

# DataStax Developer Day

## DSE Operations / Security



# Installing DSE

@DataStaxAcademy #DataStaxDeveloperDay

The power behind the moment.



# Install



- Manual with tarball
- Manual with Linux package
- Automatic with Lifecycle Manager

# Configuring DSE

# Making configuration files FUN!!!



- Let's look at important configuration files
- Why they matter

# Important configuration files



cassandra.yaml



cassandra-rackdc.properties

# cassandra.yaml

- Cassandra nodes read this file on start-up
  - restart nodes for changes to take effect
- Located in the following directories
  - Cassandra package installations:  
/etc/dse/cassandra
  - Cassandra tarball installations:  
<INSTALL\_LOCATION>/resources/cassandra/conf
- `cluster_name` (Default: “Test Cluster”)
  - Once `cluster_name` is set it cannot be changed unless the cluster is rebuilt
- `listen_address` (Default: localhost)
- `native_transport_address` (Default: localhost)
- `seeds` (Default: “127.0.01”)
  - You must have at least one seed node per datacenter

# cassandra-rackdc.properties

- Cassandra nodes read this file on start-up
- Located in the following directories
  - Cassandra package installations:  
/etc/dse/cassandra
  - Cassandra tarball installations:  
<INSTALL\_LOCATION>/resources/cassandra/conf
- dc (Default: “dc1”)
  - Once dc is set it cannot be changed unless the DC is rebuilt
- rack (Default: “rack1”)
  - Once rack is set it cannot be changed unless the nodes are moved to a new rack

# Honorable mentions

- DSE
  - dse.yaml
  - If you are trying to use something not common to both cassandra and DSE then the dse.yaml is the first place to look.
  - Settings
    - Security
      - Auth
      - Encryption
    - DSE Search
    - Performance
    - DSE Analytics
    - Advanced replication
- Cassandra
  - cassandra-env.sh
  - jvm.options
- OpsCenter
  - <cluster\_name>\*.conf
  - Agent
    - address.yaml
- Studio
  - configuration.yaml

# Exercise

Open up DataStax Studio and the notebook:

**DSE Operations & Security: Installing DSE**

@DataStaxAcademy

#DataStaxDeveloperDay



- Install DSE and add a node to an existing DSE cluster
- Use LCM to install the DSE cluster from scratch

# Security

# Security overview

Bad security results in data breaches. DSE has a number of features that can be used to secure your data at all stages.

- End to end encryption
- Data auditing
- LDAP integration
- Kerberos integration
- Role based access control
- Row-level access control

# Encryption

Encryption keeps your data safe so that even if people get their hands on it they will not be able to use it.



# Auditing

Auditing lets you keep track of what is going on with your system and who has been accessing it.



# Auditing

@DataStaxAcademy

#DataStaxDeveloperDay

**AUTH** - Logs login events

**DML** - Logs insert, update, delete and other data manipulation language (DML) events

**DDL** - Logs object and role create, alter, drop, and other data definition language (DDL) events

**DCL** - Logs grant, revoke, create role, drop role, and list roles events

**QUERY** - Logs all queries

[https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secAuditTOC.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secAuditTOC.html)

