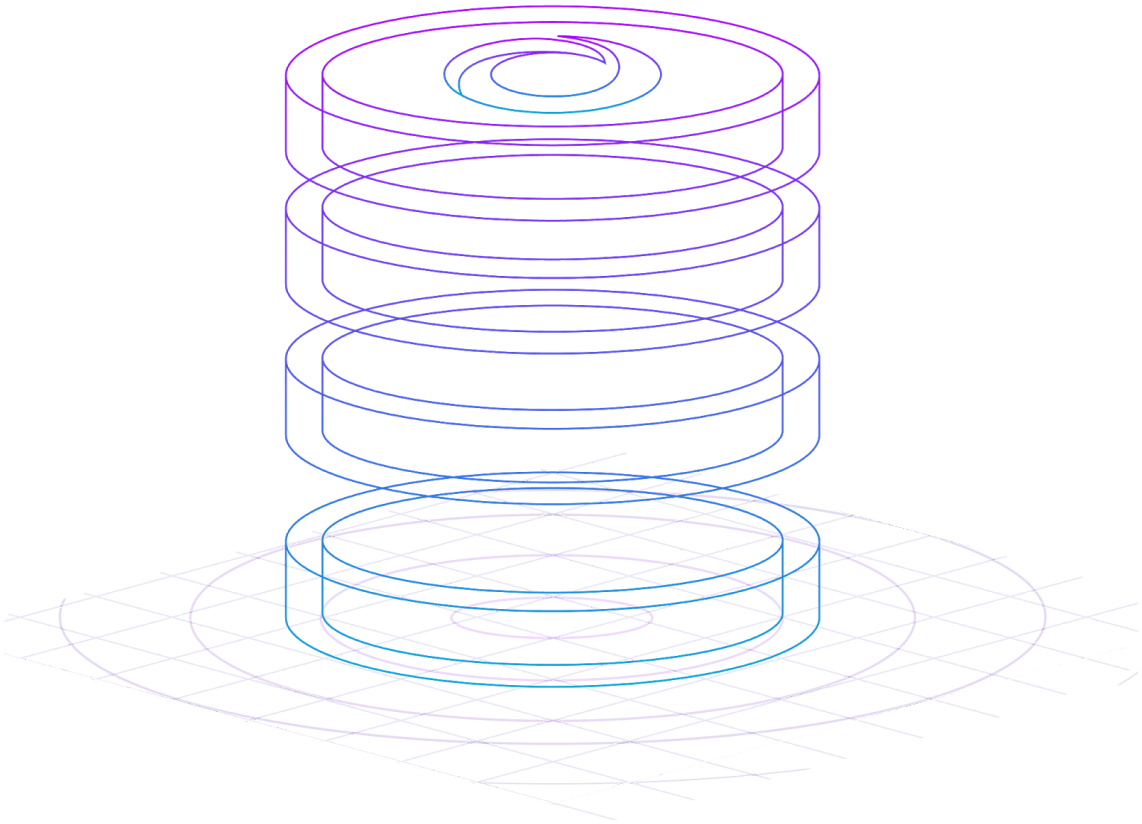
GUIDE

Proof of Concept



2020

SingleStore

Table of Contents

How to Launch a SingleStore Managed Service POC Cluster **3**

Guided Walkthrough **3**

Self-Guided Process **4**

- Create Table 4
- Load Data 4
- Test Queries 4
- Overall Concepts 4
- Third Party Integrations 4

Build a Sample Ad Tech App **5**

- Create database and columnstore table 5
- Establish a reference table 5
- Insert sample data 5
- Load data from Kafka 5
- Create the Pipeline 6
- Start Pipeline 6
- Sample Queries 6

Conclusion **7**

SingleStore Managed Service is a modern, converged data platform that allows you to leverage data of any kind: relational, geospatial, document, etc. to approach many different challenges that you may be facing.

How to Launch a SingleStore Managed Service POC Cluster

We look forward to helping you in your POC of SingleStore Managed Service. Below, you will find details on how to get started. These steps are also articulated in this video if you prefer a visual walkthrough.

1. **Click link in email**
    a. Once your POC cluster is approved, you will receive an email confirmation with the subject "You can now create a proof of concept". Click the link to take you to the "Create Cluster…" prompt.
2. **Log into the SingleStore Portal**
    a. Use your existing credentials to sign in
3. **Create Managed cluster**
    a. Click "Create Managed Cluster…" then name and tag your cluster.
4. **Select Proof of Concept**
    a. You will now see the Proof of Concept option 'Enabled'. Complete this form according to the cluster type and size that you would like to use.
5. **Edit security options**
    a. Set your cluster password and restrict access by IP.
6. **Connect to Studio**
    a. Your cluster should be ready shortly. Once it's ready, click "Load Data with SingleStore Studio".

Self-Guided Process

The following page details a sample Proof of Concept to help you get started independently. You may also use this sample POC as a guide.

Create Table

Creating tables in Managed Service is very straightforward, and also extensible. Table structure will be critical in determining the performance of your queries. SingleStore Managed Service follows the MySQL wire protocol.

Key concepts like compression, sharding, etc. are covered here: CREATE TABLE.

Load Data

This simple guide will help you through bringing your own data into SingleStore Managed Service from Amazon S3, Azure Blob, Google Cloud Storage, HDFS, Kafka, and more: SingleStore Pipelines

Test Queries

Feel free to start testing queries that you are comfortable with. To improve performance on your query, you may check out the icon and select PROFILE to get a visual explanation. In addition, you may use the Resource Usage tab to record queries and observe potential latencies.

