

Admin Training

Rev 3 Training Day

May 4, 2022

Akshay Ambekar, Field Engineer

Admin Course Objectives



Upon Completion of this Course, Learners will be able to:

Describe the Domino Architecture, Components, and Configuration.

Explain Managing the Domino Environment & Infrastructure.

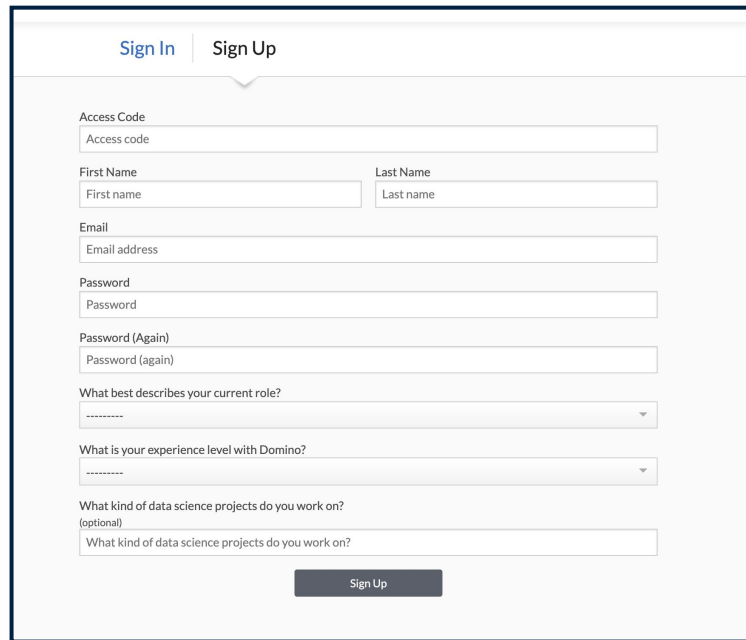
Discuss Domino Storage and how to Manage it.

Define Optimizing and Troubleshooting Domino.

Describe Domino Monitoring, Data Management, and Security.

DOMINO ACADEMY

1. Go to learn.dominodatalab.com
2. Click on “Sign Up”
3. Input your access code 2sypw-iehkb44g
4. Select the Admin Training course



The screenshot shows the 'Sign Up' page of the Domino Academy. At the top, there are two tabs: 'Sign In' and 'Sign Up', with 'Sign Up' being the active tab. Below the tabs, the form contains several input fields and dropdown menus. The fields are: 'Access Code' (with a placeholder 'Access code'), 'First Name' (with a placeholder 'First name'), 'Last Name' (with a placeholder 'Last name'), 'Email' (with a placeholder 'Email address'), 'Password' (with a placeholder 'Password'), 'Password (Again)' (with a placeholder 'Password (again)'), 'What best describes your current role?' (a dropdown menu), 'What is your experience level with Domino?' (a dropdown menu), and 'What kind of data science projects do you work on?' (with a sub-label '(optional)' and a placeholder 'What kind of data science projects do you work on?'). At the bottom right of the form is a dark grey button labeled 'Sign Up'.

Sign In | Sign Up

Access Code
Access code

First Name
First name

Last Name
Last name

Email
Email address

Password
Password

Password (Again)
Password (again)

What best describes your current role?

What is your experience level with Domino?

What kind of data science projects do you work on?
(optional)
What kind of data science projects do you work on?

Sign Up

PREREQUISITES

You are familiar with Kubernetes.

The level of required familiarity will depend on how deep you want to go.

Recommendations for learning:

- A Cloud Guru online course for surface level Kubernetes (4 hours)
- Online training from Linux Foundation to train for Kubernetes Admin Certification (35 hours + homework)



PREREQUISITES

Skillsets for Domino Administration

The following table lists some of the skillsets needed to administer Domino, and a few options for how to learn about various technologies. You may spread administration among numerous people, in which case these tasks would be shared, or you may have a Domino managed instance where you do not need to do more than application level administration work.

Technology	Recommendations
Kubernetes	Kubernetes Basics , Cloud Guru Online Course (4 hours) , Online Training from the Linux Foundation (35 hours)
Docker	Intro to Containers, VMs, and Docker , Docker for Beginners
Git	Learn Git, Git on the Command Line
Networking	Networking for Developers
Linux	Intro to Linux
Python Package Management	Overview of Python management tools , Intro to Python Modules and Packages
AWS (if using)	Understanding of the AWS services used for Domino . Consider taking the AWS Certified Cloud Practitioner exam .
Google Cloud (if using)	Understanding of the GKE services used for Domino .
Azure (if using)	Understanding of the AKS services used for Domino .

PREREQUISITES II

Admin access to Domino

During the presentation, we will take short pauses so that you can click around areas of the product to get familiar.

However, please don't modify anything during the admin training unless asked to do so.

RESPONSIBILITIES OF DOMINO ADMINISTRATORS

As *infrastructure admins*

- Set up network connectivity to data sources and other external resources
- Maintain and support infrastructure
- Manage the compute grid (hardware tiers and resource quotas)

As *application admins*

- Manage user access
- Administer global environments
- Manage application-level configurations (built-in features to Domino that require customization or enablement)
- Certify best practices and spreading those to the user community

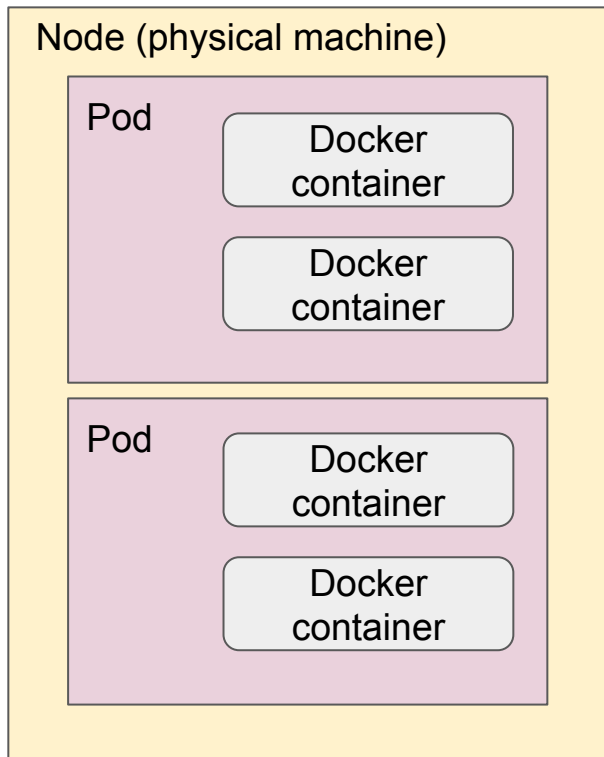
Both are responsible for

- Helping debug issues with Domino support

Introduction to Domino

KUBERNETES TERMINOLOGY REVIEW

Docker Container	A standard unit of software that packages up code and all its dependencies
Pod	A group of containers, co-located and co-scheduled, to run some application-specific logic
Node	A worker machine, which can host pods
Namespace	A virtual cluster within the larger Kubernetes cluster



Overview of Critical Product Areas

1	Workspaces
2	Environments
3	Model APIs

Namespaces



Platform

Compute

Browsers

CLI

API

Frontends

Client layer

API server

Keycloak

Dispatcher

Redis

Fluentd

HA
MongoDB

Blob Storage

Git

RabbitMQ

Docker
Registry

Elasticsearch

Postgres

Kubernetes
Masters

Shared Storage

Service layer

Image
Builds

Jobs

Workspaces

Apps

Models

Execution layer

GETTING STARTED

Lesson on navigating the Admin Panel -

<https://learn.dominodatalab.com/admin-training/855680>

Set-up Checklist - <https://learn.dominodatalab.com/admin-training/894574>

Configuration

CENTRAL CONFIGS

- Various configuration options of your Domino installation.
- In order for changes to take effect, you'll need to restart the Domino frontend process.

- **BE CAREFUL!!!**
 - Typos can cause downtime!
 - We recommend consulting with Domino support before altering any values.

The screenshot shows the 'Configuration Management' page in the Domino admin console. The top navigation bar includes links for Admin home, Projects, Users, Data, Workspaces, Executions, Infrastructure, Notifications, Advanced, and Help. The 'Advanced' menu is open, showing options like Usage, Resource Usage Report, User Activity Report, Manage Environments, Hardware Tiers, Resource Quotas, Central Config (highlighted), Feature Flags, Email Settings, Search Config, Project Stage Configuration, Jira Configuration, Restart Services, and MongoDB. Below the navigation bar, a yellow warning banner states: 'Changes here do not take effect until services are restarted. Click [here](#) to restart services.' The main content area displays a table of configuration keys.

namespace	name	key	
common		authentication.oidc.externalOrgsEnabled	
common		com.cerebro.domino.apps.contentSecurityPolicy.whiteListedImageSrcList	
common		com.cerebro.domino.auth.aws.sts.defaultSessionDuration	
common		com.cerebro.domino.auth.aws.sts.enabled	
common		com.cerebro.domino.auth.aws.sts.region	
common		com.cerebro.domino.auth.refreshTokenInRun.enabled	true
role	dispatcher	com.cerebro.domino.aws.customExecutorTagsAsJson	{}
common		com.cerebro.domino.aws.instanceProfileName	aws-staging-executor





CENTRAL CONFIG COLUMNS

Column	Description
namespace	[common] NOT K8s namespace. Directs configuration at specific components of Domino.
name	[<empty>] If namespace role or instance are specified, name applies to the specific role or instance.
key	The key of the configuration
value	The value for the configuration

Configuration Management

[Add Record](#)

Changes here do not take effect until services are restarted. Click [here](#) to restart services.

namespace	name	key	value	
common		authentication.oidc.externalOrgsEnabled	true	 
common		com.cerebro.domino.auth.aws.sts.region	us-west-2	 
<div><div>common</div><div>instance</div><div>✓ role</div></div>	dispatcher	com.cerebro.domino.aws.customExecutorTagsAsIso	{}	Save Cancel

CENTRAL CONFIG DEFAULTS

If there is no record explicitly set, then the default value is used. See our docs for keys and default values:

https://admin.dominodatalab.com/en/latest/configuration/Central_configuration.html

Admin Lesson - <https://learn.dominodatalab.com/admin-training/855692>

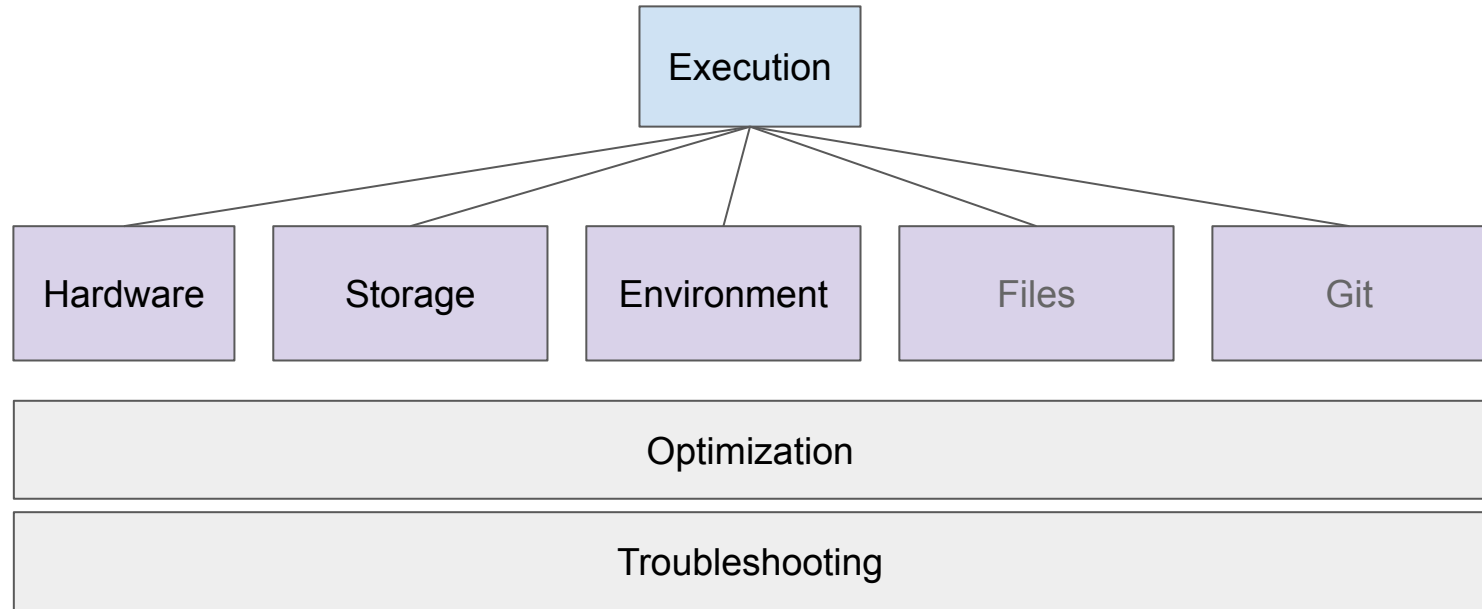
ACTIVITY: LOCATE CENTRAL CONFIG VALUES

1. Sign into the temporary training Domino instance as an Admin:
<https://admin-rev.workshop.domino.tech/>
2. Determine the central config values for the following:
 - a. Maximum executions per user
 - b. Default volume size for jobs and workspaces