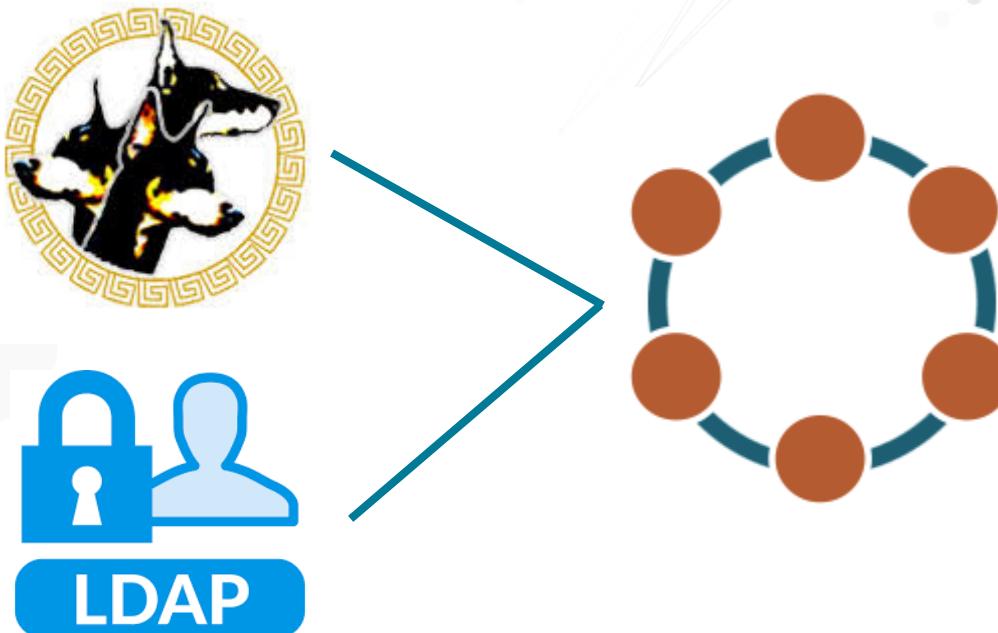
# Authentication and Authorization

What do these terms mean?

- Authentication controls who can access the cluster
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secDSEUnifiedAuthAbout.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secDSEUnifiedAuthAbout.html)
- Authorization controls what they can do on the cluster
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secRbacAbout.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secRbacAbout.html)

# LDAP & Kerberos

Account management is made easier when integrated with the existing systems you use.



# Exercise

Open up DataStax Studio and the notebook:

**DSE Operations & Security: Securing DSE**

@DataStaxAcademy

#DataStaxDeveloperDay



- Use the Lifecycle Manager to enable authentication and authorization
- Also use the Lifecycle Manager to enable audit logging
- Finally, use Lifecycle Manager to turn on in-flight encryption

# Monitoring and Performance

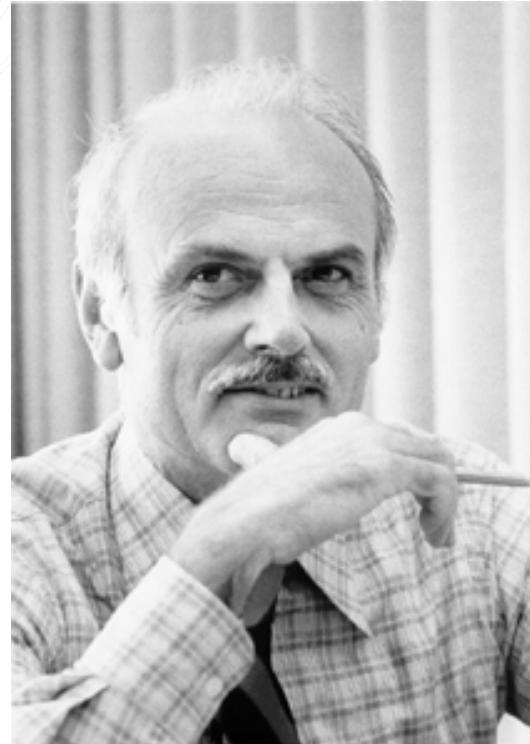
# Coming from a relational database?

- How you tune your database for performance can be very different from what you may be used to
  - Even scary, like really scary



# DSE Performance

- Storing data "correctly"
- Performance optimizations come afterwards
- We tune our queries using explain plans, things like that



# DSE Performance

## What do we do with DSE?

- Designed for OLTP operations
- Queries are meant to be fast!
  - No joins
  - Data is distributed across different nodes
  - Replicas mean that DSE has multiple nodes to read from
- Don't need explain plans

# DSE Performance

## How this gets complicated...

This is a distributed system!

