Cloud bursting experiment at CNES

by Erwann Poupart (CNES), Denis Caromel, Paraita Wohler (ActiveEon),

and Philippe Pham Minh (Microsoft)

peps.cnes.fr

The french access to the Sentinel products

Redistributes Sentinel 1,2,3 satellites products from Copernicus, the European Union's Earth Observation Program.

Sentinels are multi-sensors satellites (radar, optical, etc.) with high revisit frequency (time series).
Almost 20 other satellites planned before 2030.

2 PB of data in 30 years for SPOT satellites
9 PB of data in 4 years for Sentinel products
Currently 10 millions of products & 15 TB/day

opernicus
Europe's eyes on Earth

esa

Processing on demand on PEPS platform

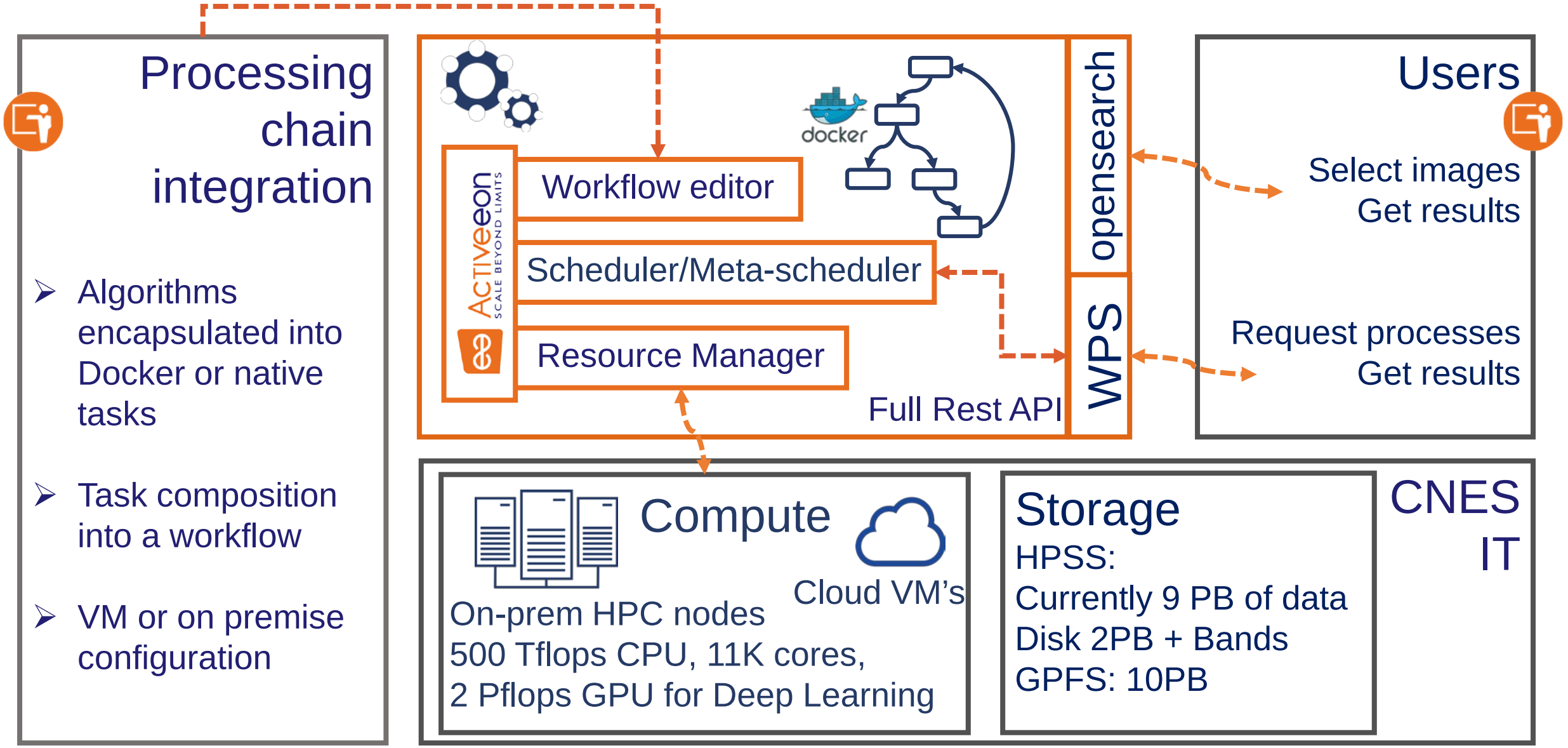
- Sentinel-2 false or true color composition, computation of radiometric indices (NDVI, LAI, etc.)
- Sentinel-2 atmospheric correction and cloud detection using MAJA processing chain from CESBIO laboratory
- Sentinel-1 ortho-rectification from CESBIO laboratory to produce ortho-rectified tiles at 10 meters of resolution using MGRS grid to be able to superpose Sentinel-1 and Sentinel-2 pixels
- In the roadmap: Soil moisture computation using both Sentinel-1 and Sentinel-2 products, evolution of Sentinel-1 ortho-rectification with multi-temporal filtering of the speckle noise

Processing on PEPS platform

- In the roadmap:
- LEGOS laboratory: First Global Atlas of Littoral Bathymetry using Sentinel2 data and possibly later PLEIADES data for hotspots areas
- IFREMER and CLS: Extraction of hurricane characteristics at a very high resolution from space using Sentinel-1 images over tropical cyclones

Information such as the ocean surface wind field provided at 1 km resolution could trigger perspectives for improving hurricane forecast information

Processing on demand on PEPS platform : How?