Hint: The docs can be found [here](#)

Managing the Compute Grid Infrastructure

COMPONENTS REQUIRED FOR AN EXECUTION

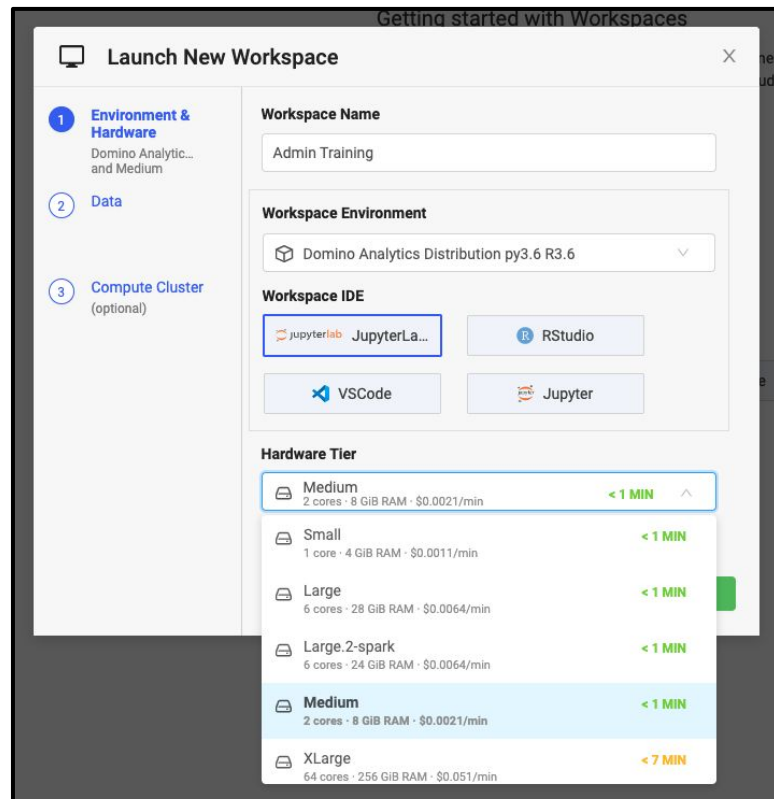


UNDERSTANDING THE STATE OF THE CLUSTER WHEN EXECUTING CODE

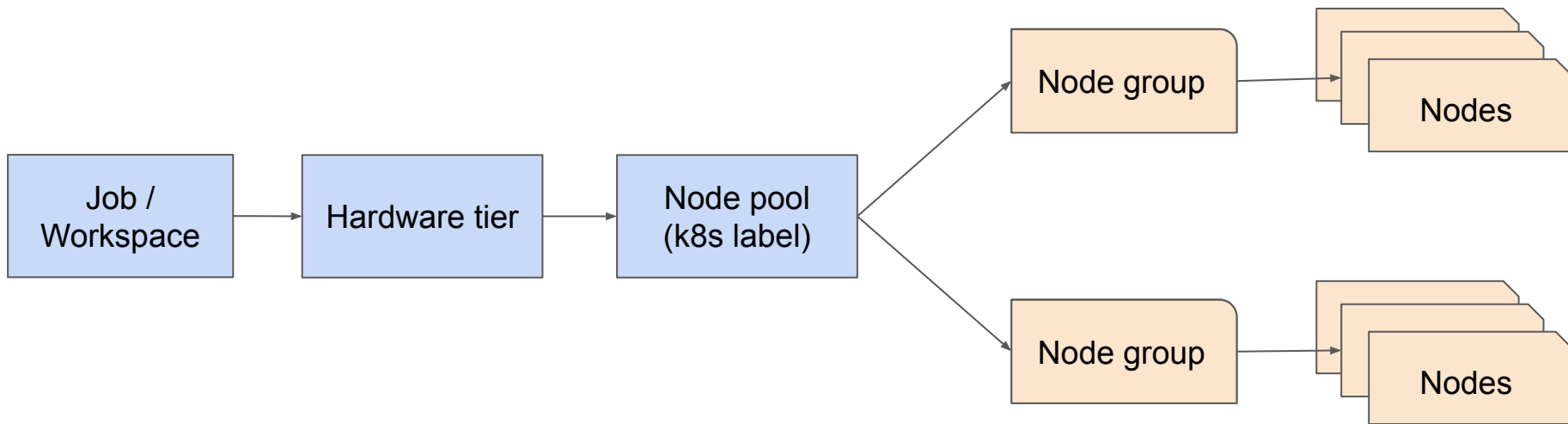
When running a Job or a Workspace, the workload gets scheduled on a hardware tier.

The hardware tier defines

- The type of machine that it will run on
- The resource limits for the pod that the run will execute in



EXECUTION LIFECYCLE



ADMIN PANEL - EXECUTIONS PAGE

Executions



Workload ▾	Execution ID ▾	Hardware Tier ▾	User ▾	Project ▾	Title ▾	Created ▾	Status ▾	Infrastructure Details	Actions
App	60413b12bf059855192f	k8s-small: 1 cores, 4 GiB	elliott_botwick	elliott_botwick/churn-demo-cleaned	App server started by e	2021-03-04 11:54:58	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	<button>Stop</button>
App	6047c3ddb0f059855192f	k8s-small: 1 cores, 4 GiB	ross_sharp	ross_sharp/customer_churn_2021	App server started by r	2021-03-09 10:52:13	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	<button>Stop</button>

[Resources](#)

Admin lesson - <https://learn.dominodatalab.com/admin-training/879952>

ADMIN PANEL - EXECUTIONS PAGE



[Admin home](#) [Projects](#) [Users](#) [Data](#) [Workspaces](#) [Executions](#) [Infrastructure](#) [Advanced](#)

[Help](#) [taylor_sale](#) ▾

Executions



Workload ▾	Execution ID ▾	Hardware Tier ▾	User ▾	Project ▾	Title ▾	Created ▾	Status ▾	Infrastructure Details	Actions
App	60413b12bf059855192f	k8s-small: 1 cores, 4 GiB	elliott_botwick	elliott_botwick/churn-demo-cleaned	App server started by e	2021-03-04 11:54:58	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	Stop
App	6047c3ddbf059855192f	k8s-small: 1 cores, 4 GiB	ross_sharp	ross_sharp/customer_churn_2021	App server started by r	2021-03-09 10:52:13	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	Stop
								Resources	

Output of kubectl describe

ADMIN PANEL - EXECUTIONS PAGE



[Admin home](#) [Projects](#) [Users](#) [Data](#) [Workspaces](#) [Executions](#) [Infrastructure](#) [Advanced](#)

[Help](#) [taylor_sale](#) ▾

Executions




Workload ▾	Execution ID ▾	Hardware Tier ▾	User ▾	Project ▾	Title ▾	Created ▾	Status ▾	Infrastructure Details	Actions
App	60413b12bf059855192f	k8s-small: 1 cores, 4 GiB	elliott_botwick	elliott_botwick/churn-demo-cleaned	App server started by e	2021-03-04 11:54:58	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	Stop
App	6047c3ddb059855192f	k8s-small: 1 cores, 4 GiB	ross_sharp	ross_sharp/customer_churn_2021	App server started by r	2021-03-09 10:52:13	Running	Resources Pod Node Logs ⌵ (CSV) Support Bundle	Stop


[Resources](#)

Output of kubectl describe

ADMIN PANEL - INFRASTRUCTURE PAGE

Admin homeProjectsUsersDataWorkspacesExecutions**Infrastructure**NotificationsAdvancedHelpakshay_ambekar ▾

Infrastructure

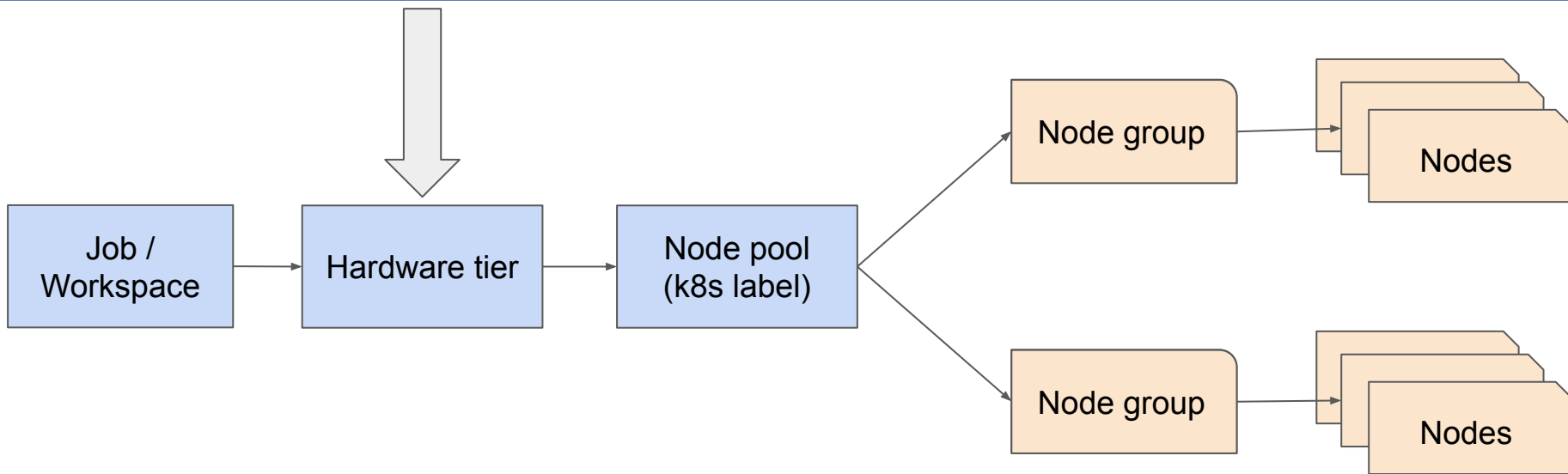
50 entries ▾

NAME ▾	INSTANCE TYPE ▾	NODE POOL ▾	BUILD NODE ▾
ip-10-0-1-21.us-west-2.compute.intern	m5.large		
ip-10-0-36-25.us-west-2.compute.intern	m5.2xlarge	platform	
ip-10-0-44-255.us-west-2.compute.inte	m5.2xlarge	platform	
ip-10-0-45-39.us-west-2.compute.intern	m5.2xlarge	default	true

Showing 1 - 4 out of 4 < **1** >

<https://admin.dominodatalab.com/en/latest/compute/compute-grid.html#how-do-i-view-the-current-nodes-in-my-compute-grid>

HOW DO YOU CREATE A NEW HARDWARE TIER?



CREATING A HARDWARE TIER

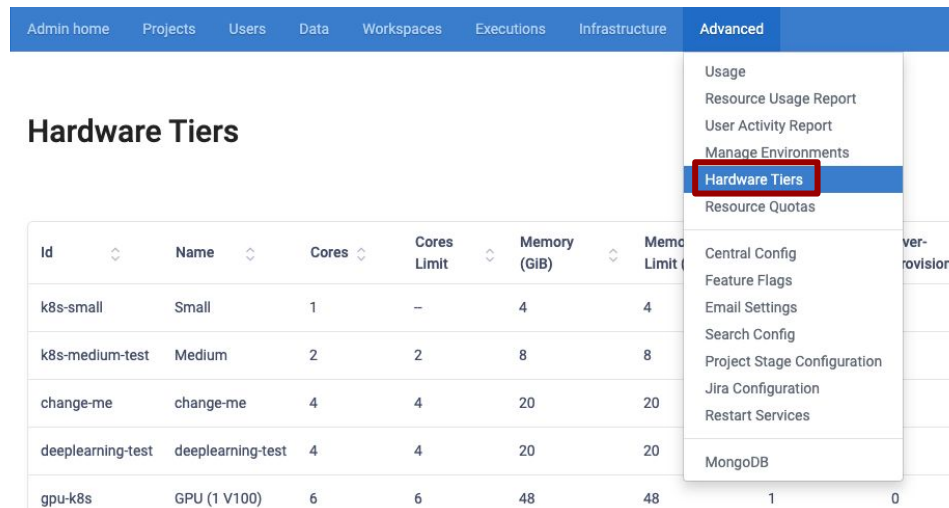
1. Create a Node Pool in Kubernetes
2. Define the Hardware Tier using the Admin Panel

Node pools lesson -

<https://learn.dominodatalab.com/admin-training/894982>

Hardware tiers lesson -

<https://learn.dominodatalab.com/admin-training/879953>



The screenshot shows the Databricks Admin Panel with the 'Advanced' tab selected. A dropdown menu is open, highlighting 'Hardware Tiers'. Below the menu, the 'Hardware Tiers' table is visible, showing columns for Id, Name, Cores, Cores Limit, Memory (GiB), and Memory Limit. The table lists several tiers: k8s-small, k8s-medium-test, change-me, deeplearning-test, and gpu-k8s.