- Data... data everywhere
- Workloads and resources on different instances

Read path and write path... know this and how read / write performance is affected by this

- Compaction
- Anti-entropy repair and node sync operations

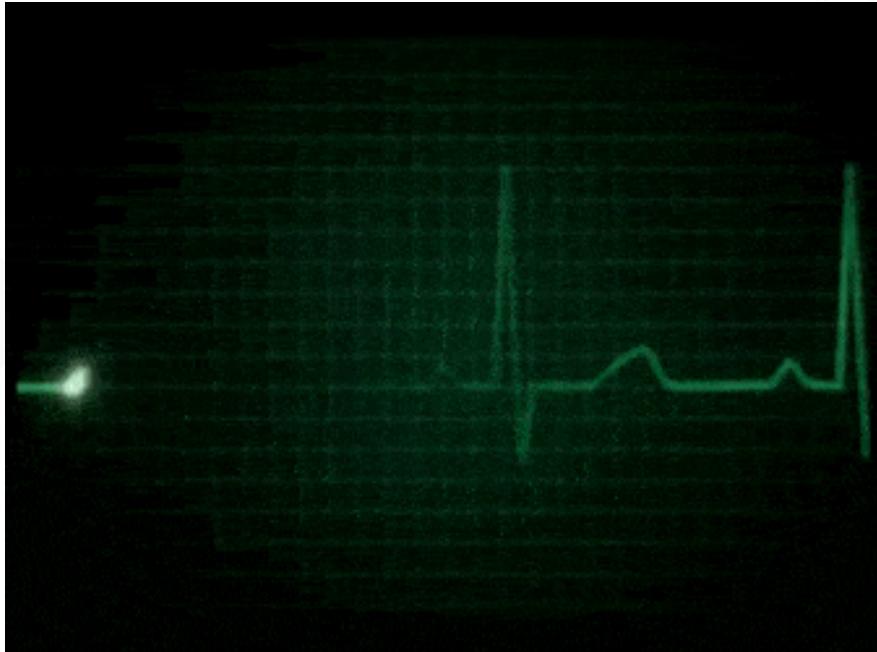
Data modeling

- **The irony - you're optimizing the data model now, not your queries**



# DSE Performance

What?



Monitoring!

Keyspace : killrvideo  
Read Count: 13516  
Read Latency: 0.09195841965078425 ms  
Write Count: 0  
Write Latency: NaN ms  
Pending Flushes: 0  
  
Table: comments\_by\_user  
SSTable count: 35  
Space used (total): 937.4  
Space used (total): 1443744  
Space used by snapshots (total): 0  
Off heap memory used (total): 5960  
SSTable Compression Ratio: 0.5464805748827252  
Number of partitions (estimate): 4510  
Memtable cell count: 0  
Memtable data size: 0  
Memtable off heap memory used: 0  
Memtable switch count: 3  
Local read count: 0  
Local read latency: NaN ms  
Local write count: 0  
Local write latency: NaN ms  
Pending flushes: 0  
Percent repaired: 0.0  
Bytes repaired: 0.000KiB  
Bytes unrepaired: 2.423MiB  
Bytes pending repair: 0.000KiB  
Bloom filter false positives: 0  
Bloom filter false ratio: 0.00000  
Bloom filter space used: 5656  
Bloom filter off heap memory used: 5648  
Index summary off heap memory used: 0  
Compression metadata off heap memory used: 312  
Compacted partition minimum bytes: 73  
Compacted partition maximum bytes: 3973  
Compacted partition mean bytes: 617  
Average live cells per slice (last five minutes): NaN  
Maximum live cells per slice (last five minutes): 0  
Average tombstones per slice (last five minutes): NaN  
Maximum tombstones per slice (last five minutes): 0  
Dropped Mutations: 0  
Failed Replication Count: null

