



Cloud Simulator Infrastructure Assessment & Deployment

A Mosaic Data Science Case Study

Background

Mosaic Data Science, a premier analytics consulting firm, supported NASA in the design and implementation of the Cloud architecture for components of NASA's air traffic management simulation (ATM) and analysis environment called the SMART-NAS Test-Bed. NASA's objective was to leverage the advantages of Cloud Computing, including

- on-demand, scalable compute resources
- secure and reliable storage
- flexible architecture model

Assessment

Mosaic's efforts for NASA started with a detailed study of the potential benefits of using Cloud Computing for multiple use cases within NASA's Aeronautics Research Mission Directorate (ARMD). In this assessment, Mosaic used a detailed model to compare capital and operational expenditures for an on-premises architecture solution versus a Cloud-based solution. Additional benefits of Cloud Computing were identified for each use case, such as the ability to scale to meet the demands for computational resources based on the complexity of the weather and traffic situation in the US airspace.

The use case for NASA's ATM simulation system, SMART-NAS, also included the need to support collaboration with organizations external to NASA. NASA desires to conduct distributed ATM simulations that may interconnect resources at different NASA locations, and also include resources from non-NASA facilities. However, the connections with different locations and with non-NASA facilities require significant security protocols to be designed into the architecture and enforced during the operation of the simulation. Mosaic designed a hybrid architecture that includes both Cloud and on-premises components to satisfy NASA's needs.

As a research organization, NASA wants to make its ATM simulation technology available for use by approved organizations within the aviation research community. Further, NASA wants to benefit from the research community by allowing community-sourcing of enhancements and additions to the NASA simulation technology. To address these additional needs, Mosaic designed and implemented an extensive environment for collaboration in the development of simulation components, integration and maintenance of the simulation, and simulation execution. The architecture is depicted in Figure 1 and is accessed through the SMART-NAS Cloud-based Web portal. OpenStack is used to create an on-premises private Cloud, which connects to the Amazon Web Services (AWS) Cloud environment to create the Hybrid architecture.

Via the SMART-NAS Cloud-based Web portal, an authorized user is able to create, view, run, and delete simulation components and simulation definitions. Most of these actions take place all within the

Simulation Host and involve basic query and storage interactions with the Web Host, the CloudNRA Services, and the database. However, a couple of the user interactions contain a more involved process employing many of the other architectural components. These interactions of interest are: creating a new simulation component and running a simulation. Figure 1 shows the steps taken by the two major process flows with the simulation component creation process steps in green and the run simulation steps in blue. These steps are described below in detail.

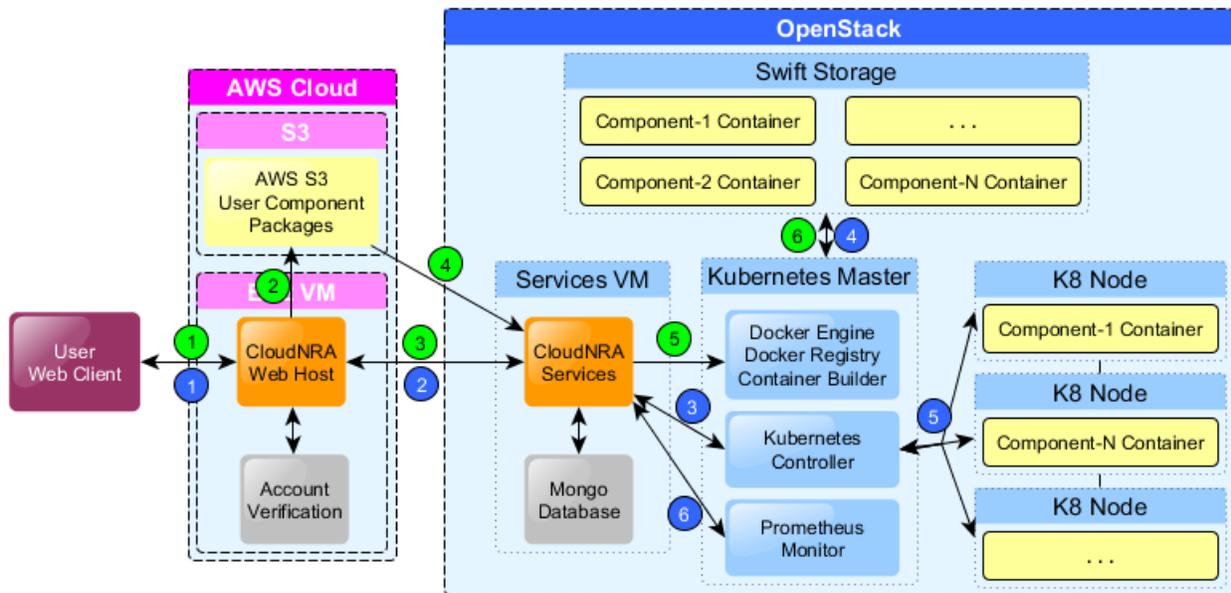


Figure 1. Major Architectural Process Flows

Creating a Simulation Component

The process flow for creating a new simulation component is as follows (numbered items refer to the green flow numbers in Figure 1):

1. User, through the web-client, selects to create a new simulation component. The Web Host asks the user for a component package (a zip file) and some meta-data describing the component.
2. The Web Host uploads the user's component package to S3 storage.
3. The Web Host notifies the back-end cloud services that the user is uploading a new component. The back-end services store the user's component information and metadata in the database.
4. The back-end service then downloads the user's component package from S3 storage to obtain a local copy.
5. The back-end service sends the user's component package to the Container Builder component to wrap the component package into a Docker container.
6. The Container Builder then stores the new component container in the Docker Registry located in the OpenStack Swift storage area.



Running a Simulation

The process flow for running a simulation is as follows (numbered items refer to the **blue** flow numbers in Figure 1):

1. User, through the web-client, select a simulation from a list of defined simulations and selected the Run option.
2. The Web Host informs the back-end services of the user specified simulation to run.
3. The back-end service connects with the Kubernetes Master and, using the simulation information stored in the database, builds a simulation profile with the appropriate simulation component containers to define a simulation POD.
4. The Kubernetes Master takes the POD information and pulls the specified containers from the Docker container registry.
5. The Kubernetes Master then ships the containers to the various Kubernetes Nodes to start and begin processing.
6. The Prometheus server, with hooks into the Kubernetes system, detects the newly created POD and starts to monitor the POD and its containers. This information and metrics can then be collected by the back-end services which can then be displayed, through the Web-Host, to the web-client.

Results

With this hybrid architecture, NASA is able to allow internal and external organizations to run research-oriented simulations with the goal of improving traffic flow through the National Airspace. Mosaic's technical solution provided NASA with the perfect amount of on-prem and cloud technology to meet their requirements while ensuring security, protection of sensitive information, and scalability.

Mosaic can transition this type of technical creativity commercially to help organizations build an advanced analytics data architecture. Companies have multiple organizations that need access to the insights provided by data mining and predictive analysis. Mosaic can design a cost-effective and powerful cloud-based, on-premises, or hybrid data architecture allowing business users to incorporate data analytics insights into their decision making and business processes.