Id	Name	Cores	Cores Limit	Memory (GiB)	Memory Limit
k8s-small	Small	1	—	4	4
k8s-medium-test	Medium	2	2	8	8
change-me	change-me	4	4	20	20
deeplearning-test	deeplearning-test	4	4	20	20
gpu-k8s	GPU (1 V100)	6	6	48	48

HARDWARE TIER OVERHEAD

When defining hardware tiers, leave room for overhead.

As a rule of thumb, **node overhead is 1.5 cores and 2 GiB of RAM.**

Additional **overhead per execution is about 1 core and 1.5 GiB.**

Taking account of overhead is especially important when you have only 1 concurrent execution per node.

ACTIVITY: DIAGNOSE ERROR WITH HARDWARE TIER

1. Sign into the temporary training Domino instance as an Admin:
<https://admin-rev.workshop.domino.tech/>
2. Find the logs for this failed workspace:
3. Determine why the hardware tier didn't work

Managing Model API Resources

[Access & Sharing](#)**Deployment**[Environment](#)[Invocation](#)[Advanced](#)

Deployment Configuration

Configure the details of your model deployment.

Scale Configuration

Number of Instances

1 Instance

Routing Mode

☐ Basic mode

In this mode, the traffic simply gets redirected to the latest deployed version. You cannot activate an older version which would take over the traffic from the latest version.

☒ Advanced mode

This mode allows you to have two deployed versions concurrently. You can promote a version to production, which would be accessible from a labeled route. You have to manually deactivate the model versions that are no longer in use.

Regenerate routes

Regenerate

Resource Quota

Compute Resources per Instance

- ✓ small: 1.0 cores / 1.0 GB RAM
- medium: 1.9 cores / 1.9 GB RAM
- large: 3.8 cores / 3.8 GB RAM

Usually, Model APIs will not need significant computational resources, since they are doing inference rather than training.

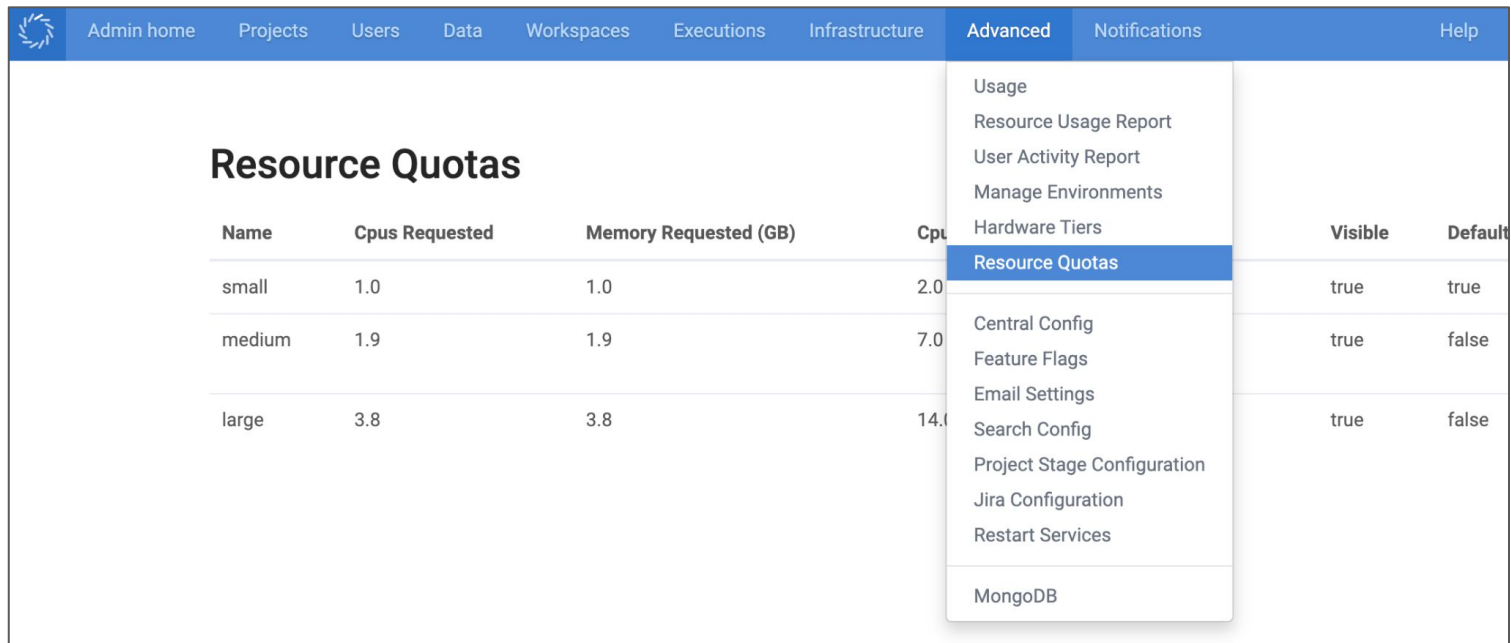
Occasionally, a Model API will break this mold. Users will need access to more computational resources.

Admins control the options listed here.

RESOURCE QUOTAS

Lessons on creating Resource Quotas and other configs for Model APIs:

<https://learn.dominodatalab.com/admin-training/894358>

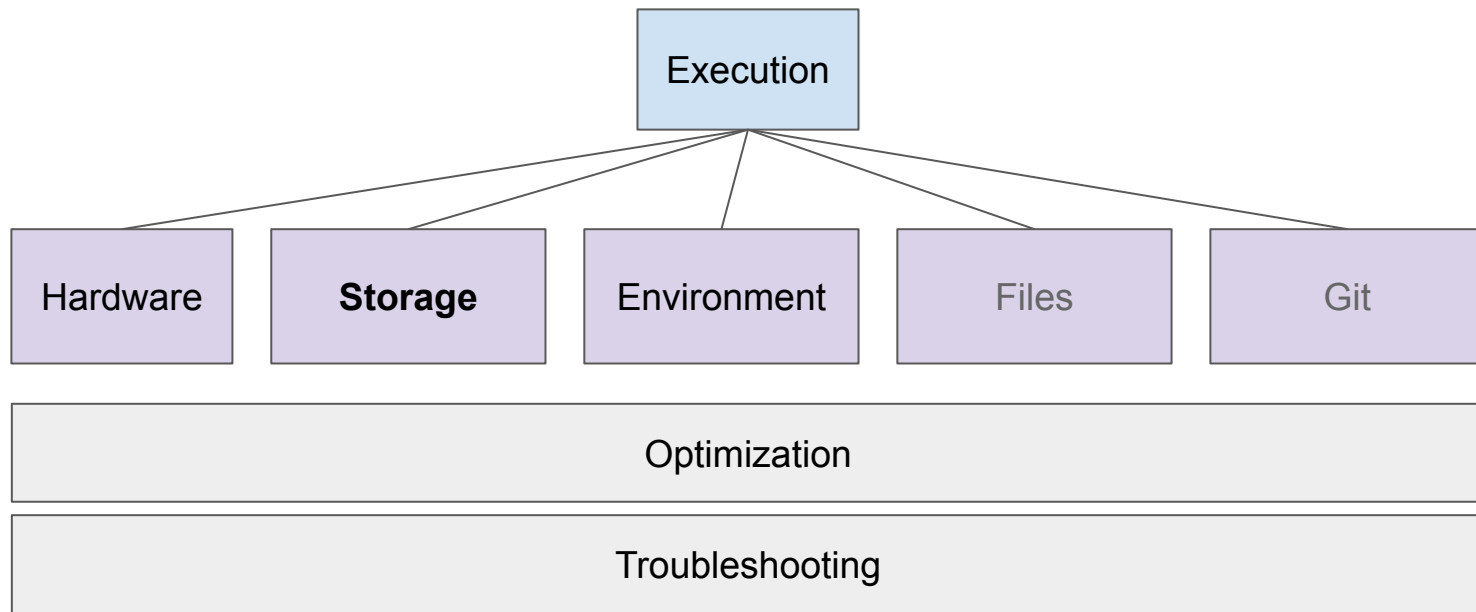


The screenshot shows the Databricks Admin console interface. The top navigation bar includes links for Admin home, Projects, Users, Data, Workspaces, Executions, Infrastructure, Advanced, Notifications, and Help. The 'Advanced' menu is expanded, showing options like Usage, Resource Usage Report, User Activity Report, Manage Environments, Hardware Tiers, Resource Quotas (highlighted), Central Config, Feature Flags, Email Settings, Search Config, Project Stage Configuration, Jira Configuration, Restart Services, and MongoDB.

Resource Quotas

Name	Cpus Requested	Memory Requested (GB)	Cpus	Visible	Default
small	1.0	1.0	2.0	true	true
medium	1.9	1.9	7.0	true	false
large	3.8	3.8	14.0	true	false

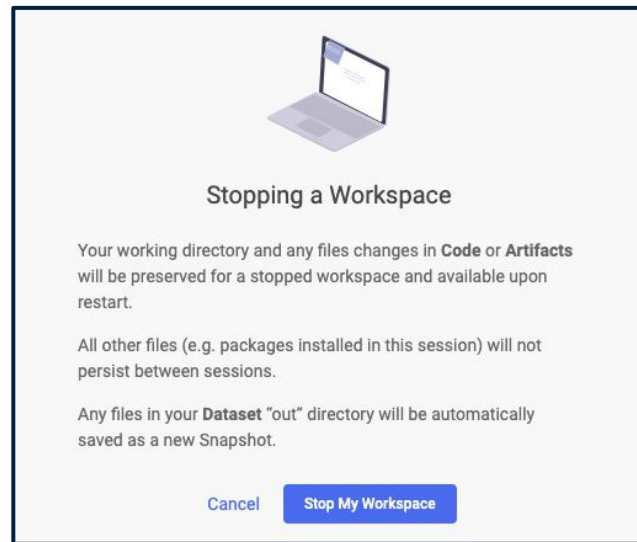
COMPONENTS REQUIRED FOR AN EXECUTION



Persistent Volume Management - Workspaces

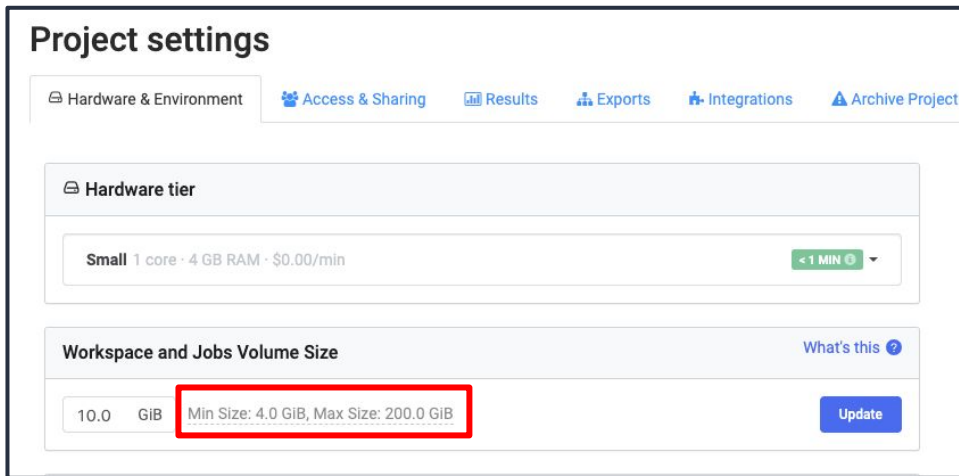
WORKSPACES

- You can stop and start workspaces without committing files
- Everything in the directory mapped to the *Files* (or *Code* and *Artifacts*) folders will persist to next session
- Objects in memory and files outside of the */mnt* directory (including pip installed packages) will not persist
- See docs [here](#) for user info



VOLUME SIZE

- Users can control the amount of storage available to workspaces and jobs in their projects
- This is updated in project settings
- Default limits are shown - this is configurable as an Admin



The screenshot displays the 'Project settings' page with a navigation bar at the top containing links for Hardware & Environment, Access & Sharing, Results, Exports, Integrations, and Archive Project. The 'Hardware & Environment' section is active, showing a 'Hardware tier' dropdown set to 'Small 1 core · 4 GB RAM · \$0.00/min' with a '< 1 MIN' indicator. Below this, the 'Workspace and Jobs Volume Size' section is highlighted, showing a current value of '10.0 GIB' and a red-bordered box containing the text 'Min Size: 4.0 GiB, Max Size: 200.0 GiB'. A 'What's this ?' link is present to the right of the section header, and an 'Update' button is located at the bottom right of the volume size configuration area.

