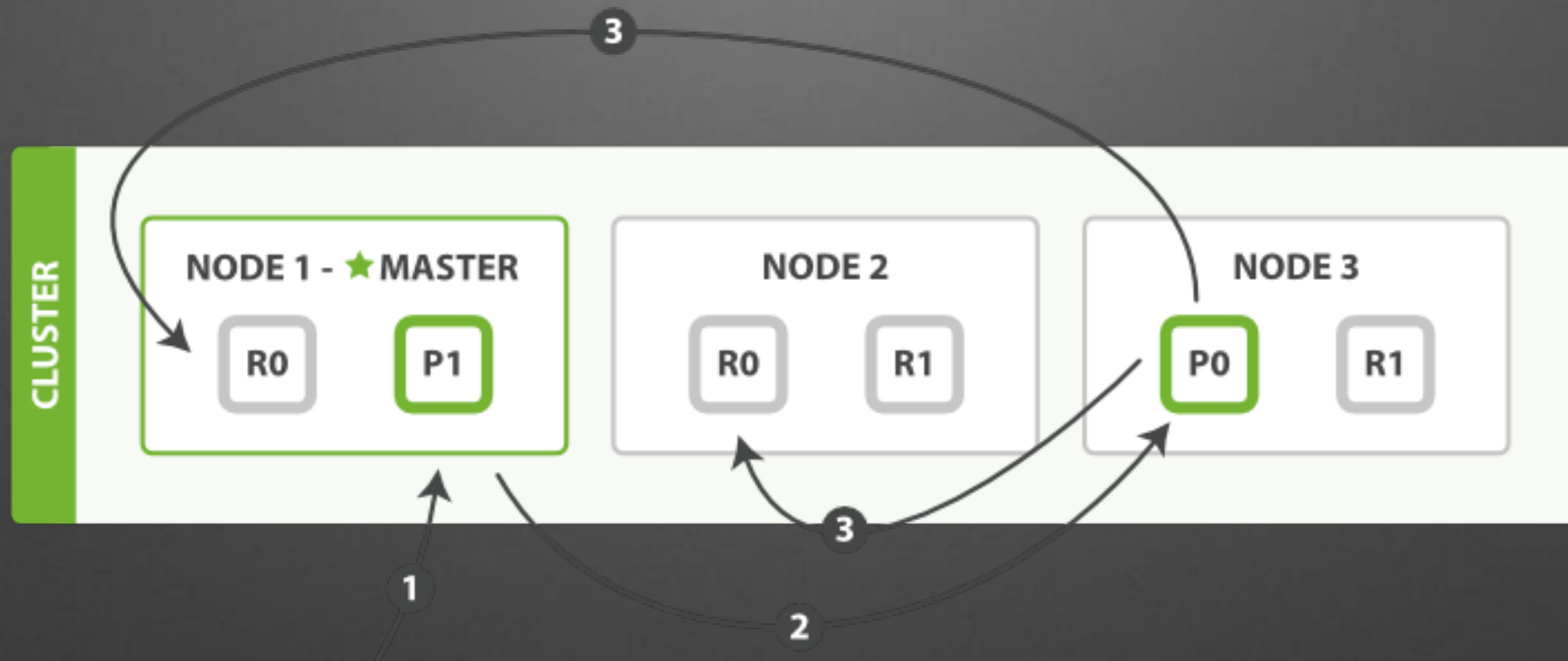
Quorum

$\text{int}((\text{primary} + \text{number_of_replicas}) / 2) + 1$

number_of_replicas 3

One Node Will Fail All Write Operations!

