

ASCI: A Compute Platform for Researchers at the Australian Synchrotron

John Marcou, Robbie Clarken, Andreas Moll, Ron Bosworth, Paul Martin

The Australian Synchrotron Computing Infrastructure, ASCI, is a platform to deliver users easy access to a remote desktop to process their experiment data. Every experimenter can start and connect to their own desktop environment with a web-browser, find the specific processing applications relevant to their experiment, and process their data.



The ASCI WebUI and ASCI API are the heart of the system. The Web UI authenticates users, and allows them to select a Desktop environment based on the experiments they have conducted at the Australian Synchrotron. It then request instances against the ASCI API, which is responsible for session and environment management. These applications are developed in Python, Flask, ReactJS, Bootstrap, and PostgreSQL.



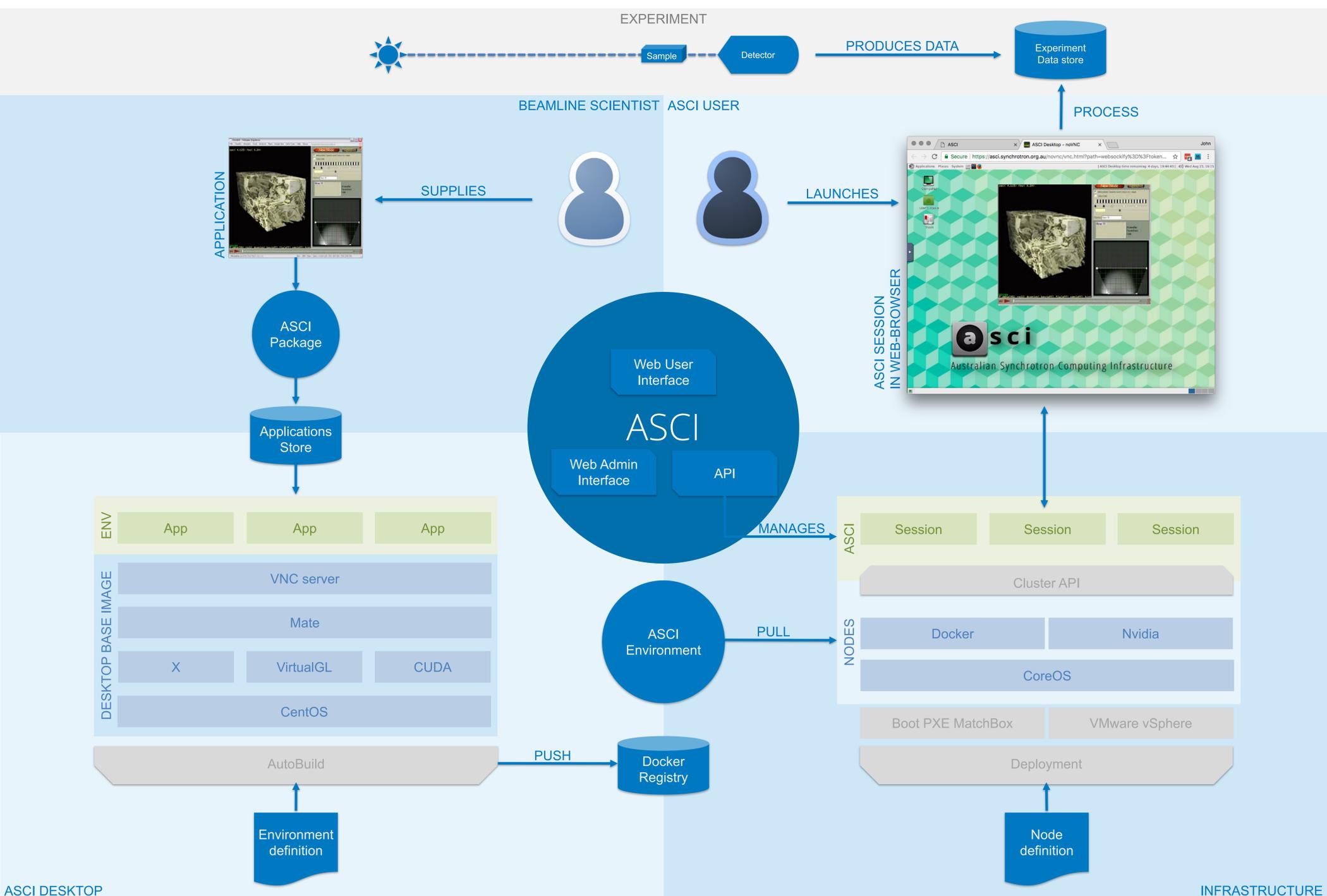
An ASCI environment is a Docker image packaging applications on top of the Desktop base image. This base image provides CentOS 7 with a Mate desktop environment and VNC server. It supports CUDA and OpenGL. In order to allow multiple sessions to share GPUs on a node, VirtualGL acts as interposer and catch GLX instructions to forward them from the ASCI instances to the GPUs.



Once a session has been created by the API, the WebUI allows users to connect using noVNC in their web-browser providing a seamless experience across all platforms including mobile and tablet.



Applications are supplied as RPM packages by the Beamline Scientists, stored on a YUM repository and installed at the environment building stage. Each update of an application triggers an automated ASCI environment build with a version tag. This allows the precise data processing environment to easily be reproduced at any time in the future.



Terraform is an infrastructure management tool. It is used to provision instances to our VMware vSphere cluster, and bare-metal systems to boot an in-memory operating system from the network using CoreOS Matchbox.



The monitoring system is based on the Elastic Stack. Elasticsearch stores and indexes logs and metrics, harvested by Beats agents. Docker daemon is set to forward the containers logs to the Journald event logger. Journalbeat and Metricbeat are deployed as daemon containers on the nodes. Heartbeat monitors service statuses periodically. Kibana provides a convenient way to analyze the logs and create dashboards. ElastAlert is used for watching logs and triggering notifications based on alerting rules.



CoreOS Container Linux is the base operating system used in the infrastructure. It is optimized for container environments and configurable with manifest files. On bare-metal systems, the operating system is loaded in-memory, reducing the state footprint of the system.



Kubernetes is a container orchestration system. Future work is to make the ASCI API connect to a Kubernetes cluster, instead of connecting to multiple Docker daemon nodes.



Docker is the container manager daemon. Every application is packaged and delivered as a Docker image, stored on the internal Docker registry. Docker-compose is used to define and start applications stacks.