# DSE Performance

## The Open Source way

- Low level stuff
  - JMX Mbeans - lots of metrics
  - Nodetool commands
    - Might be running scripts to capture, collect, and aggregate output
    - Cron jobs

# DSE Performance Service

Queries and statistics collected and saved in tables

- DSE will monitor your queries
- Slow queries? Will let you know what's going on
- Just pull out of the database and use wherever you want

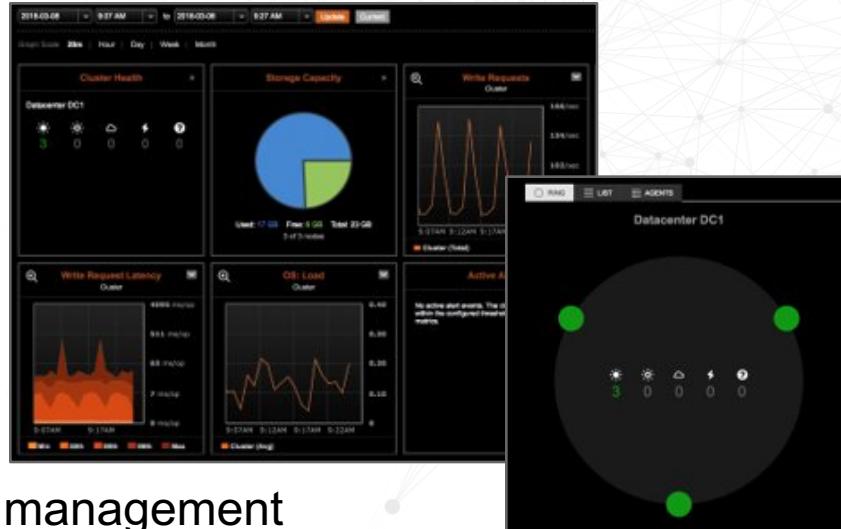


# DSE Performance Service

- What else...
  - CQL Slow Query Log
  - Latency metrics on tables
  - Per node and cluster lifetime metrics by table and keyspace
  - Recent and lifetime statistics about tables, sstable count, partition size, etc
  - Read / write activity by per-client and per-node
  - DSE Search bottlenecks
  - DSE Analytics resource usage
  - Monitor specific DSE Analytics applications

# OpsCenter overview

- Browser-based DSE cluster tool for:
  - Configuring
  - Monitoring
  - Managing
- Two major components tied together:
  - OpsCenter Monitoring - monitoring and management
  - Life Cycle Manager (LCM) - mostly configuration and deployment



The screenshot shows the 'Nodes' tab of the LCM interface for 'TestCluster'. It displays node health (3 Normal, 0 Medium, 0 High, 0 Down, 0 Unknown), data size (Total Size: 420 kB, Avg. Node: 140 kB, +/- 11 kB), repair status (Not running), backup status (No Scheduled Backups, Recent: 0 successful, 0 failed), performance metrics (Table Metrics, Thread Pool Stats, Slow Query Log), and best practices (Rules passed: 0, Rules failed: 0).

# DSE Performance

**Metrics on this Graph**

Metric	Nodes	Tables
--------	-------	--------

Add Metric

**Save Graph**   **Cancel**

**Add Metric**

Metric: Click to choose a metric...

Node: Click to choose one or more nodes...

Table: Click to choose one or more tables...

**Add Metric**   **Cancel**

Click to choose a metric...