Write Durability



Write Durability

refresh

flush

transaction log

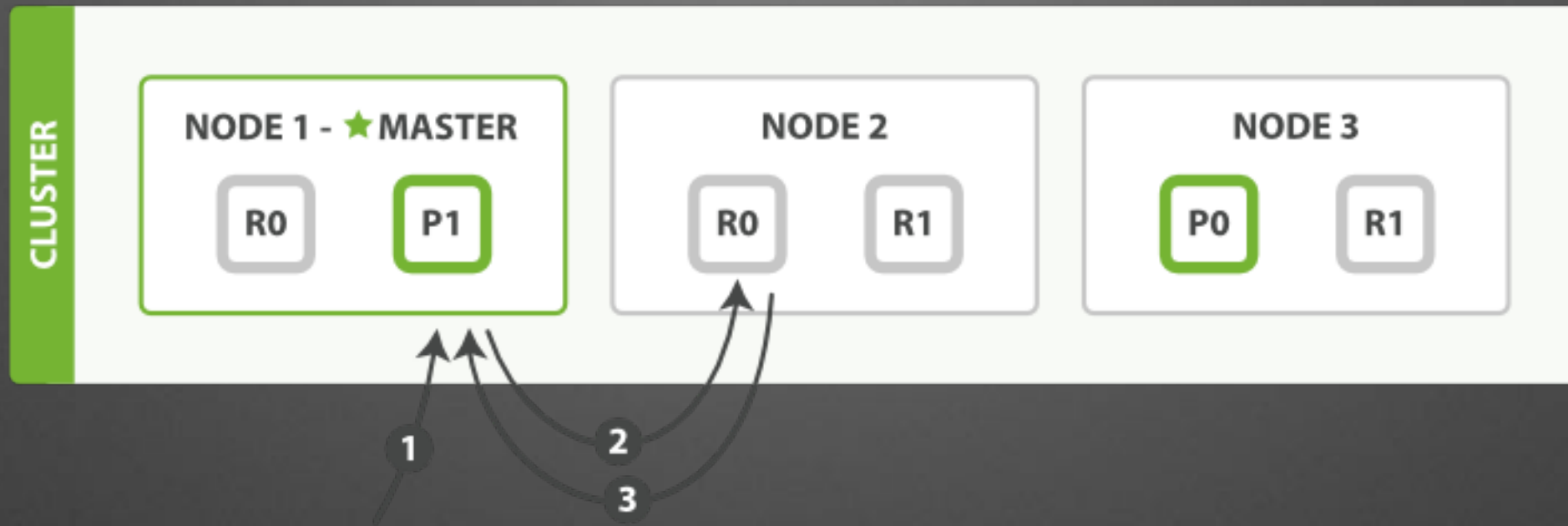
Update Based On Version

Read

Fetch One Document By Id

Fetch Documents By Conditions

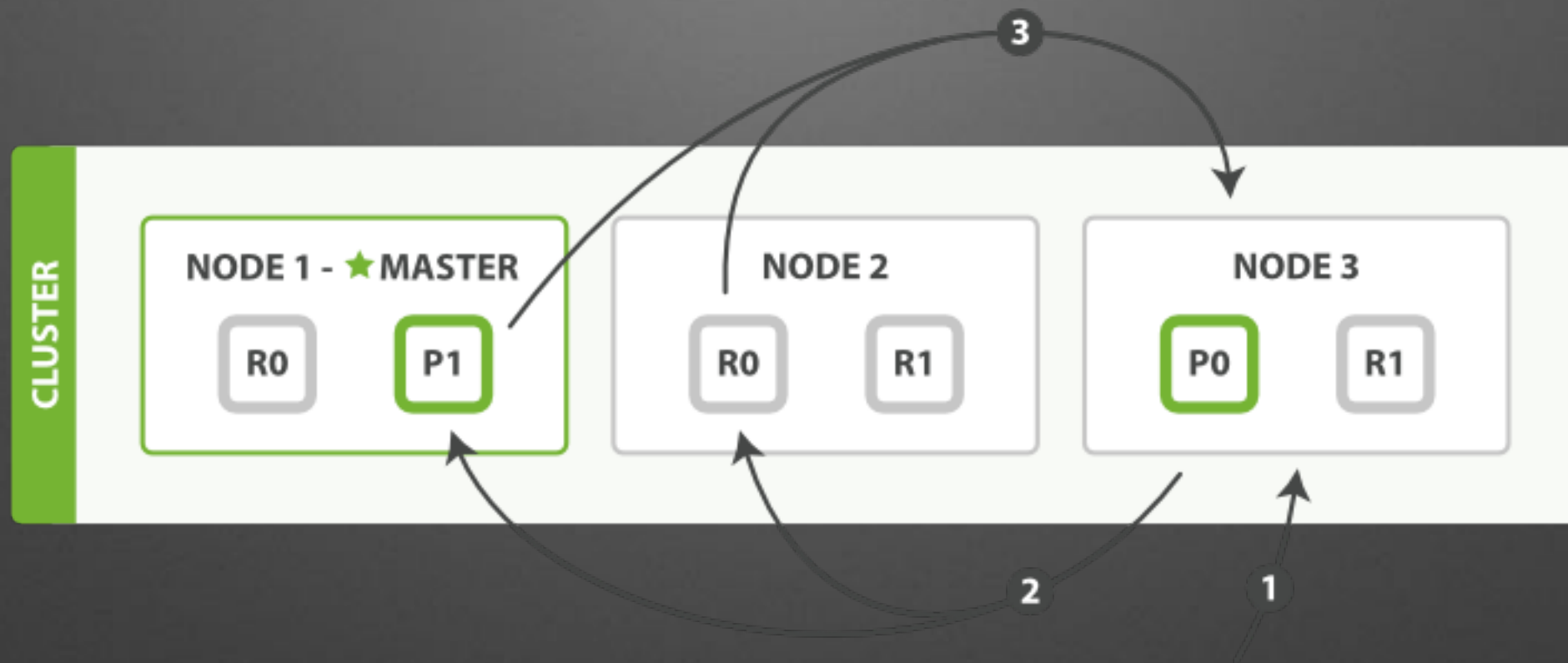
Fetch One Document By Id



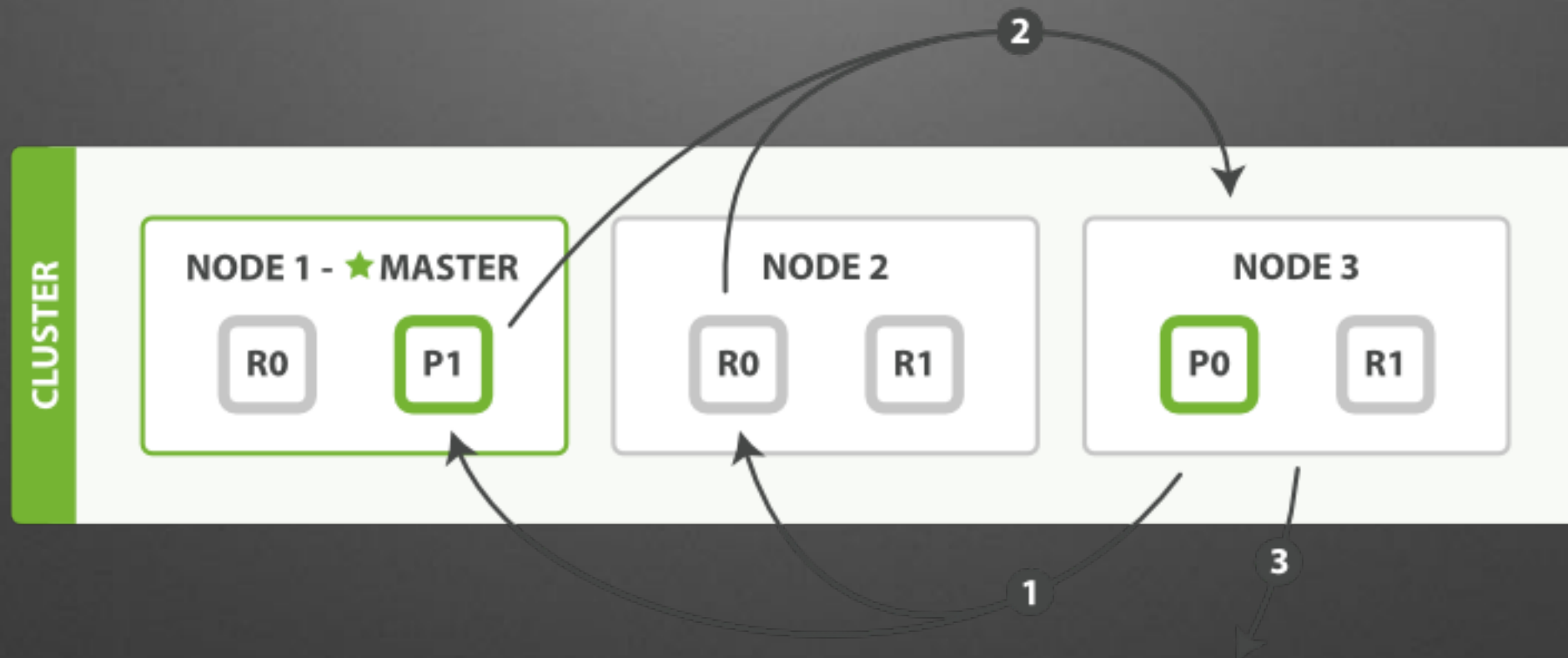
Paging

```
GET /_search
{
  "from": 90,
  "size": 10
}
```

Query Phase



Fetch Phase



Search Options

routing

GET /_search?routing=user_1,user2

search_type

query_then_fetch

count

query_and_fetch

scan

Sort

~~Inverted Index~~

~~Term -> Document~~

Load Field Values To Memory!

FieldData

Thanks!