

Instaclustr

Instaclustr's Open Source Tools for Apache Cassandra

**Adam Zegelin, Co-founder @ Instaclustr
adam@instaclustr.com**

©Instaclustr Pty Limited, 2019

`/usr/bin/
whoami`

Instaclustr

- **Managed Cassandra, Spark, Kafka & Elasticsearch in the cloud**
- **AWS, Azure, GCP & SoftLayer**
- **24x7x365 support**
- **Provide support for private DC/on-prem**
- **Manage and support ~3k+ nodes.**

Agenda

- **Instarepair & cassandra-sstable-tools**
- **Cassandra Kerberos Plugin**
- **Cassandra LDAP Plugin**
- **Backup & Restore Tool**
- **Cassandra Exporter for Prometheus**
- **Cassandra Operator for Kubernetes**

Instarepair

- Repairs for when repairs don't work
- Leverages read repair
- Pause and resume
- Throttle

- <https://github.com/instaclustr/instarepair/>

cassandra-sstable-tools

- Repairs for when repairs don't work
 - Leverages read repair
 - Pause and resume
 - Throttle
-
- <https://github.com/instaclustr/instarepair/>

Cassandra Kerberos Authenticator

- Application single sign-on for Cassandra CQL users.
- Cassandra plugin & Client driver extensions.
- GSSAPI SASL mechanism & JAAS
- Supports transit protection/Kerberos QoP
 - TLS not required for auth
- Open-source (v.s. DSE)
- Cassandra 3.0.x and 3.11.x

Usage

- **Unique DNS + reverse DNS for each node.**
- **Kerberos 5 KDC server.**
- **Kerberos client libraries installed on each node.**
- **NTP configured on nodes and KDC.**
- **JCE Unlimited Strength Policy files installed (Oracle JVM only).**
- **Kerberos service principals and keytabs for each node.**
- **Install JAR into C* classpath (`$CASSANDRA_HOME/lib`)**
- **Set authenticator in `cassandra.yaml`:**
 - `authenticator: com.instaclustr.cassandra.auth.KerberosAuthenticator`
- **Switch clients to use `KerberosAuthProvider`**

Cassandra Kerberos Authenticator (cont.)

- **Authenticator:**

<https://github.com/instacluster/cassandra-java-driver-kerberos/>

- **Java Driver extension:**

<https://github.com/instacluster/cassandra-java-driver-kerberos>

- **Introduction blog post:**

<https://www.instacluster.com/kerberos-authenticator-for-apachecassandra/>

Cassandra LDAP Authenticator

- **Verify usernames/passwords against LDAP/AD**
 - Centralized user management.
- **Username/password sent in plain text to C***
 - Enable client → node TLS.
- **Credentials cached (using C* built-in cache)**
 - Prevents LDAP server overload.
- **Supports Cassandra 2.2, 3.0.x, 3.11.x**

Usage

- **Checkout from Git and compile (no releases yet, sorry!)**
- **Place built JAR into C* classpath (\$CASSANDRA_HOME/lib)**
- **Modify ldap.properties, specifying:**
 - LDAP server address
 - Service account details
- **Set authenticator in cassandra.yaml:**
 - authenticator: `com.instaclustr.cassandra.ldap.LDAPAuthenticator`
- **No application code changes necessary**

Cassandra LDAP Authenticator

- **Authenticator**

<https://github.com/instaclustr/Cassandra-ldap>

- **Introduction blog post:**

<https://www.instaclustr.com/apache-cassandra-ldap-authentication/>

Cassandra Backup & Restore Tool

- **Backup and restore**
- **Take snapshots**
- **Uploads snapshots to cloud storage or remote file systems**
 - AWS S3, Azure Blob Storage, GCP Cloud Storage
- **Transfer throttling**
- **Schema and token range backups**
- **Automatic “de-duplication”**
 - SSTables that already exist in the destination are skipped.

Cassandra Backup & Restore Tool (cont.)

- **Used by**
 - Instaclustr managed service
 - Many customers
 - Instaclustr Cassandra K8s operator

Cassandra Backup & Restore Tool (cont.)

