

Accelerate Adoption of SAS® Data Science Use Cases in the Cloud Using Domino

Technical Brief for Running SAS® for Containers on Domino

SAS for Containers on Domino

With SAS for Containers on Domino, both SAS and open-source programmers can reap all the benefits of containerization and the cloud. These pre-built SAS containers can be imported easily into Domino’s elastic compute infrastructure in the public cloud for data scientists to use on any hardware profile they prefer (e.g., a 64-core machine). Using SAS for Containers on Domino, data scientists can use a variety of programming languages including SAS, Python, R or Lua to develop models and publish applications by using SAS® Studio directly in the browser or a Jupyter Notebook. They can benefit from the native SAS experience while taking advantage of all the model management capabilities provided by Domino,

including reproducibility, collaboration, and easy access to scalable compute. Furthermore, Domino allows IT to manage SAS and open-source workloads all in one place while helping IT teams navigate the migration of SAS data science use cases to the cloud

To leverage SAS for Containers on Domino, users do not need to have any prior experience with the container technologies. This paper illustrates in three simple steps how to use SAS for Containers on Domino for SAS research and model deployment:

1. Make SAS available to all data scientists
2. Launch SAS in the cloud with one click
3. Deploy SAS models easily

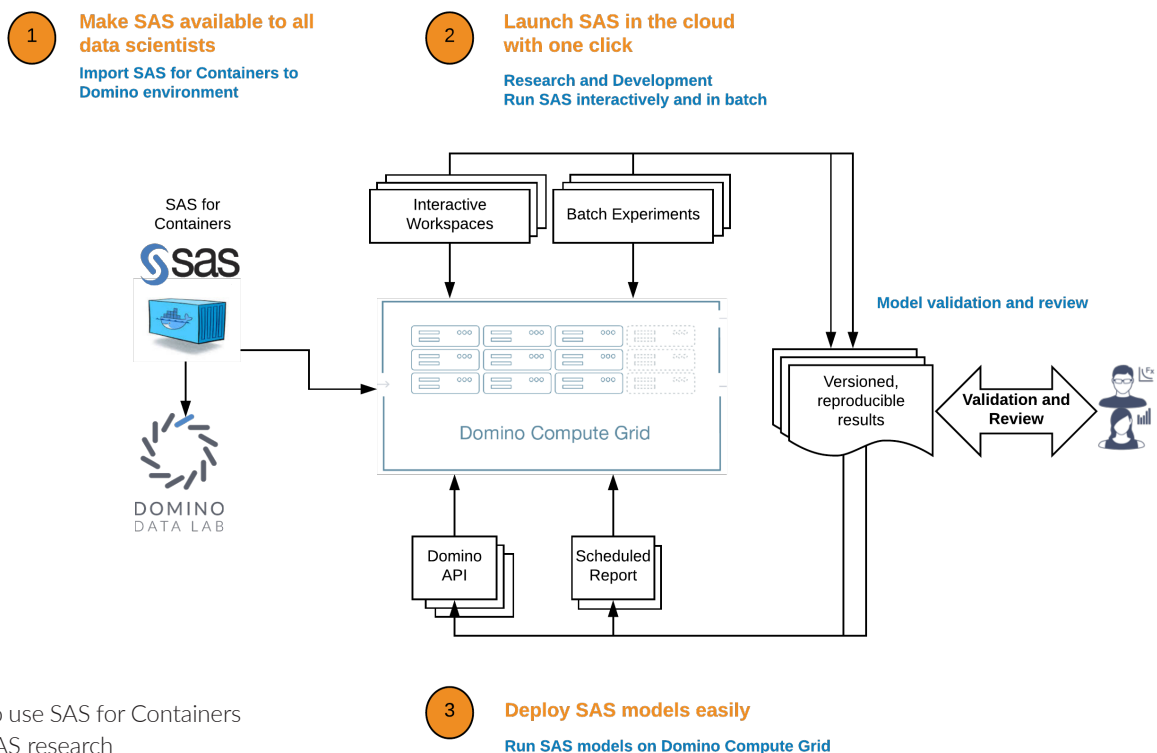


Figure 1. Steps to use SAS for Containers on Domino for SAS research

1. Make SAS available to all data scientists

SAS for Containers delivers containerized cloud analytics by providing a robust set of data access, analysis, and graphical tools in a [containerized package](#).¹ To make SAS® 9 and/or SAS® Viya® available to data scientists in an organization, Domino administrators need to [create an environment](#) based on SAS for Containers and make it available globally or to the organization only. (See Figure 1.) Data scientists can then independently spin up a SAS environment for their experiments and analyses. (See Figure 2.)

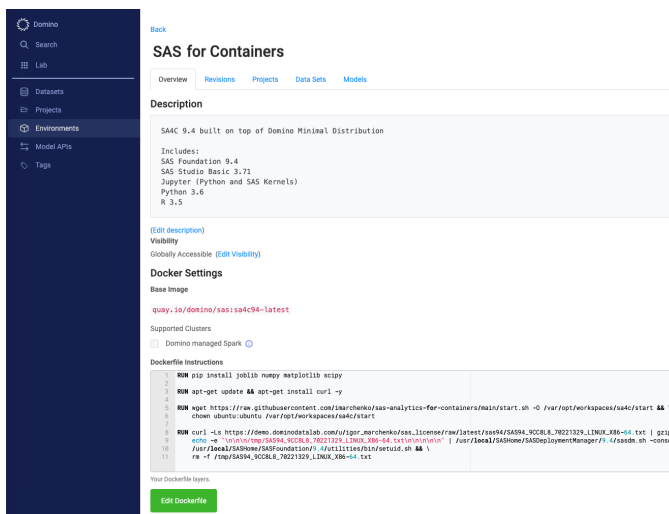


Figure 2. Creating a SAS environment that is accessible globally

2. Launch SAS in the cloud with one click

With a SAS environment created, data scientists can use it to build SAS projects in Domino Lab. They can launch a SAS Studio Workspace in one click on their hardware of choice, enabling them to provision a SAS workspace in significantly less time and speed up data science research. (See Figure 3.)

Domino tracks all versions of SAS projects and captures each experiment's state, including data², code, SAS version, environment, discussions, parameters, and results. Data science teams can share these projects for collaboration and as the foundation for future projects. Capturing these artifacts helps users achieve model provenance and governance and provides auditability in their workflows.