Workspace and Jobs Volume Size	
10.0 GIB	Min Size: 4.0 GiB, Max Size: 200.0 GiB

WORKSPACE LIMITS

Even *Stopped* Workspaces use storage space

Default limits are in place for cost management:

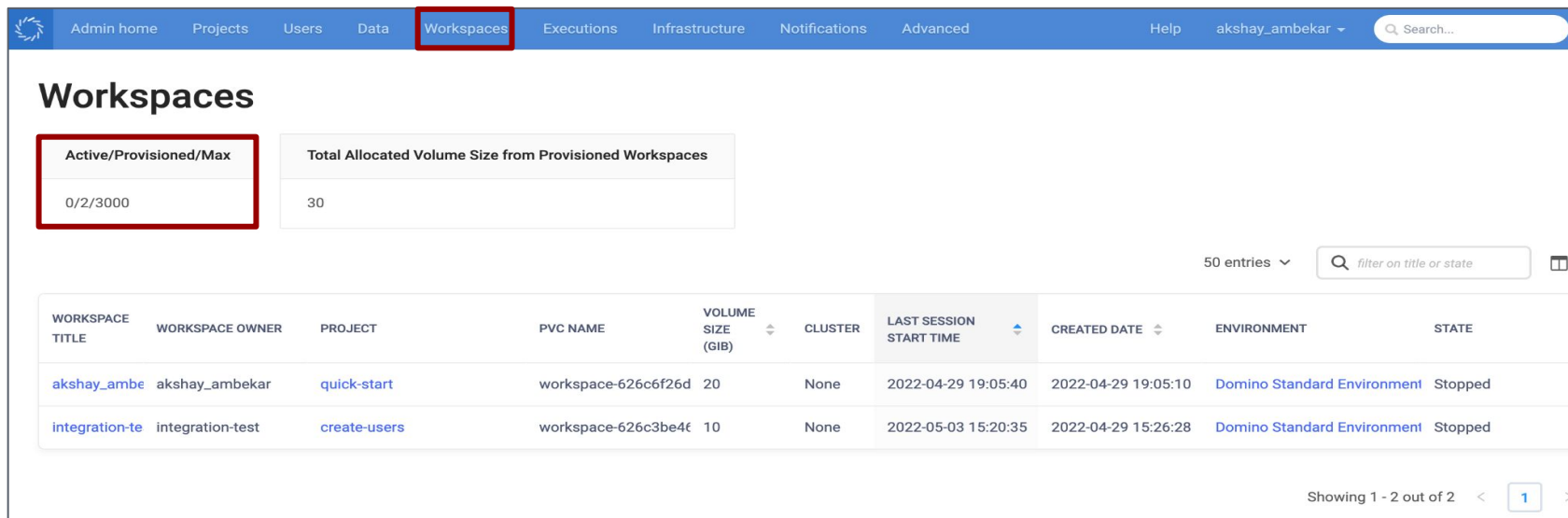
- Four workspaces per *user* per *project*
- Sixteen workspaces *total* per user
- 3000 workspaces in *deployment*

Deleting a workspace stops it permanently and removes it from limit count

Admin lesson - <https://learn.dominodatalab.com/admin-training/880015>

WORKSPACES ADMIN PAGE

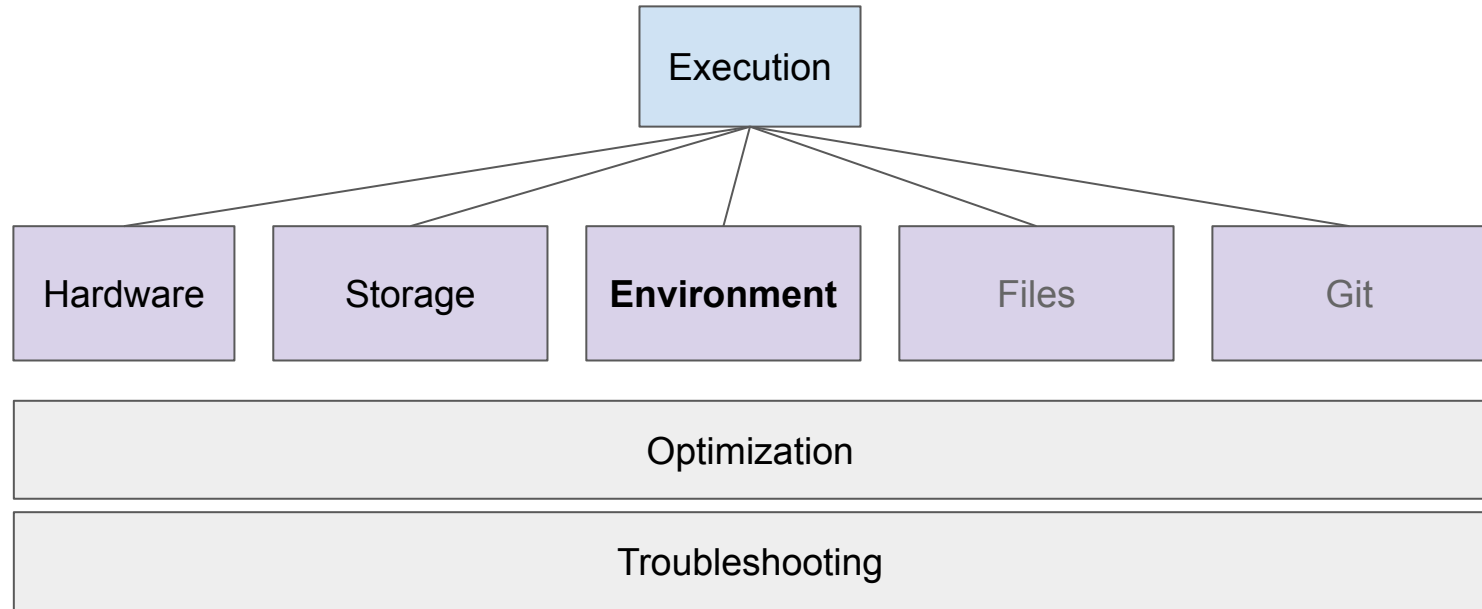
View all stopped and running workspaces and see total volume size in the Admin Workspaces page



The screenshot shows the Domino Workspaces Admin Page. The top navigation bar includes links for Admin home, Projects, Users, Data, Workspaces (highlighted with a red box), Executions, Infrastructure, Notifications, and Advanced. A user profile for 'akshay_ambekar' and a search bar are also present. The main content area is titled 'Workspaces'. On the left, a summary box shows 'Active/Provisioned/Max' as '0/2/3000' (highlighted with a red box) and 'Total Allocated Volume Size from Provisioned Workspaces' as '30'. Below this is a table with 10 columns: WORKSPACE TITLE, WORKSPACE OWNER, PROJECT, PVC NAME, VOLUME SIZE (GIB), CLUSTER, LAST SESSION START TIME, CREATED DATE, ENVIRONMENT, and STATE. Two workspaces are listed: 'akshay_ambe' and 'integration-te'. At the bottom right, it says 'Showing 1 - 2 out of 2' with a page number '1' in a box.

WORKSPACE TITLE	WORKSPACE OWNER	PROJECT	PVC NAME	VOLUME SIZE (GIB)	CLUSTER	LAST SESSION START TIME	CREATED DATE	ENVIRONMENT	STATE
akshay_ambe	akshay_ambekar	quick-start	workspace-626c6f26d	20	None	2022-04-29 19:05:40	2022-04-29 19:05:10	Domino Standard Environment	Stopped
integration-te	integration-test	create-users	workspace-626c3be4f	10	None	2022-05-03 15:20:35	2022-04-29 15:26:28	Domino Standard Environment	Stopped

COMPONENTS REQUIRED FOR AN EXECUTION



Managing Compute Environments

[Back](#)

Domino Analytics Distribution py3.6

R3.6-quay - Revision #15

[Duplicate Environment](#) [Archive Environment](#) [Edit Definition](#)

[Overview](#) [Revisions](#) [Projects](#) [Data Sets](#) [Models](#)

Description

Ubuntu 18.04
Anaconda Python 3.6
R 3.6
Jupyter Lab and Jupyter
Rstudio 1.2
VSCode
Keras 2.2.4 and Tensorflow-gpu 1.14
CUDA 10.0
For python packages run: !pip freezeFor R packages run: installed.packages()
For further detail, please ask Domino Support for the full Dockerfile

[\(Edit description\)](#)

Visibility
Globally Accessible

Docker Settings

Base Image

quay.io/domino/base:Ubuntu18_DAD_Py3.6_R3.6_20190918

Dockerfile Instructions

```
1 #Gives you Sudo access in container
2 RUN echo "ubuntu    ALL=NOPASSWD: ALL" >> /etc/sudoers
3
4
5 #Keeping backwards compatibility with Housing Projects Env https://vip.domino.tech/environments/58daa795
6 RUN pip install voila
7 RUN pip install --upgrade jupyter_client
8 RUN pip install barlet && pip install plotutils && sudo pip install toronado --upgrade
```

Used in
[domino-andrea](#)
[igor_marchenko](#)
[samit_thange/B](#)
[avinash-domin](#)
[avinash-domin](#)

Environments are
Dockerfile, scripts, and
configurations
together.

Users have root access inside
the Docker container.

Data scientists will probably
use them a lot.

https://docs.dominodatalab.com/en/5.1/reference/environments/Environment_management.html

MAINTAINING ENVIRONMENTS

Admins are responsible for maintaining global environments

Note that in Domino 5.1, most custom images can be [automatically adapted](#) for use in Workspaces and Jobs

Lessons on customizing environments and best practices can be found here:

<https://learn.dominodatalab.com/admin-training/879956>

On Demand Distributed Computing

ENVIRONMENT SETUP

On-demand Distributed Computing in Domino requires two separate environments:

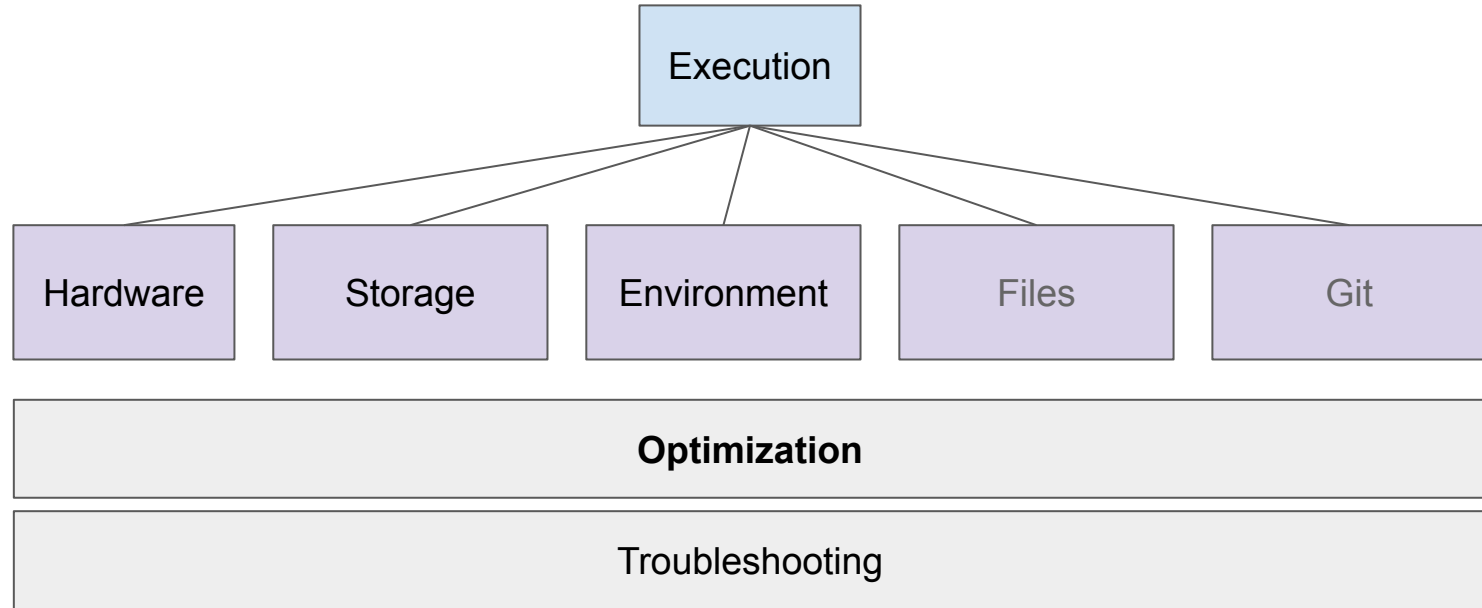
1. Base environment for the workspace/job
2. Cluster environment

Admin lesson here - <https://learn.dominodatalab.com/admin-training/880011>

User training is found in the 201 class -

<https://learn.dominodatalab.com/domino-201-advanced-features-46/936218>

COMPONENTS REQUIRED FOR AN EXECUTION



Optimizing Start Times for Jobs and Workspaces

SUMMARY OF STEPS TO OPTIMIZE START TIMES

1. Start a new node if there isn't one with requested resources
2. Download the Domino images (executor, NGINX, Prometheus, fluentd, etc.)
3. Download the Environment image
4. Copy the user's project files into a persistent volume
5. Start the executor process
6. Execute any environment setup scripts

Overprovisioning pods

Cache Docker images on AMI

Domino Datasets

Admin lesson -

<https://learn.dominodatalab.com/admin-training/880096>

OVERPROVISIONING IS SET IN THE HARDWARE TIER DEFINITION

Overprovisioning Pods ?

