



# Configuring a Cassandra node

DataStax Academy:

Installing and configuring Cassandra

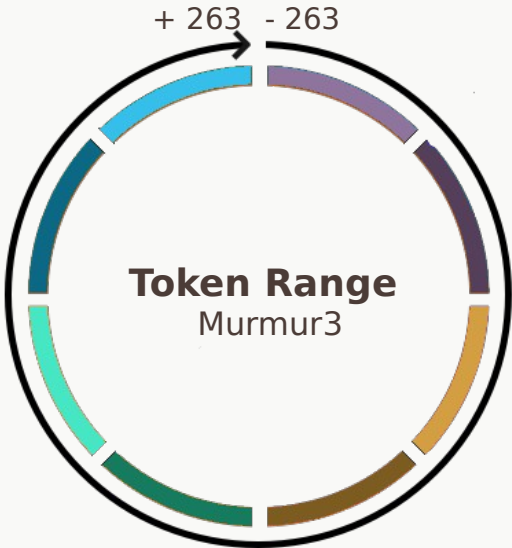
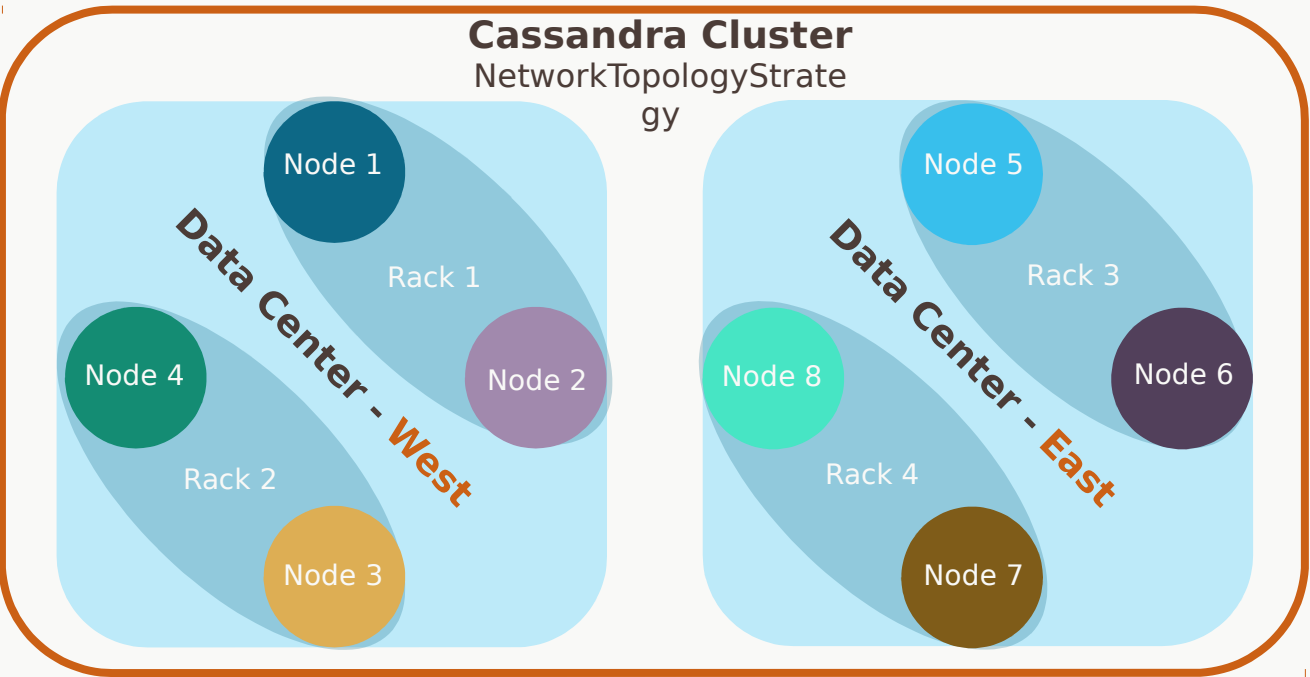
Leo Schuman