- Get it here:

<https://github.com/instaclustr/cassandra-backup>

Cassandra Prometheus Exporter

- **Exporter of Cassandra metrics**
 - Aka, a scrape target for Prometheus

Prometheus

- **Systems monitoring and alerting toolkit**
- **Multi-dimensional data model + PromQL = powerful queries**
- **Pull-based collection over HTTP**

- **Time series metrics database**
- **CNCF project**
- **De facto standard for K8s monitoring**

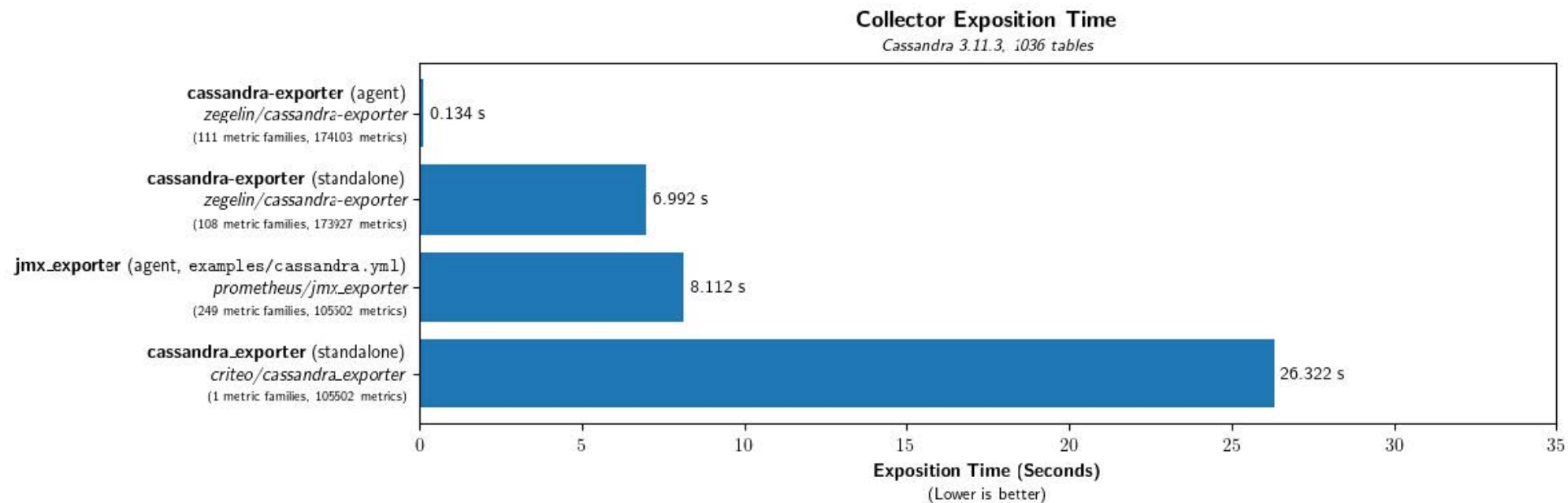
Cassandra Prometheus Exporter (cont.)

- *FAST*
- **Runs in-process (JVM Agent)**
- **Bypasses JMX**
 - JMX is slow
 - Really slow and burns CPU
- **Easy to use**

Existing Exporters

- **Generic**
 - don't know about Cassandra specific metrics
 - EstimatedHistograms, etc.
- **Not everything is available as a Dropwizard metric/Metric Mbean**
 - Failure detector stats (endpoint PHI, etc.)
 - Gossip stats
- **Don't follow Prometheus best practices**
 - Produce "garbage"
 - No labels, all metrics under one family, etc.
 - Hard to query

Did I mention fast?



Cassandra Prometheus Exporter (cont.)

- **Schema size is no longer a concern.**
- **Clean metrics/hand-tuned output.**
 - Easy to query & work with.
 - Maps JMX attributes → labels + custom labels.
 - i.e., table_type (index, view, etc.) – only query stats for views!
 - Groups appropriate metrics into families.
 - Scales/transforms values where appropriate.
 - Prometheus has common units (seconds, bytes, etc.)
 - Maps metrics to appropriate Prometheus types (counters, meters, etc.)
- **Handles Cassandra's histograms (inc. EstimatedHistograms), timers, etc.**
- **Prometheus text + custom JSON format supported.**
 - Easy (enough) to add more.

Cassandra Prometheus Exporter (cont.)

- **Metrics collected using Java Streams.**
 - Allows parallel collection.
 - Lower memory usage.
- **Light-weight Netty HTTP server.**
 - Originally Jersey (JAX-RS) was used, but too much overhead.
 - Custom Netty handlers that stream metrics (chunked) to socket.
- **Optimized text exposition format stream writer:**
 - Labels stored pre-formatted and UTF-8 encoded.
 - Slowest part is Float → text conversion!