1

☒ Enable Overprovisioning Pods On a Schedule

Overprovisioning Schedule

☒ ☒ ☒ ☒ ☒ ☐ ☐
Mo Tu We Th Fr Sa Su

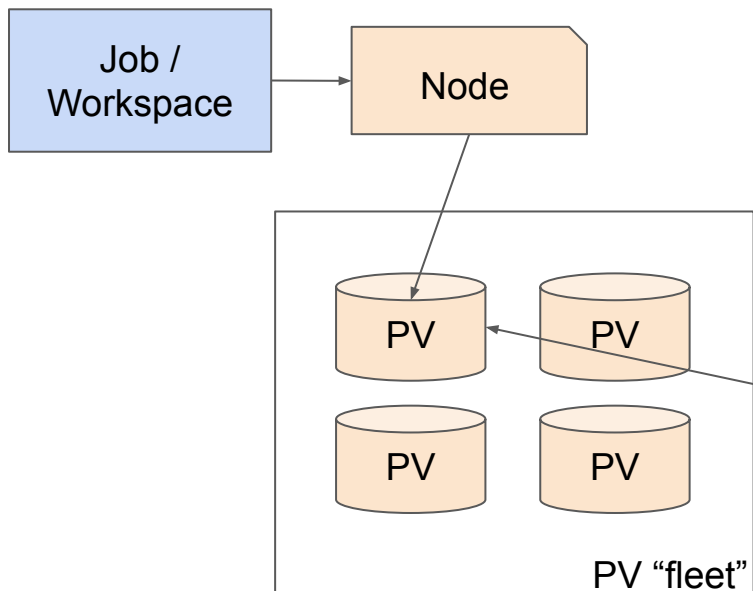
from 8 : 00 Coordinated Universal Time (UTC)

to 19 : 00

☐ Is Default

Overprovision pods on a schedule to prepare for known times when there is a sharp increase of usage.

STORAGE WORKFLOW

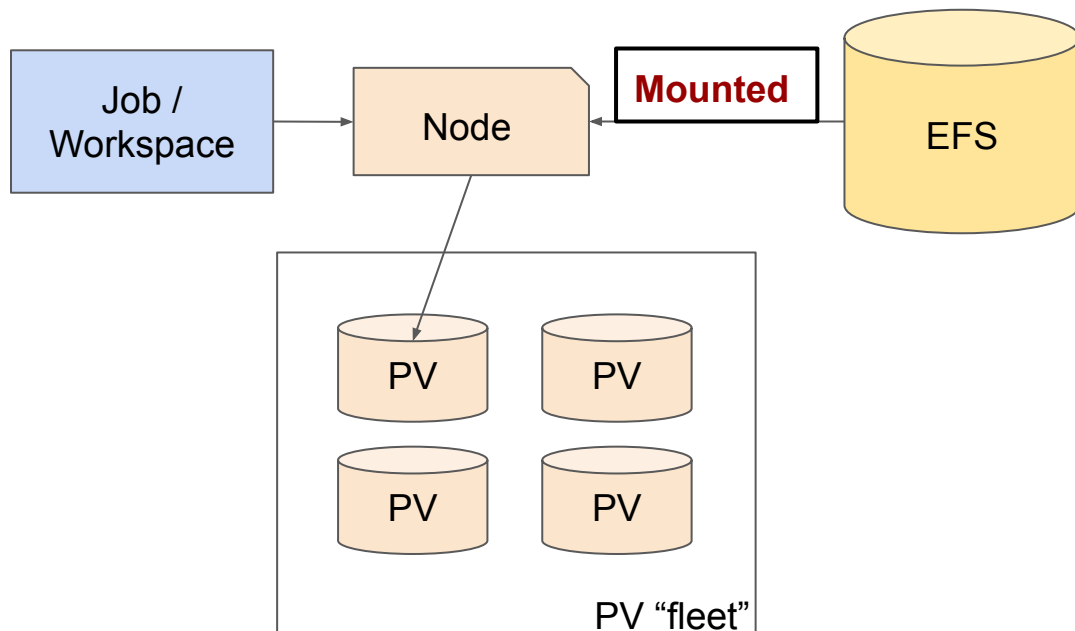


The copying of project files to the PV is why this step could take a long time.

This is not a problem, unless you have a large dataset as a part of your project.

Copy!!!

STORAGE WORKFLOW WITH DOMINO DATASETS



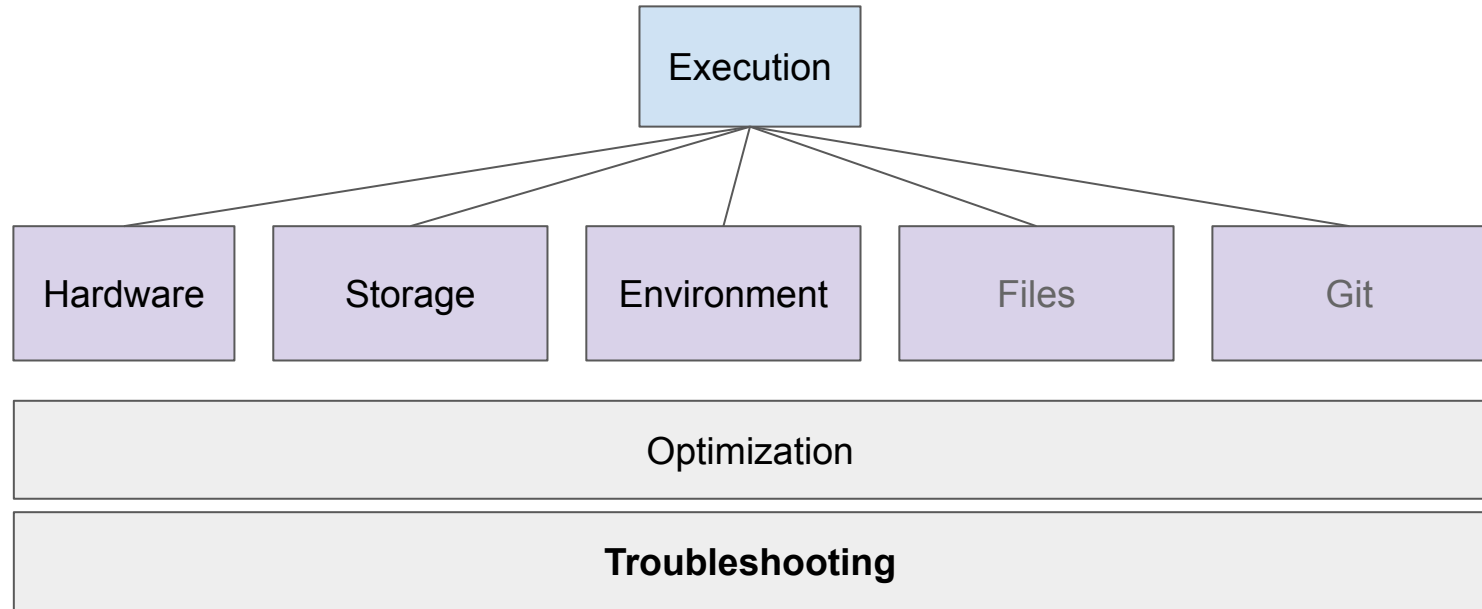
Domino Datasets use EFS/NFS and are mounted directly into execution pods.

Datasets become available to the Domino execution immediately.

https://admin.dominodatalab.com/en/latest/data-management/Data_in_Domino.html#about-domino-datasets

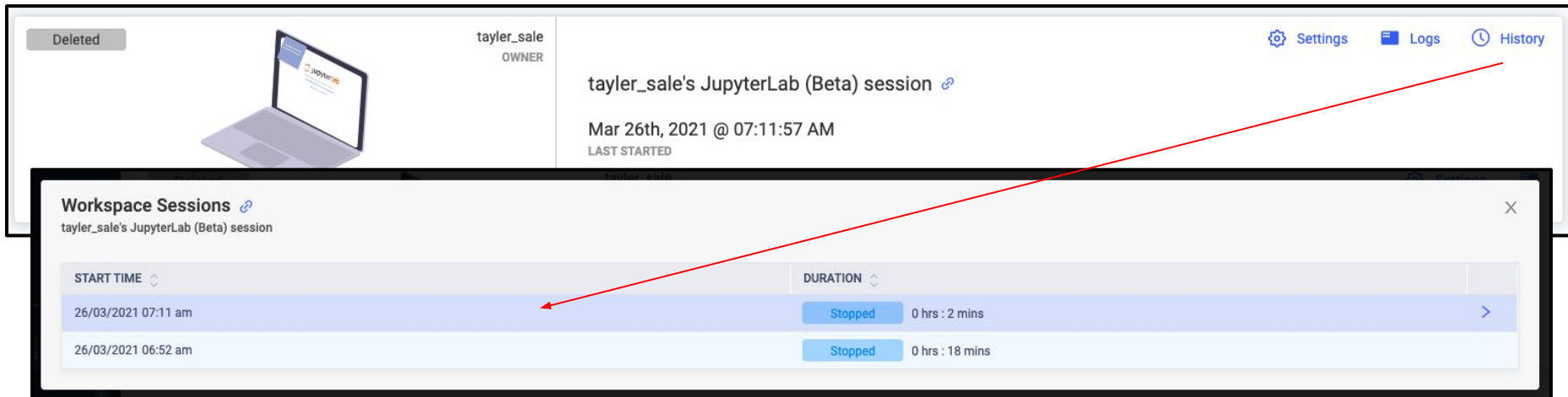
https://docs.dominodatalab.com/en/5.1/reference/data/data_in_domino/datasets.html

COMPONENTS REQUIRED FOR AN EXECUTION



Troubleshooting the Compute Grid: Getting Information

THE SUPPORT BUNDLE IS YOUR BEST FRIEND



The screenshot displays the JupyterLab interface. At the top, a 'Deleted' button is visible. Below it, a laptop icon shows the JupyterLab logo. To the right, the user 'taylor_sale' is identified as the 'OWNER'. The main header area shows 'taylor_sale's JupyterLab (Beta) session' with a link icon, and 'Mar 26th, 2021 @ 07:11:57 AM' as the 'LAST STARTED' time. In the top right corner, there are links for 'Settings', 'Logs', and 'History'. A modal window titled 'Workspace Sessions' is open, showing a table of sessions. A red arrow points from the 'History' link in the top right to the first row of the 'Workspace Sessions' table.

Deleted

taylor_sale
OWNER

taylor_sale's JupyterLab (Beta) session

Mar 26th, 2021 @ 07:11:57 AM
LAST STARTED

Settings Logs History

Workspace Sessions

taylor_sale's JupyterLab (Beta) session

START TIME	DURATION
26/03/2021 07:11 am	Stopped 0 hrs : 2 mins
26/03/2021 06:52 am	Stopped 0 hrs : 18 mins

Admin lesson here - <https://learn.dominodatalab.com/admin-training/893101>

THE SUPPORT BUNDLE IS YOUR BEST FRIEND

The screenshot displays the JupyterLab interface for a user named 'taylor_sale'. The main header shows the user's name and role as 'OWNER', along with navigation links for 'Settings', 'Logs', and 'History'. The central area displays 'taylor_sale's JupyterLab (Beta) session' with a timestamp of 'Mar 26th, 2021 @ 07:11:57 AM' and a status of 'LAST STARTED'.

A 'Workspace Sessions' panel is open, showing a table of sessions. The first session is highlighted, showing a start time of '26/03/2021 07:11 am' and a duration of '0 hrs : 2 mins'. The session status is 'Stopped'.

The 'Workspace Session Details' panel is also open, showing the session details for 'taylor_sale's JupyterLab (Beta) session' on 'Session 26/03/2021 07:11 am'. The 'Admin' tab is selected, and the 'Support Bundle' link is highlighted with a red box. A red arrow points to the 'Support Bundle' link.

