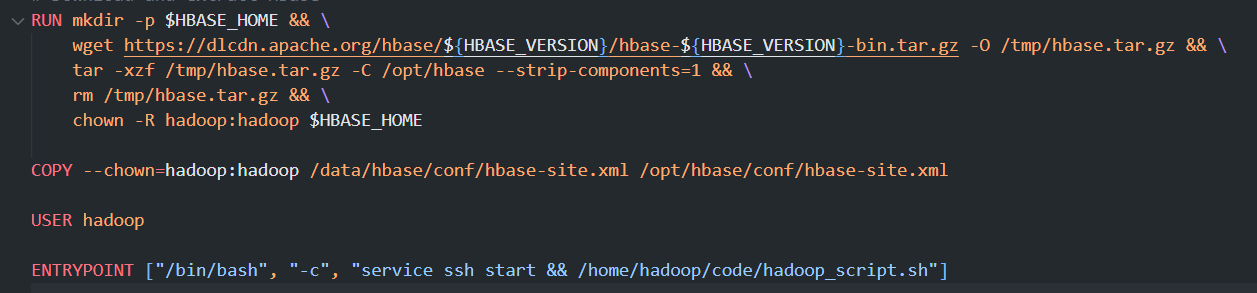
This Dockerfile sets up an environment that includes **Hadoop**, **Zookeeper**, and **HBase** on an Ubuntu 22.04 base image, with additional setup for Python3 and several related Python packages. It also configures a user (hadoop) and sets up SSH and directories to allow the Hadoop cluster to run smoothly. Below is a detailed breakdown of the Dockerfile and its components, with a focus on **HBase**.

This Dockerfile is designed to:

1. Set up an Ubuntu-based container with necessary packages.
2. Install and configure **Hadoop 3.3.6**.
3. Install and configure **Zookeeper 3.8.4**.
4. Install **HBase 2.5.11**.
5. Set up a non-root user (hadoop) to run the services.
6. Configure SSH and Hadoop directories.
7. Copy necessary configuration files.
8. Set up environment variables and paths.
9. Execute a custom Hadoop bootstrap script on container startup.