Datacenter Messaging Latency  
Node Messaging Latency  
Write Request Latency (percentiles)  
Write Requests  
Read Failures  
Read Request Latency (percentiles)  
Read Requests  
Read Timeouts  
Read Unavailable Errors

- We collect metrics. Query metrics? Got it. JVM metrics. Got it. OS metrics? You know it!
  - Can divide by datacenter, node, table, etc

# DSE Performance

## OpsCenter Services

- Want to know when your nodes are slowing down? We'll send you a message
- Best practices service, performance service, lots of stuff
- Even capacity planning

Developer Day Cluster: Services				
	SERVICE	DESCRIPTION	STATUS	
	Backup Service	Schedule and restore backups (local or cloud). Also supports commit log backup for point in time restore.	SCHEDULED: 0	<a href="#">Details</a>
	NodeSync Service	NodeSync is a data validation and syncing service managed by DataStax Enterprise. NodeSync must be configured in DSE. NodeSync status can be viewed in OpsCenter after enabling keyspaces and tables.	ON	<a href="#">Details</a>
	Repair Service	Manage automatic repairs of all clusters; keeps data consistent on all nodes without impacting performance.	OFF	<a href="#">Configure</a>
	Best Practice Service	Automatically scans clusters using a set of pre-defined best practice rules and provides expert advice on resolving issues.	PASS: 12 FAIL: 0	<a href="#">Details</a>
	Capacity Service	Automatically collects statistical data for trend analysis and forecasts future trends.	ON	
	Performance Service	Combines OpsCenter metrics with CQL-based diagnostic tables populated by the DSE Performance Service to help understand, tune, and optimize cluster performance. Visually enable the performance objects and analyze the results within OpsCenter. Performance Service is available in DataStax Enterprise version 4.6 and higher.	ON	<a href="#">Details</a>



2018-08-22 11:13 AM to 2018-08-22 11:33 AM

Update

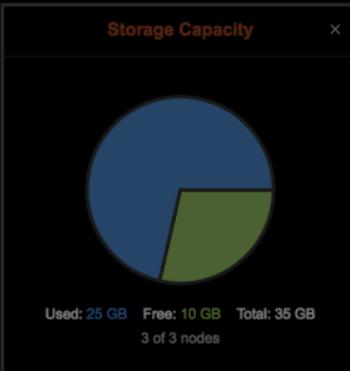
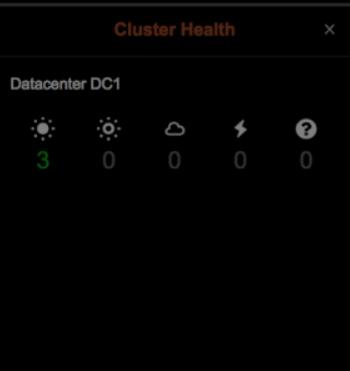
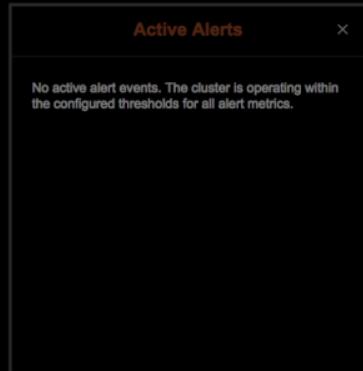
Current



Default

Graph Scale 20m | Hour | Day | Week | Month

Add Graph



# DEMO fun times!

# Wrapping Up

# Useful resources

- Security checklist and features
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secChecklists.html#secChecklists](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secChecklists.html#secChecklists)
- Encryption
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/encryptClientNodeSSL.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/encryptClientNodeSSL.html)
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secEncryptEnable.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secEncryptEnable.html)
- Auditing
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secAuditTOC.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secAuditTOC.html)
- Authentication
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secAuthTOC.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secAuthTOC.html)
- Role based access control
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secRbacAbout.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secRbacAbout.html)
- Row-level access control
  - [https://docs.datastax.com/en/dse/6.0/dse-admin/datastax\\_enterprise/security/secRlac.html](https://docs.datastax.com/en/dse/6.0/dse-admin/datastax_enterprise/security/secRlac.html)

# Continue Your Learning at DataStax Academy

- [DS210: DataStax Enterprise 6 Operations with Apache Cassandra™](https://academy.datastax.com)
  - Free self-paced course
  - Cluster sizing
  - Node management
  - Repair and nodesync
  - Compaction
  - Data loading

<https://academy.datastax.com>



Thank You