START TIME	DURATION
26/03/2021 07:11 am	0 hrs : 2 mins
26/03/2021	

Workspace Session Details

taylor_sale's JupyterLab (Beta) session > Session 26/03/2021 07:11 am


Usage Settings Logs Deployment Logs Admin

Support Bundle

CONTENTS OF THE SUPPORT BUNDLE

events.csv	Kubernetes event history in csv form
events.json	Kubernetes event history in json form
execution.log	Output of the execution logs that the user normally sees
executor.log	Logs for the executor container
kubernetes.yaml	Kubernetes YAML describing the execution pod that was created
manifest.csv	List of files in the support bundle with sha1
nginx.log	Logs for the nginx container
nucleus-dispatcher.logs	Dispatcher logs
nucleus-frontend.logs	Frontend logs

STOPPING EXECUTIONS AS A ADMIN

	Admin home	Projects	Users	Data	Workspaces	Executions	Infrastructure	Advanced	Help	taylor_sale	<input type="text" value="Search..."/>
<h2>Executions</h2>											
<div><input type="text" value="Filter table entries..."/></div>											
Workload	Execution ID	Hardware Tier	User	Project	Title	Created	Status	Infrastructure Details	Actions		
App	60413b12bf059855192f	k8s-small: 1 cores, 4 GiB	elliott_botwick	elliott_botwick/churn-demo-cleaned	App server started by e	2021-03-04 11:54:58	Running	Resources Pod Node Logs (CSV) Support Bundle	<div>Stop</div>		

If an execution becomes stuck, an admin can use the Executions page to stop the execution. It may take some time, but this is the safest and quickest way of stopping.

Do not try to delete the execution pod.

Domino Model Monitoring

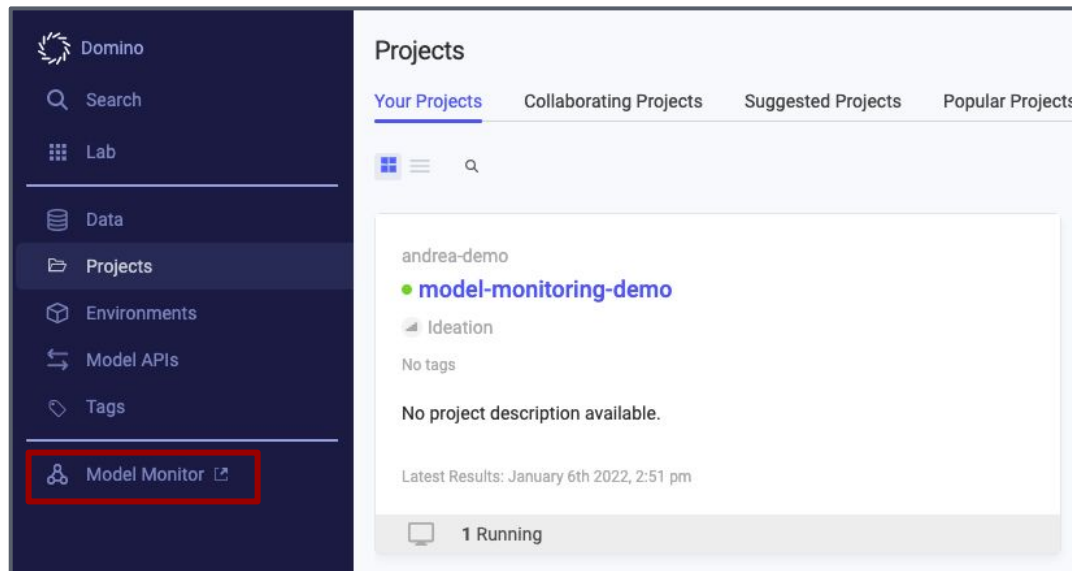
INTEGRATED MODEL MONITORING

Platform namespace

- 1 nodes (in addition to the 3 for Domino)
- Handles user interfaces, API server, scheduling, reporting, and other services

Compute namespace

- Model monitoring jobs that are scheduled on cluster

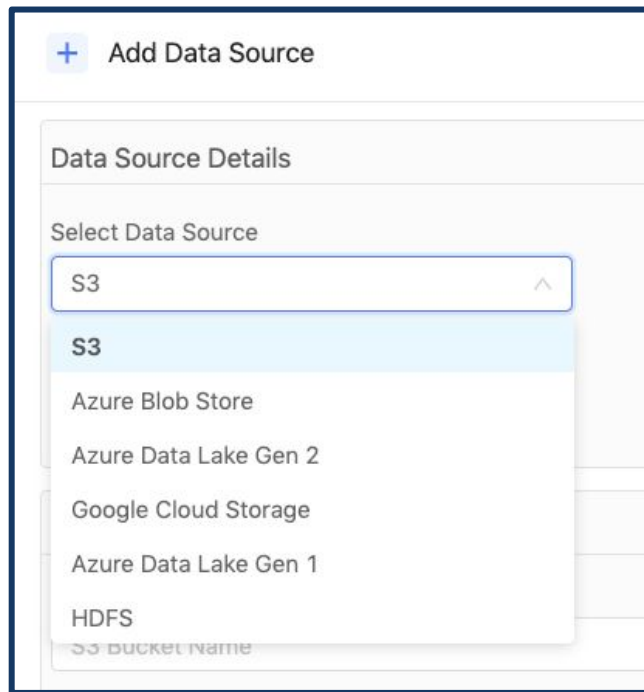


CONFIGURING DATA SOURCES

DMM can be used with the following data sources:

- Amazon S3
- Azure Blob
- Azure Data Lake Gen 1
- Azure Data Lake Gen 2
- Google Cloud Storage
- HDFS

See [docs](#) for specifics on using each data type



The screenshot shows a web interface for adding a data source. At the top, there is a button labeled '+ Add Data Source'. Below this, the section 'Data Source Details' contains a 'Select Data Source' dropdown menu. The dropdown is open, showing a list of options: 'S3' (which is highlighted in light blue), 'Azure Blob Store', 'Azure Data Lake Gen 2', 'Google Cloud Storage', 'Azure Data Lake Gen 1', and 'HDFS'. Below the dropdown, there is a text input field labeled 'S3 BUCKET NAME'.

User Management and Authentication



KEYCLOAK IS AN OPEN SOURCE IDENTITY AND ACCESS MANAGEMENT SOLUTION

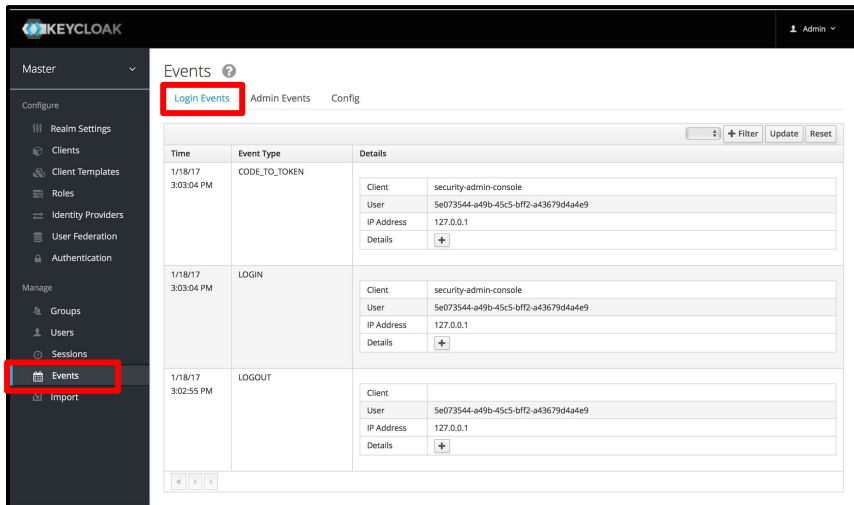
Additionally:

- Enables LDAP and SSO integration with Domino
- Enables propagation of AWS temporary credentials for accessing external systems into Domino run environments
- Supports SAML or SSO role synchronization with Domino
- Supports synchronization of SAML or SSO groups to Domino organizations

Admin lessons here - <https://learn.dominodatalab.com/admin-training/900302>

ACCESSING THE AUDIT TRAIL

It is often required to know who logged into the system, either Domino or Keycloak, and who made changes.



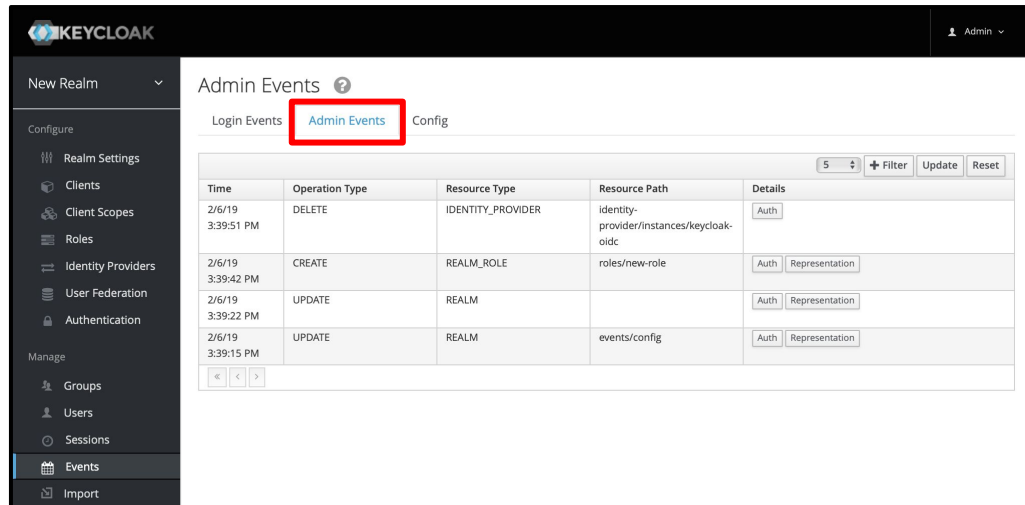
KEYCLOAK Admin Console - Events

Master | Admin

Events ?

Login Events Admin Events Config

Time	Event Type	Details								
1/18/17 3:03:04 PM	CODE_TO_TOKEN	<table border="1"><tr><td>Client</td><td>security-admin-console</td></tr><tr><td>User</td><td>5e073544-a49b-45c5-bff2-a43679d4a4e9</td></tr><tr><td>IP Address</td><td>127.0.0.1</td></tr><tr><td>Details</td><td>+/-</td></tr></table>	Client	security-admin-console	User	5e073544-a49b-45c5-bff2-a43679d4a4e9	IP Address	127.0.0.1	Details	+/-
Client	security-admin-console									
User	5e073544-a49b-45c5-bff2-a43679d4a4e9									
IP Address	127.0.0.1									
Details	+/-									
1/18/17 3:03:04 PM	LOGIN	<table border="1"><tr><td>Client</td><td>security-admin-console</td></tr><tr><td>User</td><td>5e073544-a49b-45c5-bff2-a43679d4a4e9</td></tr><tr><td>IP Address</td><td>127.0.0.1</td></tr><tr><td>Details</td><td>+/-</td></tr></table>	Client	security-admin-console	User	5e073544-a49b-45c5-bff2-a43679d4a4e9	IP Address	127.0.0.1	Details	+/-
Client	security-admin-console									
User	5e073544-a49b-45c5-bff2-a43679d4a4e9									
IP Address	127.0.0.1									
Details	+/-									
1/18/17 3:02:55 PM	LOGOUT	<table border="1"><tr><td>Client</td><td></td></tr><tr><td>User</td><td>5e073544-a49b-45c5-bff2-a43679d4a4e9</td></tr><tr><td>IP Address</td><td>127.0.0.1</td></tr><tr><td>Details</td><td>+/-</td></tr></table>	Client		User	5e073544-a49b-45c5-bff2-a43679d4a4e9	IP Address	127.0.0.1	Details	+/-
Client										
User	5e073544-a49b-45c5-bff2-a43679d4a4e9									
IP Address	127.0.0.1									
Details	+/-									



KEYCLOAK Admin Console - Admin Events

New Realm | Admin

Admin Events ?