Cassandra Version Compatability

2.x	Not yet! Pull requests appreciated
3.x	Yes
3.11.x	Yes
4.0	Work started

Usage

1. On each node:

a. Download latest release JAR and install into `$CASSANDRA_HOME/lib`

b. Edit `cassandra-env.sh`, and append:

```
JVM_OPTS="$JVM_OPTS -javaagent:$CASSANDRA_HOME/lib/cassandraexporter-agent-<version>.jar"
```

c. (Re)start Cassandra

2. On Prometheus host:

a. Add Cassandra nodes as scrape targets to `prometheus.yml`

3. Query, graph, and fascinate the executives (and monitor your cluster)

Cassandra Prometheus Exporter

- Get it here:

<https://github.com/instaclustr/cassandra-exporter>

Cassandra Operator for Kubernetes

- It's an Operator for running Cassandra on K8s!

HUH?!!

Adam's Crash Course on Containers and K8s

Containers

- **Not virtual machines.**
 - Sometimes called “lightweight VMs”.
- **Processes running on a host.**
- **Leverage kernel features to “fake the world”.**
- **Like chroot on steroids.**

Containers (cont.)

- **Provide processes an alternate view of:**
 - Filesystem
 - Network
 - Other processes
 - Resource limits
 - ... and more
- **Docker, rkt, containerd, K8s all provide sane defaults + tooling.**
 - Make containers behave somewhat like Vms.
 - Images, management, etc...

Containers (cont.)

- **A separation of concerns.**
 - Bundle dependencies (libraries, runtimes, userland tools) into one image.
 - Good for deployment.
- **Reproducible artefacts that are the same across all environments.**
- **Simple package management.**
- **A building block for a microservices architecture.**

Containers (cont.)

- **Cassandra in a container? Sure!**

But:

- Use host networking.
- Mount a XFS volume into the container.
- Data persistence.
- Some host tuning required:
 - sysctls
 - tcp & virtual memory tuneables
 - disk IO scheduler & read ahead
 - disable swap

Kubernetes

- **Service for running and managing containers.**
- **Across a cluster of servers.**
- **Runs anywhere (all major cloud providers, on-prem)**

- **Clean room implementation of Borg by Google.**
- **Google is one of many contributors**

Kubernetes (cont.)

- **Manages dependent/related containers**
- **Manages storage**
- **Distribution of secrets (keys, certs, etc)**
- **Monitoring application health**
- **Replication & scaling**
- **Load balancing**
- **Version updates**
- **RBAC!**

Kubernetes (cont.)

- **It won the “war”:**
 - AWS EKS
 - Google Cloud Kubernetes Engine
 - Azure Kubernetes Service
 - Red Hat OpenShift
 - Pivotal Kontainer Service
 - CoreOS
 - Mesosphere
 - Docker Swarm

... and more

Cassandra Operator for Kubernetes (cont.)

- Simplifies provisioning and management of Cassandra on K8s
- Cassandra-as-a-Service on top of Kubernetes
- Instaclustr in a box
- Scale safely
- Backups
- Repairs
- Security (encryption, certs, user management)
- Out-of-the-box Prometheus integration (monitoring)

Cassandra Operator for Kubernetes (cont.)

- **A custom resource definition (CRD) allows end users to create “Cassandra” objects in Kubernetes.**
 - Contains configuration options for Cassandra
 - e.g. node count, JVM tuning options, security settings, etc.
 - **The Controller listens to state changes on the Cassandra objects.**
 - **Modifies/creates resources to match:**
 - StatefulSets
 - Services
 - Secrets & ConfigMaps
 - Prometheus ServiceMonitors
- ... and more.**

Cassandra Operator for Kubernetes (cont.)

- Get it here:

<https://github.com/instaclustr/cassandra-operator>

Summary

- **Instarepair & cassandra-sstable-tools**
- **Cassandra Kerberos Plugin**
- **Cassandra LDAP Plugin**
- **Backup & Restore Tool**
- **Cassandra Exporter for Prometheus**
- **Cassandra Operator for Kubernetes**

instaclustr

Adam Zegelin

Co-founder @
Instaclustr

adam@instaclustr.com

©Instaclustr Pty Limited, 2019

<https://www.instaclustr.com/company/policies/terms-conditions/>

Except as permitted by the copyright law applicable to you, you may not reproduce, distribute, publish, display, communicate or transmit any of the content of this document, in any form, but any means, without the prior written permission of Instaclustr Pty Limited