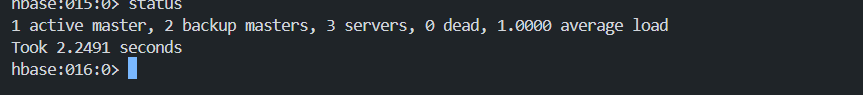
**Downloading and Installing HBase**



**Then adding the hbase-site.xml**

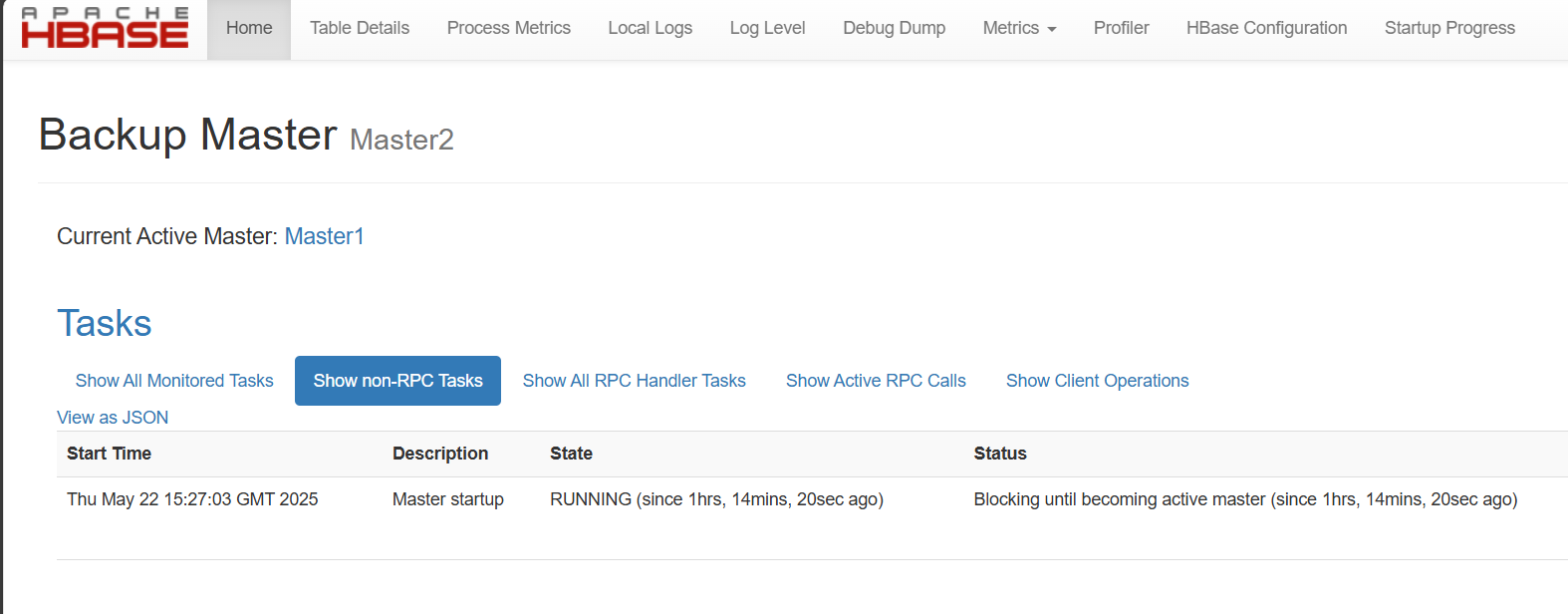
****

**Testing for failover**

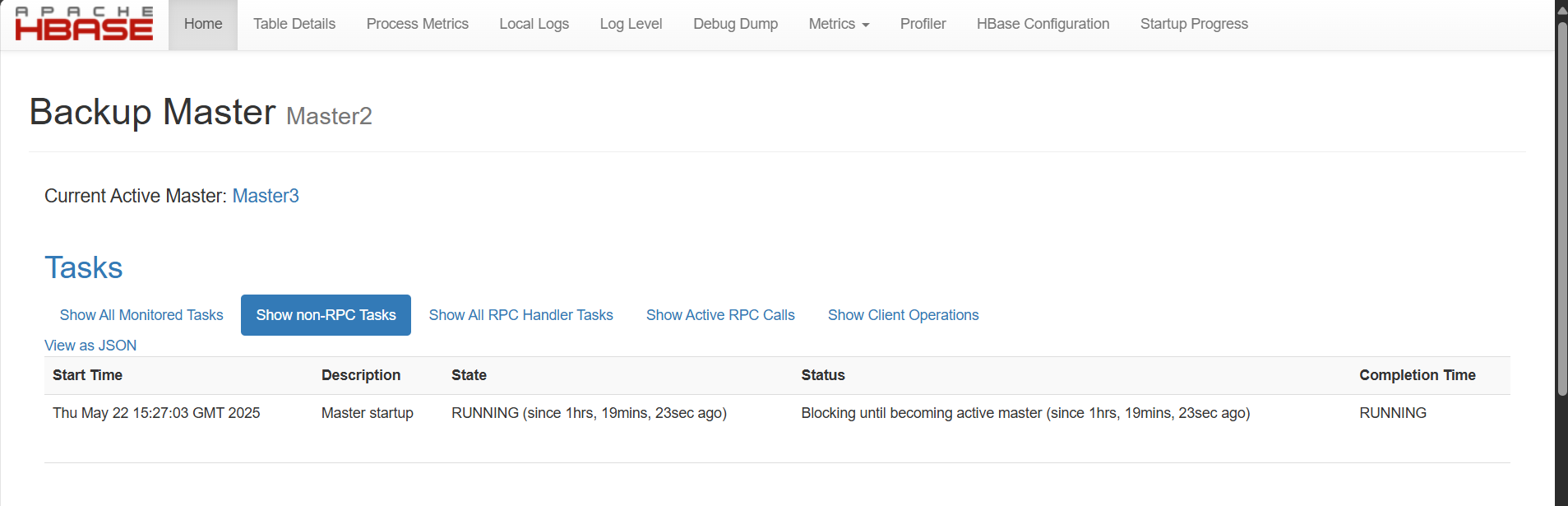
****

**Having one master two stand by and three region server**

**Current active master is master1**

****

**After stoping the master1**

****

**Table Design & Implementation**

**Using this format**

<domain>!<url>

Column Families

| **Family** | **Purpose** | **Versions** | **TTL** |
| --- | --- | --- | --- |
| content | HTML page content | 3 |  |
| metadata | Page metadata | 1 |  |
| outlinks | Outbound links | 2 |  |
| inlinks | Inbound links | 2 |  |

Putting data and retrieve