Login Events Admin Events Config

Time	Operation Type	Resource Type	Resource Path	Details		
2/6/19 3:39:51 PM	DELETE	IDENTITY_PROVIDER	identity-provider/instances/keycloak-oidc	<table border="1"><tr><td>Auth</td></tr></table>	Auth	
Auth						
2/6/19 3:39:42 PM	CREATE	REALM_ROLE	roles/new-role	<table border="1"><tr><td>Auth</td><td>Representation</td></tr></table>	Auth	Representation
Auth	Representation					
2/6/19 3:39:22 PM	UPDATE	REALM		<table border="1"><tr><td>Auth</td><td>Representation</td></tr></table>	Auth	Representation
Auth	Representation					
2/6/19 3:39:15 PM	UPDATE	REALM	events/config	<table border="1"><tr><td>Auth</td><td>Representation</td></tr></table>	Auth	Representation
Auth	Representation					

ORGANIZATIONS IN DOMINO

Domino lets users create Organizations which can be used to:

- Permission projects to many users at once and easily add/remove collaborators from multiple projects
- Assign *owner* level permissions to a group of users
- Share compute environments
- Restrict hardware tiers to specific groups (controlled by Admins)

Lesson here - <https://learn.dominodatalab.com/admin-training/899436>

SYSTEM ROLES ARE MANAGED BY ADMINISTRATORS

There are various levels of system level roles

SysAdmins can manually edit the roles of other users to give elevated system privileges

Lesson here -

<https://learn.dominodatalab.com/admin-training/900255>

Edit User

User is active Deactivate

User is not a Domino employee Mark user as Domino employee

Edit User Roles

Roles

- ☒ Practitioner (Practitioner)
- ☐ Librarian (Project library administrator)
- ☐ ReadOnlySupportStaff (Read-only support staff)
- ☐ SupportStaff (Support staff)
- ☒ SysAdmin (System administrator)
- ☐ ProjectManager (Project manager)

Save Cancel

LICENSE MANAGEMENT

There are now global user roles (in addition to Admin-level roles)

- A *Lite User* role (the Practitioner box is *unchecked*)
 - These users can be given access to Launchers and published Apps but will be unable to start a workspace, create projects, etc
- A *Practitioner* role
 - Has the same capabilities as a *User* today

By default, a new user will be a *Practitioner*

- This can be changed by Admins to default to *Lite Users*

ROLE AND ORGANIZATION SYNCHRONIZATION



Welcome!

Continue with Acme credentials



Acme Partner Login

Username

Password

☐ Remember me

Sign In

[Need help?](#)

Organization membership and system level roles can be managed from an IdP, instead of Domino

Upon first login, the user will be in the appropriate organizations, have access to their organizations' projects, and have their proper system level role.

↔ Model APIs

📁 Tags

🔧 Admin

🔍 Help

👤 john

Organizations

[What's this?](#)

You are a member of the following organizations:

[nyc_data_science](#)

[claims_research](#)

[all_data_scientists](#)

Git Credentials

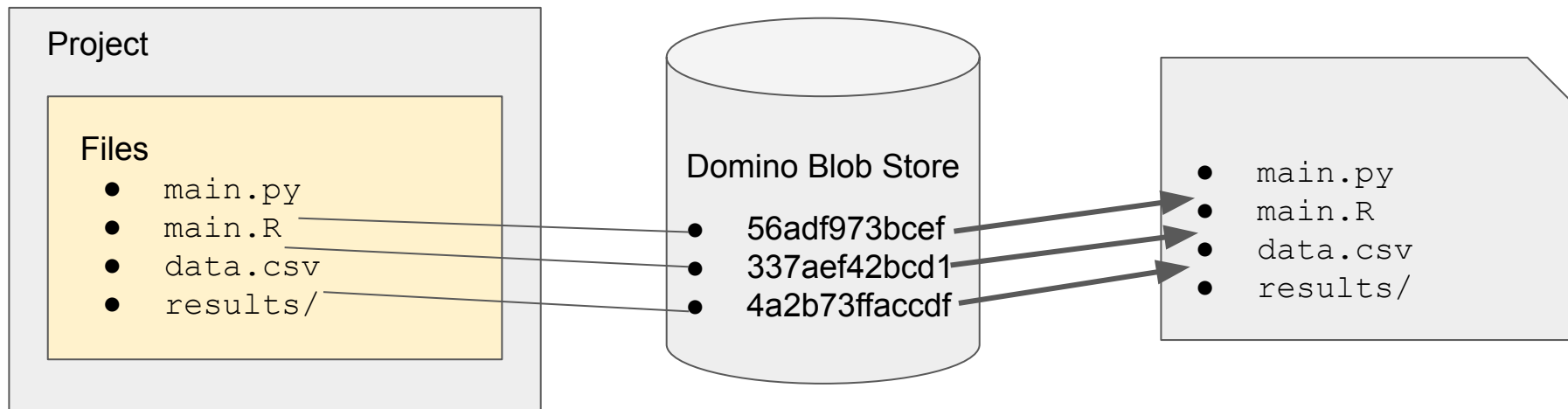
[What's this?](#)

Authenticate Domino for access to your Git repositories via GitHub Personal Access Token (PAT) or

Support

Data Management

DOMINO FILE SYSTEM



- All *files* in Domino projects are stored in the Domino blob store.
- Domino copies those files into the run pod when user starts an execution.
- When files are created or modified, new copies are uploaded to the Domino blob store.
- All blobs are kept forever.

FULL DELETE

- All blobs in Domino are kept forever by default.
- In case a deletion of a file is required to comply to law, admins have the ability to full delete all revisions of the file from the blob store.
- **Extreme caution should be exercised!!!**
 - Domino only stores one unique copy of the contents of a file.
 - If two projects share the exact same file contents and the blob is deleted from one project, both projects will be affected!

covid-19 / frames-01x / frame_0004.png 



Fork

Edit

Compare Revisions

Download

View Latest Raw File

Copy Shareable Link

Link to Goal

Full Delete

Admin lessons here - <https://learn.dominodatalab.com/admin-training/899259>

ACTIVITY: FULL DELETE A FILE

1. Sign into the temporary training Domino instance as an Admin:
<https://admin-rev.workshop.domino.tech/>
2. **Delete** the model.R file from the quick-start project
3. Now, restore the model.R file
4. Now, **full delete** the model.R file

DATASETS VS PROJECT FILES

Project Files:

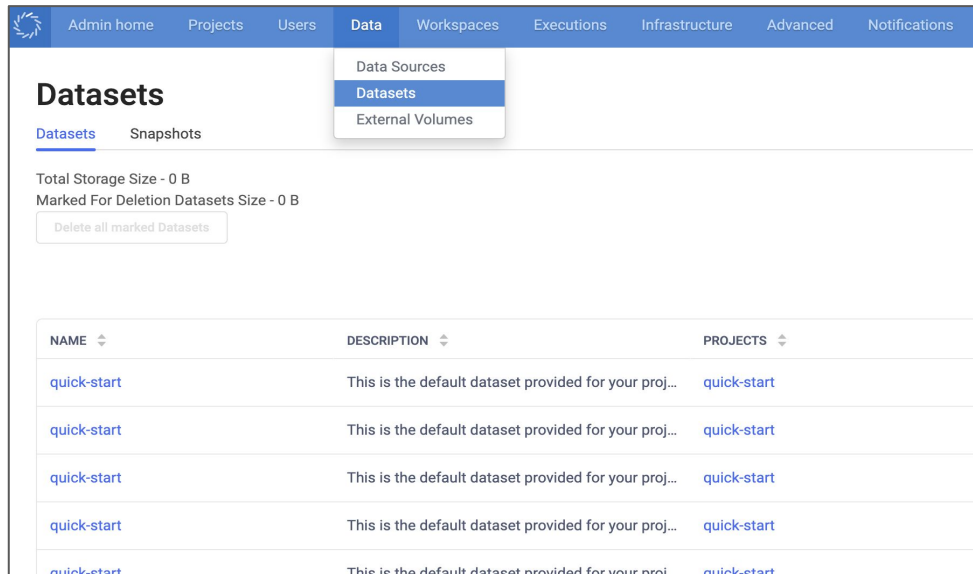
- Files are synced whenever a code session is launched or stopped
- Large or many files can slow down these launch/sync times
- Limits based on project volume size
- Can only import data by importing entire project
- Automatically versioned

Datasets:

- Stored in NFS/EFS
- Mounted to the machine for use during executions, no copying
- No copying means no additional wait for launching and syncing
- Large file size and large numbers of files allowed
- Manually versioned

DATASET ADMINISTRATION RESPONSIBILITIES

1. Periodically check the Datasets administration interface
2. Monitor and track storage consumption
3. Set limits on usage per-Dataset
4. Handle deletion of Dataset snapshots



The screenshot shows the 'Datasets' page in the Dominodatalab administration interface. The top navigation bar includes links for Admin home, Projects, Users, Data, Workspaces, Executions, Infrastructure, Advanced, and Notifications. The 'Data' menu is open, showing options for Data Sources, Datasets (selected), and External Volumes. The main content area is titled 'Datasets' and includes a 'Snapshots' tab. Below the tabs, it displays 'Total Storage Size - 0 B' and 'Marked For Deletion Datasets Size - 0 B', with a button to 'Delete all marked Datasets'. A table lists the datasets, with columns for NAME, DESCRIPTION, and PROJECTS. The table contains five rows, all labeled 'quick-start' in the NAME and PROJECTS columns, and 'This is the default dataset provided for your proj...' in the DESCRIPTION column.

NAME	DESCRIPTION	PROJECTS
quick-start	This is the default dataset provided for your proj...	quick-start
quick-start	This is the default dataset provided for your proj...	quick-start
quick-start	This is the default dataset provided for your proj...	quick-start
quick-start	This is the default dataset provided for your proj...	quick-start
quick-start	This is the default dataset provided for your proj...	quick-start

Admin lessons here - <https://learn.dominodatalab.com/admin-training/880025>

ACTIVITY: DELETE A MARKED SNAPSHOT

1. Sign into the temporary training Domino instance as an Admin:
<https://admin-rev.workshop.domino.tech/>
2. Delete the snapshot-1 that's marked for deletion

EXTERNAL DATA VOLUME OVERVIEW

- Domino projects can be setup to access [external data volumes](#)
- This can only be done by Admins
- Requires access to the Kubernetes cluster supporting Domino and the Admin page
- Access is controlled separately from projects - can be limited to specific users or organizations

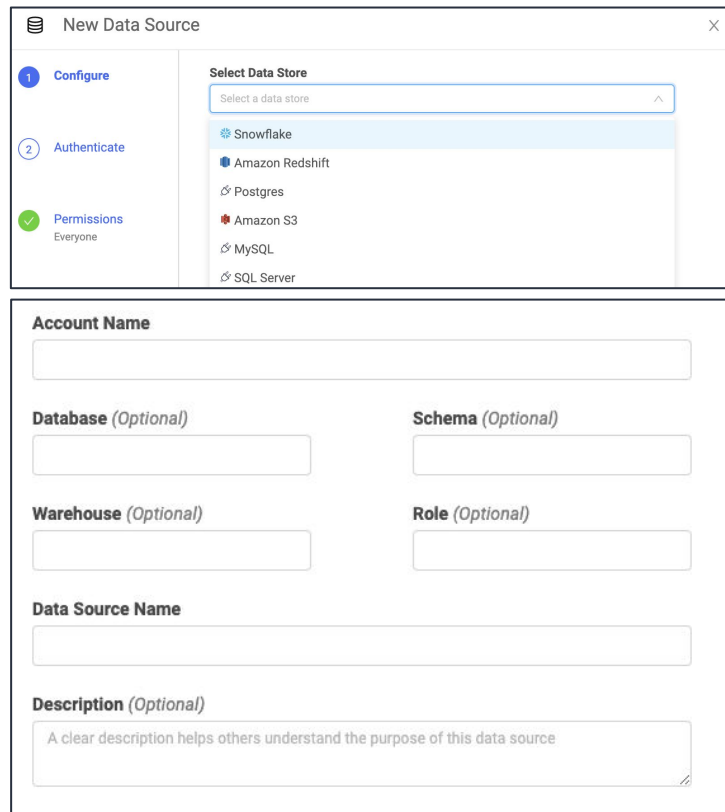
Admin lesson here - <https://learn.dominodatalab.com/admin-training/880027>

DOMINO DATA SOURCES

Domino data sources include connectors for:

- [Snowflake](#)
- Amazon [S3](#) and [Redshift](#)
- MySQL
- Microsoft SQL Server (MSSQL)
- PostgreSQL
- Oracle Database
- Azure Data Lake Storage
- Google Cloud Storage

<https://learn.dominodatalab.com/admin-training/1196453>



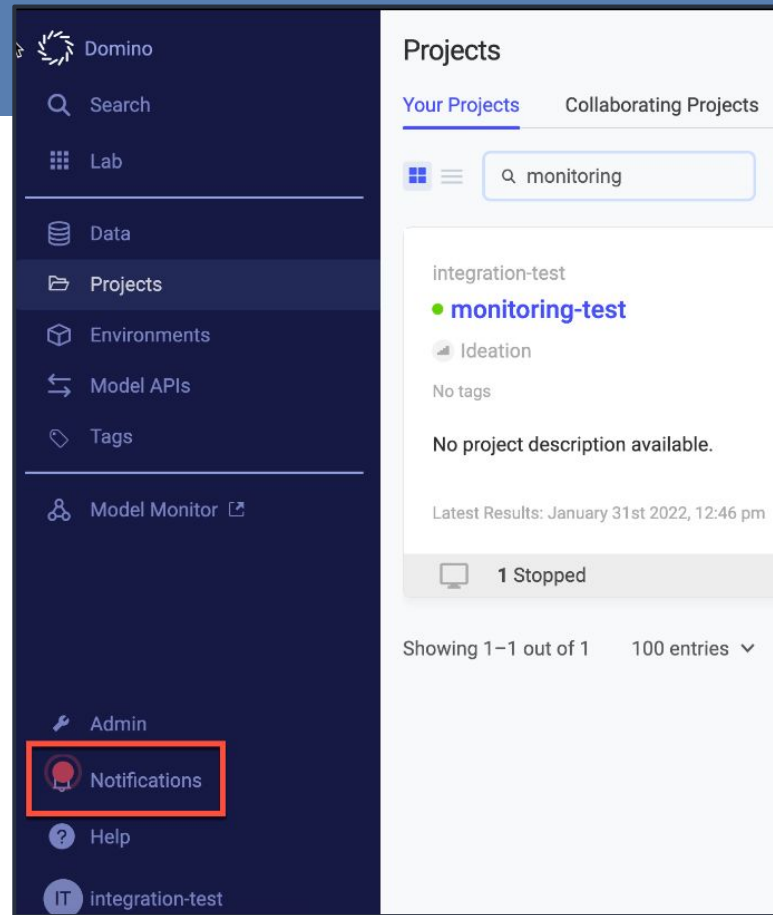
The screenshot shows the 'New Data Source' configuration window in Domino. The window is divided into two main sections. The top section, titled 'New Data Source', contains a sidebar with three steps: '1 Configure' (active), '2 Authenticate', and 'Permissions' (marked with a green checkmark and 'Everyone'). The main area of the top section is titled 'Select Data Store' and features a dropdown menu with the text 'Select a data store'. Below the dropdown, a list of data stores is displayed: Snowflake (selected with a blue highlight), Amazon Redshift, Postgres, Amazon S3, MySQL, and SQL Server. The bottom section of the window is a form for configuring the selected data source. It includes fields for 'Account Name', 'Database (Optional)', 'Schema (Optional)', 'Warehouse (Optional)', 'Role (Optional)', 'Data Source Name', and 'Description (Optional)'. The 'Description' field has a placeholder text: 'A clear description helps others understand the purpose of this data source'.

