

How to use the Redback Operations Data Lakehouse

Firstly, access the Deakin virtual machine

To access the virtual machine requires an account, set up by someone with Admin to the Virtual machine backend. You can contact a mentor in the capstone unit to arrange this.

1. Login to the anyconnect VPN

Enter Deakin username and password where required and if required, complete the Duo two-factor authentication.

2. Open cmd prompt.

To access the virtual machine enter the following commands:

3. `ssh [yourname]@redback.it.deakin.edu.au`

Then your password (the VM is Linux, and automatically hides the password when entered)
(pwd supplied on virtual machine setup, by admin)

Accessing the Lakehouse

Check if the programs are up and running on docker.

In cmd:

```
docker ps
```

this will list the programs running in docker in the vm

If there is no listing of:

```
nessie, minioserver and Dremio
```

You can check the last time the programs were running with:

```
docker ps -a
```

If the result is similar to below:

```
e9e360861525  projectnessie/nessie      "/usr/local/s2i/run"    2 weeks ago  Exited (143) 4 days ago
ed227966b6c3  minio/minio               "/usr/bin/docker-ent..." 2 weeks ago  Exited (0) 4 days ago
e1a55567d611  dremio/dremio-oss:latest  "bin/dremio start-fg"   2 weeks ago  Exited (143) 4 days ago
```

Then it will require a restart

To restart, enter in cmd while in the VM:

```
docker start dremio
```

```
docker start minioserver
```

```
docker start nessie
```

This will boot up the images in the .yml (stored in the VM) and programs will be running again.

```
kaylin@redback1:~$ docker start dremio
dremio
kaylin@redback1:~$ docker start minioserver
minioserver
kaylin@redback1:~$ docker start nessie
nessie
kaylin@redback1:~$ docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS
5dc0f07aaa2e	wazuh/wazuh-dashboard:4.5.0	"/entrypoint.sh"	4 days ago	Up 4 days	443/tcp, 0.0.0.0:443->5601/tcp, :::443->5601/tcp
e8a753e4eaf6	wazuh/wazuh-manager:4.5.0	"/init"	4 days ago	Up 4 days	0.0.0.0:1514-1515->1514-1515/tcp, :::1514-1515->1514-1515/tcp
55000/tcp, 1516/tcp	single-node-wazuh.manager-1				
b941bdb4f549	wazuh/wazuh-indexer:4.5.0	"/entrypoint.sh open..."	4 days ago	Up 4 days	0.0.0.0:9200->9200/tcp, :::9200->9200/tcp
e9e360861525	projectnessie/nessie	"/usr/local/s2i/run"	2 weeks ago	Up 6 seconds	8080/tcp, 8443/tcp, 0.0.0.0:19120->19120/tcp
ed227966b6c3	minio/minio	"/usr/bin/docker-ent..."	2 weeks ago	Up 13 seconds	0.0.0.0:9000-9001->9000-9001/tcp, :::9000-9001->9000-9001/tcp
e1a55567d611	dremio/dremio-oss:latest	"bin/dremio start-fg"	2 weeks ago	Up 24 seconds	0.0.0.0:9047->9047/tcp, :::9047->9047/tcp
5678/tcp	dremio				

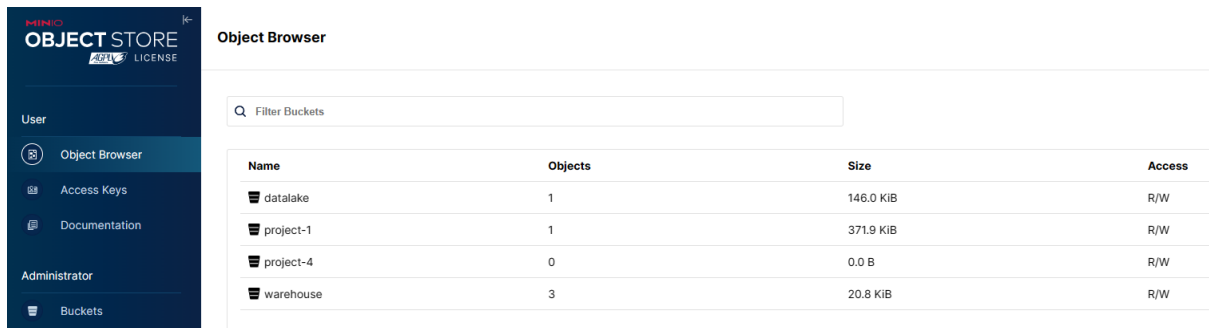
```
kaylin@redback1:~$
```

And accessible through the browser with the appropriate access, (set-up handled by a data warehouse team member

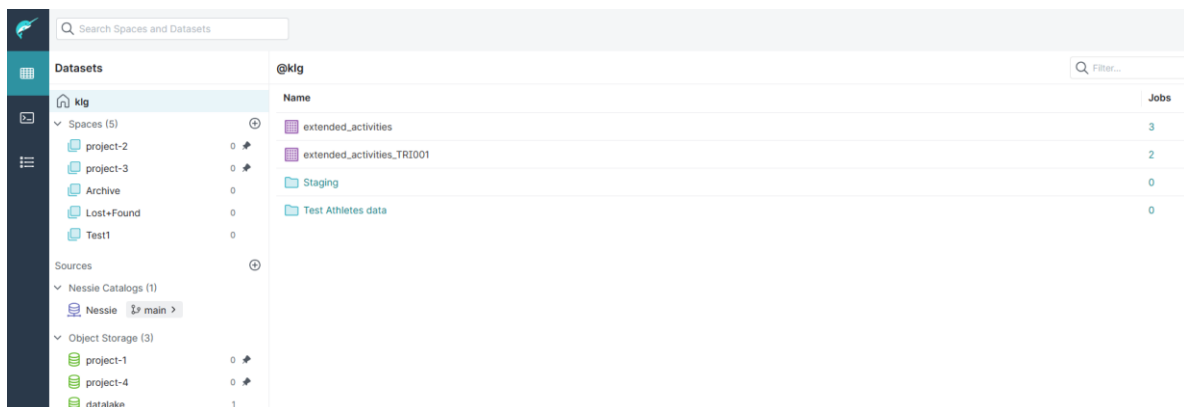
Dremio: <http://10.137.0.149:9047/>

Minio: <http://10.137.0.149:9001/login>

From here you can control the programs from the UI



Name	Objects	Size	Access
datalake	1	146.0 KiB	R/W
project-1	1	371.9 KiB	R/W
project-4	0	0.0 B	R/W
warehouse	3	20.8 KiB	R/W



Name	Jobs
extended_activities	3
extended_activities_TRI001	2
Staging	0
Test Athletes data	0

Any issues or account set-up requests please contact the Data warehouse project team

The .Yaml code to run the tools is below:

```
version: "3.9"

services:

  dremio:
    platform: linux/x86_64
    image: dremio/dremio-oss:latest
    ports:
      - 9047:9047
      - 31010:31010
      - 32010:32010
    container_name: dremio

  minioserver:
    image: minio/minio
    ports:
      - 9000:9000
      - 9001:9001
    environment:
      MINIO_ROOT_USER: minioadmin
      MINIO_ROOT_PASSWORD: minioadmin
    container_name: minioserver
    command: server /data --console-address ":9001"

  nessie:
    image: projectnessie/nessie
    container_name: nessie
    ports:
      - "19120:19120"

networks:
  default:
    name: iceberg_env
    driver: bridge
```