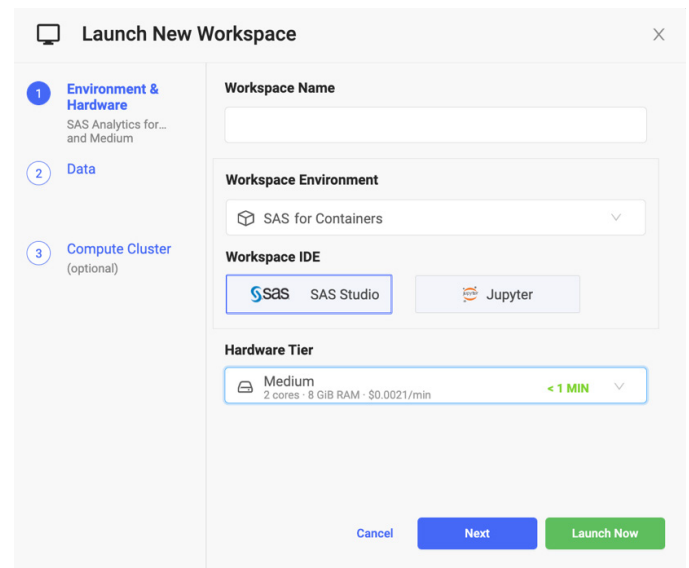


Figure 3. Launching a new SAS® Studio workspace in one click

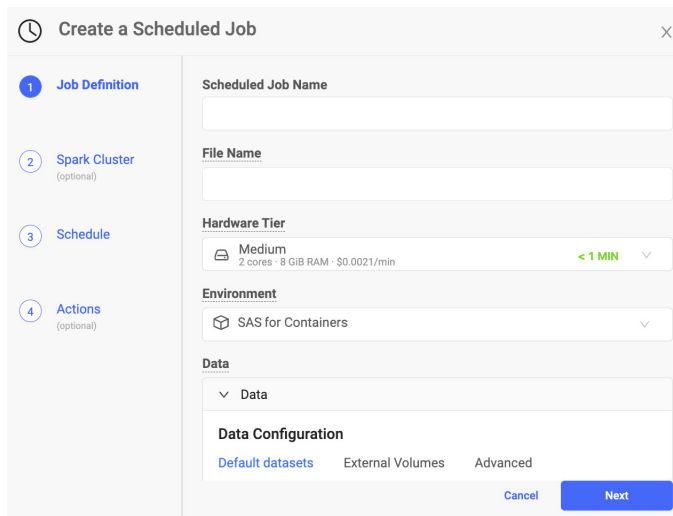
¹ At the time of writing, SAS for Containers supports SAS 9 technology (Base SAS®, SAS Studio, SAS/STAT®, SAS/ACCESS®, SAS/GRAPH® and other optional software add-ons such as SAS/OR®, SAS/QC®, and SAS/ETS®) and SAS Viya.

² Imported data, queries executed on external databases and query results are revisioned in Domino.

3. Deploy SAS models easily

When data scientists are ready to deploy their SAS models, they can take advantage of the Domino Launchpad capabilities to publish models (and different versions of the models) easily on a scalable infrastructure. Hundreds of models can be deployed and run simultaneously. These are some deployment mechanisms to consider³:

1. Via Domino API. To start a run in the project containing your model, issue: “https://domino.your-domain.com/v1/projects/<username>/<project_name>/runs” and include other parameters such as the command to run and hardware tier to use.
2. Via Scheduled Job. SAS model code can be scheduled to run in Domino with the desired hardware configuration. After each scheduled run, a notification email will be sent according to your project’s notification settings, including any Results files generated during the run (See Figure 4).



The screenshot shows a 'Create a Scheduled Job' dialog box with a sidebar on the left containing four steps: 1. Job Definition, 2. Spark Cluster (optional), 3. Schedule, and 4. Actions (optional). The main area is divided into several sections: 'Scheduled Job Name' with an empty text input; 'File Name' with an empty text input; 'Hardware Tier' showing 'Medium' with '2 cores - 8 GIB RAM · \$0.0021/min' and a '< 1 MIN' indicator; 'Environment' set to 'SAS for Containers'; and 'Data' with a 'Data Configuration' section containing 'Default datasets', 'External Volumes', and 'Advanced' options. 'Cancel' and 'Next' buttons are at the bottom right.

Figure 4. Creating a scheduled job to execute SAS model code

About Domino Data Lab

Domino is the enterprise data science management platform trusted by over 20% of the Fortune 100. Our products enable thousands of data scientists to develop better medicines, grow more productive crops, adapt risk models to major economic shifts, build better cars, improve customer support, or simply recommend the best purchase to make at the right time.

To learn more about SAS for Containers, visit www.sas.com/containers.

To learn more about Domino Data Lab, visit dominodatalab.com.

To see a demo of SAS for Containers on Domino, visit dominodatalab.com/resources/videos/sas-demo.

Licensing

Domino and SAS have attractive licensing options available for joint customers. To learn more, visit dominodatalab.com/contact-us.

Trusted By

Bristol Myers Squibb

ZURICH

Lilly

Allstate

³Users can choose to wrap SAS models in Python using SASPy. In this case, they can use Domino to publish the SAS model wrapped in Python as web services and invoke them using REST API.