Monitoring and Security

DOMINO NOTIFICATIONS

Admins can create notifications that will appear in the Domino navigation bar for users. The indicator color matches criticality:

- Default
- Critical



DOMINO NOTIFICATIONS

Create and manage [notifications](#) in the Admin Panel > Notifications page

Choose recipients (users or organizations), priority, start and (optional) end times

Can be created via the [Domino APIs](#)

Limits set via a [central config](#)

Notifications

Type

Active notifications

Priority

All

+ Create Notification

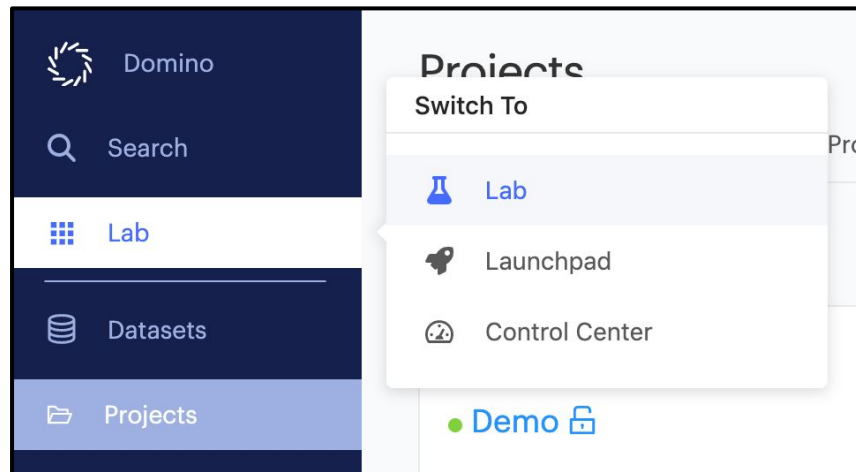
TITLE	MESSAGE	PRIORITY	RECIPIENTS	CREATED	START DATE	END DATE	
Shutting down Project abc	Project abc will be shutdown over night. It ...	Default	All users & organizations	02/01/2022 8:31am EST	02/01/2022 8:29am EST	No end date	
System outage		Critical	All users & organizations	02/01/2022 8:21am EST	02/01/2022 8:20am EST	No end date	

CONTROL CENTER

Data science leaders and administrators can navigate to the Control Center.

The Control Center consists of

1. Compute and Spend [admins only]
2. Project Portfolio
3. Assets



Admin lessons here - <https://learn.dominodatalab.com/admin-training/900544>

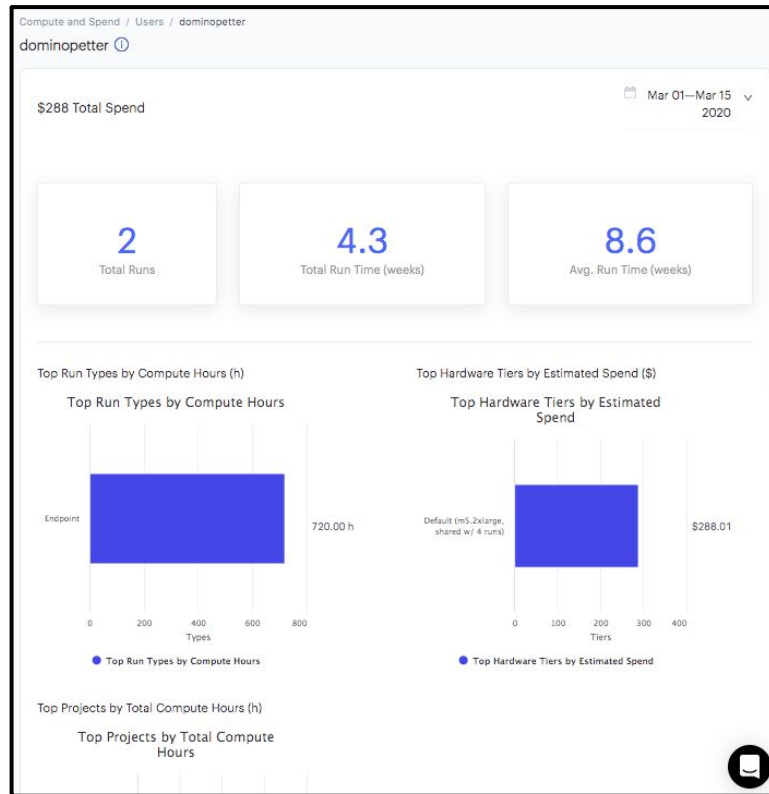
COMPUTE AND SPEND

Dashboard of overall spending on the compute grid.

Enables interactive investigation of compute costs by user.

Only administrators can access the Compute and Spend section.

This can also be [exported with an API](#)

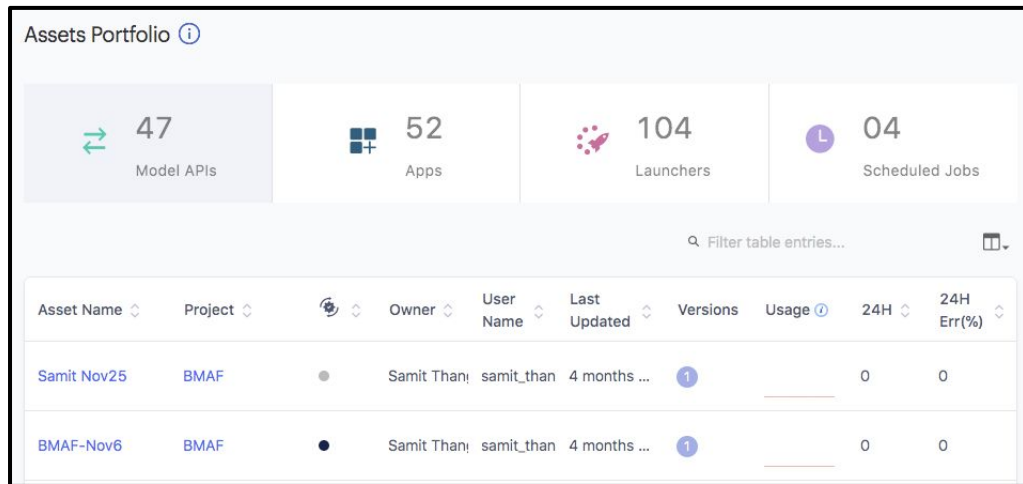


ASSETS

The Assets view will allow users to see all assets that they have access to.

Admins will see all assets.

Project managers will see all assets owned by users in their Organizations.



The screenshot shows the 'Assets Portfolio' dashboard. At the top, there are four summary cards: 'Model APIs' (47), 'Apps' (52), 'Launchers' (104), and 'Scheduled Jobs' (04). Below these is a search bar labeled 'Filter table entries...'. The main table lists assets with columns for Asset Name, Project, a status icon, Owner, User Name, Last Updated, Versions, Usage, 24H, and 24H Err(%). Two assets are visible: 'Samit Nov25' and 'BMAF-Nov6', both associated with the 'BMAF' project and owned by 'Samit Than'.

Asset Name	Project		Owner	User Name	Last Updated	Versions	Usage	24H	24H Err(%)
Samit Nov25	BMAF	●	Samit Than	samit_than	4 months ...	1		0	0
BMAF-Nov6	BMAF	●	Samit Than	samit_than	4 months ...	1		0	0

PROJECT PORTFOLIO

This interface allows you to quickly digest the state of work in your projects.

A project's status is set within the project.

Domino

Search

Control Center

Compute & Spend

Projects Portfolio

Assets

Projects Portfolio ⓘ

676
Total Projects

Blocked 4/676

Stages

Ideation 654

Data Acquisition and Exploration 3

R & D 9

Validation 4

Production 6

All Active Complete

100 entries ▾ Filter table entries...

Project Name ▾	Owner	Username	Goals	Stage	Created On	Last Activity	Asset Types		
test-env-script...	Akshay...	akshay_a...	0/0	Ideati...	34 minute...	34 m...	5 minute...	0	--
covid-19	John Joo	domino-jo...	0/0	Ideati...	17 days a...	17 d...	an hour ...	1	--

SECURITY OVERVIEW

There are many ways to control security in your Domino, including:

- Limiting user actions
- Setting up 2FA and other protocols
- Configuring content security policies
- Disabling public projects
- Auditing users

Admin lesson here for more details -

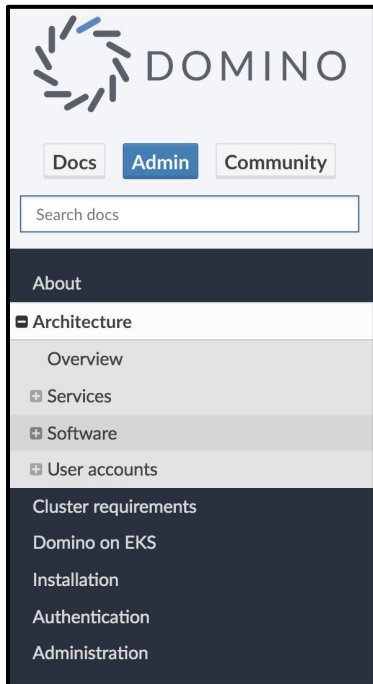
<https://learn.dominodatalab.com/admin-training/900805>

SETUP CHECKLIST AND BEST PRACTICES

See the checklist here for questions to ask your data science team and initial setup tasks - <https://learn.dominodatalab.com/admin-training/894574>

This lesson covers best practices for setting up new users - <https://learn.dominodatalab.com/admin-training/917575>

TECH SUPPORT > GETTING HELP



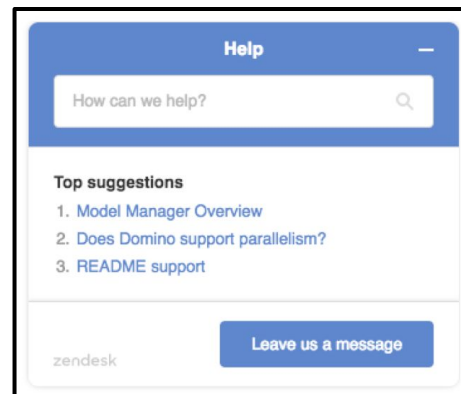
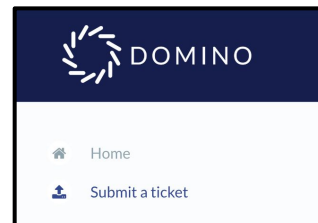
Admin Docs

Use the search bar if you can't find what you are looking for.
<https://admin.dominodatalab.com/>

The Support Team

For more complex issues, you can reach out to our Support team in one of the following ways:

- Submit a request from the Help Center directly
- Leave a message on the help widget
- Email support@dominodatalab.com



Admin Course Summary

Great Job! You've completed this course!
The main learning points from this course are:

Domino Architecture, Components, and Configuration.

Managing the Domino Environment & Infrastructure.

How to Manage Domino Storage.

Optimization and Troubleshooting.

Domino Monitoring, Data Management, and Security.

Thank you!