Overall Concepts

For more help with concepts within SingleStore Managed Service that may assist you in loading data, query tuning, etc. please see the guide here: SingleStore Managed Service Concepts.

Third Party Integrations

SingleStore has integrations with common ingestion, reporting and data governance tools. For more information on what tools are available and how to set them up, look here.

Build a Sample Ad Tech App

In this example, you will see how to monitor the effectiveness of ad spend in real-time. You will create tables, connect to a live data stream, and use pre-configured queries to see the speed and scale of SingleStore Managed Service. The platform can then be used to drive critical business decisions around targeting and spend through dashboarding and downstream decisioning.

Create database and columnstore table

```
DROP DATABASE IF EXISTS adtech;
CREATE DATABASE adtech;
USE adtech;
CREATE TABLE events ( user_id int,
event_name varchar(128),
advertiser varchar(128), campaign int(11),
gender varchar(128),
income varchar(128), page_url varchar(128),
region varchar(128),
country varchar(128),
KEY adtmidx (user_id,event_name,advertiser,campaign)
USING CLUSTERED COLUMNSTORE, SHARD KEY user_id (user_id));
```

Establish a reference table

```
CREATE REFERENCE TABLE campaigns ( campaign_id smallint(6) NOT NULL DEFAULT '0',
campaign_name varchar(255) CHARACTER SET utf8 COLLATE utf8_general_ci
DEFAULT NULL, PRIMARY KEY (campaign_id));
```

Insert sample data

```
INSERT INTO `campaigns` VALUES (1,'demand great'),(2,'blackout'),(3,'flame
broiled'),(4,'take it from a fish'),(5,'thank you'),(6,'designed by
you'),(7,'virtual launch'),(8,'ultra light'),(9,'warmth'),(10,'run
healthy'),(11,'virtual city'),(12,'online lifestyle'),(13,'dream
burger'),(14,'super bowl tweet');
```

Load data from Kafka

```
Kafka topic: public-kafka.memcompute.com/ad_events
```

Create the Pipeline

```
CREATE or REPLACE PIPELINE events_pipeline AS LOAD DATA KAFKA
'public-kafka.memcompute.com:9092/ad_events'
BATCH_INTERVAL 2500
INTO TABLE events
FIELDS TERMINATED BY '\t' ENCLOSED BY '' ESCAPED BY '\\'
LINES TERMINATED BY '\n' STARTING BY ''
(user_id,event_name,advertiser,campaign,gender,income,page_url,region,country);
```

Start Pipeline

```
/* Add ALTER for latest offsets */ ALTER PIPELINE events_pipeline SET OFFSETS
earliest;
START PIPELINE events_pipeline;
```

Sample Queries

```
/* Query 1 - How many events have we processed? */
SELECT COUNT(*) FROM events;

/* Query 2 - How many users stands in a specific income range of '100K+.' */
SELECT user_id,advertiser, event_name,gender,country FROM events
WHERE income = "100k+"
group by campaign
ORDER BY advertiser desc;

/* Query 3 - Find the traditional funnel campaigning information for the advertiser
named Walgreens. */
SELECT
Campaign, Campaign_name, impression_count,click_count,downstream_conversion_count,
click_count / impression_count AS conv_1,
downstream_conversion_count / click_count AS conv_2, downstream_conversion_count /
impression_count AS all_conv
FROM (
       SELECT campaign,
       SUM(CASE WHEN (event_name="Impression") THEN 1 ELSE null END) AS
       impression_count, SUM(CASE WHEN (event_name="Click") THEN 1 ELSE null  END)
```

```
AS click_count,
SUM(CASE WHEN (event_name="Downstream Conversion") THEN 1 ELSE null END) AS
downstream_conversion_count
FROM events
WHERE advertiser = "Walgreens"
group by campaign) tab
LEFT JOIN campaigns ON campaigns.campaign_id = campaign
ORDER BY all_conv desc;

/* Query 4 - Find the conversion metrics information for the advertiser named
Walgreens. */
SELECT campaign, advertiser, country, SUM(CASE WHEN (event_name="Impression") THEN
1 ELSE null END) AS impression_count,
SUM(CASE WHEN (event_name="Click") THEN 1 ELSE null END) AS click_count, SUM(CASE
WHEN (event_name="Downstream Conversion") THEN 1 ELSE null END) AS
downstream_conversion_count FROM events
WHERE advertiser = "Walgreens";


/* Query 5 - Targeted campaign information for advertiser named McDonalds. */
SELECT user_id, page_url,region, country, SUM(CASE WHEN (event_name="Click") THEN 1
ELSE null END) AS click_count, SUM(CASE WHEN (event_name="Downstream Conversion")
THEN 1 ELSE null END) AS downstream_conversion_count FROM events
WHERE advertiser = "McDonalds"
group by campaign
ORDER BY user_id desc;
```

After you've built these tables, check out how to connect your favorite dashboard tools here!
Tableau
Looker
PowerBI

Conclusion

As you've seen in this guide, getting started with SingleStore Managed Service is super easy and the platform allows you to approach your business challenges with the speed of real-time, limitless scale and the ease of SQL.

We look forward to engaging with you further in your journey with the Database of Now. At this point, you may be interested in trying SingleStore Managed Service with the help of a SingleStore Engineer. You can contact us via singlestore.com/managed-service/ for next steps.