# Continue your Free Self-Paced Apache Cassandra Training at DataStax Academy

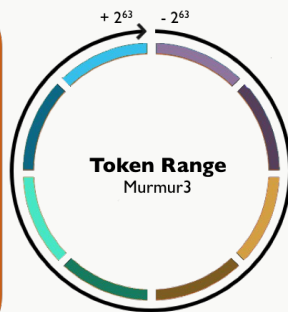
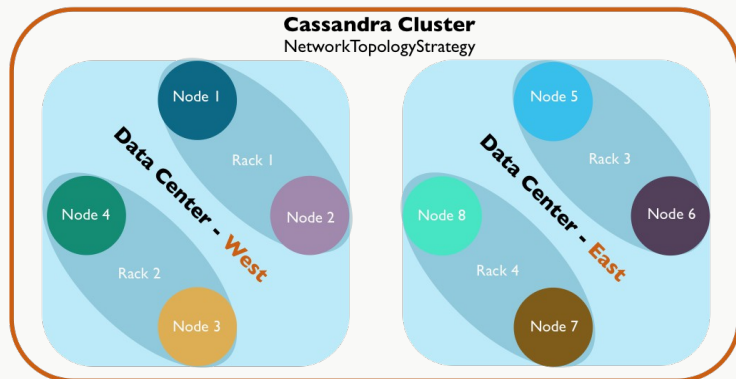
- Instructional Videos
- Learning Exercises
- Tutorials
- Demos

**<http://academy.datastax.com>**

# What are the main elements of a Cassandra cluster?



# What are the main elements of a Cassandra cluster?



- A peer to peer set of nodes with no single point of failure

- **Node** – a Cassandra instance
  - **Partition** – a uniquely identified data structure on a node
- **Rack** – a logical set of one or more nodes
- **Data Center** – a logical set of one or more racks
- **Cluster** – a logical set of one or more data centers

# What key properties are set in

## *cassandra.yaml*?

- **cluster\_name** (default: 'Test Cluster')
  - All *nodes* in a *cluster* must have the same value.
- **listen\_address** (default: localhost)
  - IP address or hostname other nodes use to connect to this *node*
- **commitlog\_directory** (default: /var/lib/cassandra/commitlog)
  - Best practice to mount on a separate disk in production (unless SSD)
- **data\_file\_directories** (default: /var/lib/cassandra/data)
  - Storage directory for data tables (*SSTables*)
- **saved\_caches\_directory** (default: /var/lib/cassandra/saved\_caches)
  - Storage directory for *key caches* and *row caches*
- **rpc\_address / rpc\_port** (default: localhost / 9160)
  - listen address / port for *Thrift* client connections
- **native\_transport\_port** (default: 9042)
  - listen address for *Native CQL Driver* binary protocol

# What key properties are set in *cassandra-env.sh*?

## • JVM Heap Size settings

- **MAX\_HEAP\_SIZE**="value"
  - Max recommended in production is 8G due to current limitations in Java garbage collection

| System Memory    | Heap Size                                |
|------------------|--|
| Less than 2GB    | 1/2 of system memory                     |
| 2GB to 4GB       | 1GB                                      |
| Greater than 4GB | 1/4 system memory, but not more than 8GB |

- **HEAP\_NEWSIZE**="value"
  - Generally set to  $\frac{1}{4}$  of MAX\_HEAP\_SIZE

## • More information

- **Documentation** ["Tuning Java resources"](#)
- **Training** ["Apache Cassandra: Operations and Performance Tuning"](#)

# What key properties are set in *logback.xml*?

- Cassandra system.log location
  - Default location is `/var/log/cassandra/system.log`
  - system.log is numerically renamed as it grows over time
- System.log configuration changed to logback.xml in Cassandra

```
<jmxConfigurator />
<appender name="FILE" class="ch.qos.logback.core.rolling.RollingFileAppender">
  <file>${cassandra.logdir}/system.log</file>
  <rollingPolicy class="ch.qos.logback.core.rolling.FixedWindowRollingPolicy">
    <fileNamePattern>${cassandra.logdir}/system.log.%i.zip</fileNamePattern>
    <minIndex>1</minIndex>
    <maxIndex>20</maxIndex>
  </rollingPolicy>

  <triggeringPolicy class="ch.qos.logback.core.rolling.SizeBasedTriggeringPolicy">
    <maxFileSize>20MB</maxFileSize>
  </triggeringPolicy>
```

- Cassandra logging level (most to least verbose)

# Demo: Configure a Cassandra node



# Continue your Free Self-Paced Apache Cassandra Training at DataStax Academy

- Instructional Videos
- Learning Exercises
- Tutorials
- Demos

**<http://academy.datastax.com>**