Transitioning Profiler Manager Service into SDX (Preview)

Date published: 2022-02-24

Legal Notice

Contents

Transitioning Profiler Manager Service into SDX Infrastructure

The current Data Catalog architecture includes profilers is a separate DataHub cluster that comprises of two types of components:

1. Profiler Manager Service (Profiler Admin/Profiler Metrics/Data Discovery Service) [PMS]
2. Profiler Scheduler
   - Profiler Services (Admin, Metrics, and Scheduler)
   - Profiler jobs - The Profiler services (Admin and Metrics) are required for serving user interactions from the Data Catalog app. This implies the profiler cluster cannot be shutdown without loss of service to users.

   Considering the above scenarios, the PMS is moved to the SDX infrastructure which results in retaining the scheduler and profiler jobs in a separate DataHub cluster.

These changes have the following benefits:

- Allows the user to completely shutdown the profiler cluster if needed to save time and cost.
- Allows the user to handle other platform capabilities like upgrade, High Availability and other pertinent features in a consistent manner for services needed by the Data Catalog App.

Known limitations

Some of the known limitations seen during the transition:

- Profiler Manager service on SDX is only supported in medium duty and it is behind an entitlement **"CDP_DATA_LAKE_MEDIUM_DUTY_WITH_PROFILER"**.

- If the entitlement is enabled and the user starts using the MEDIUM duty SDX cluster, Data Catalog creates a profiler cluster without Profiler Manager Service. Else, as usual, Profiler Manager Service would be created on the DataHub cluster.