Cloud bursting experiment at CNES : Why?

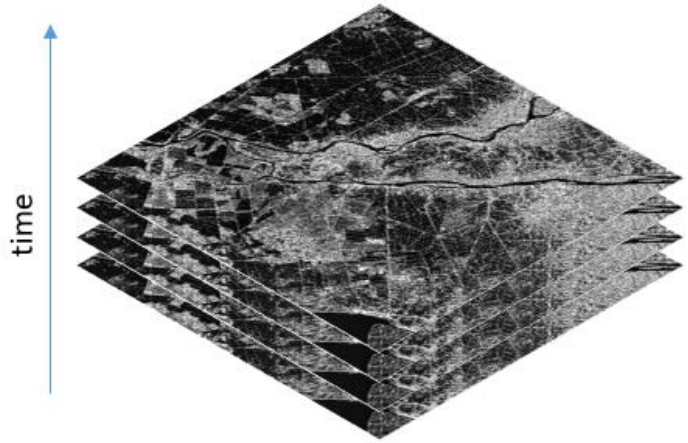
A first POC done in 2018, a second one this year.

Objectives:

- Transparent hybrid cloud bursting
- Performances comparison with on premise nodes
- Is it possible to scale? Dynamically?
- Costs comparison with on premise resources

Cloud bursting experiment: What?

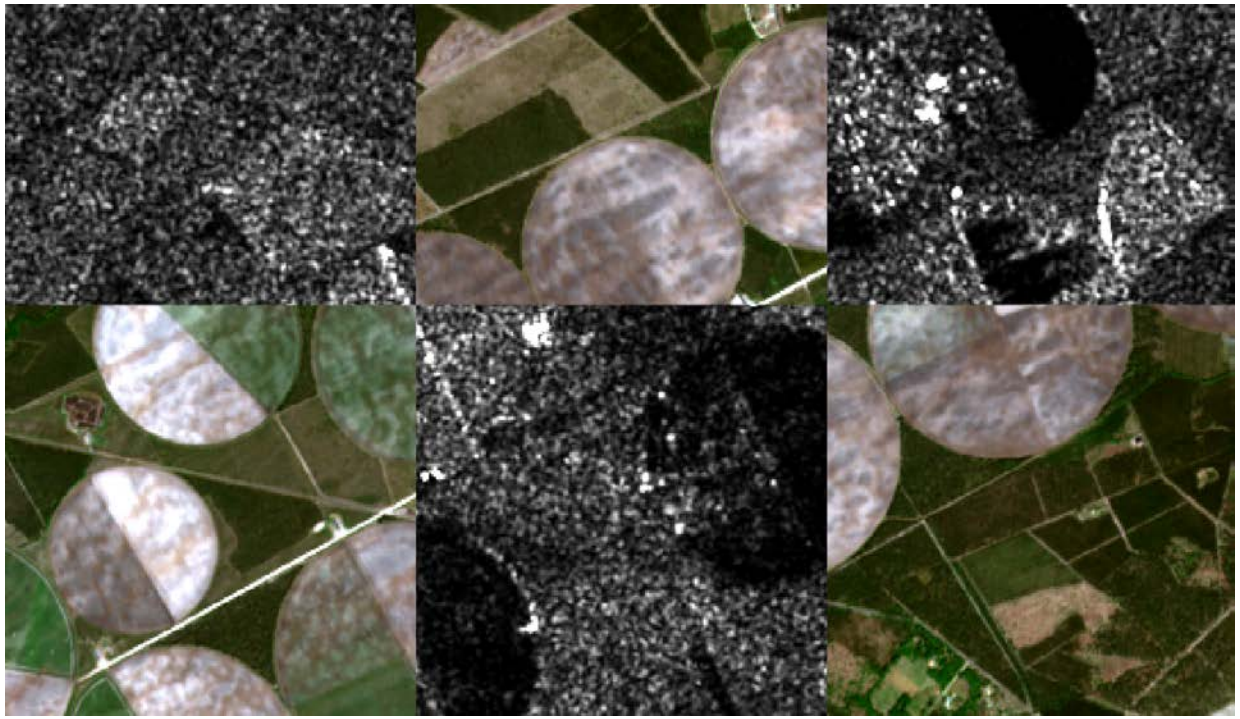
- Sentinel-1: The only system to provide SAR images regularly on all lands on the planet. Access to these time series of images opens an extraordinary range of applications



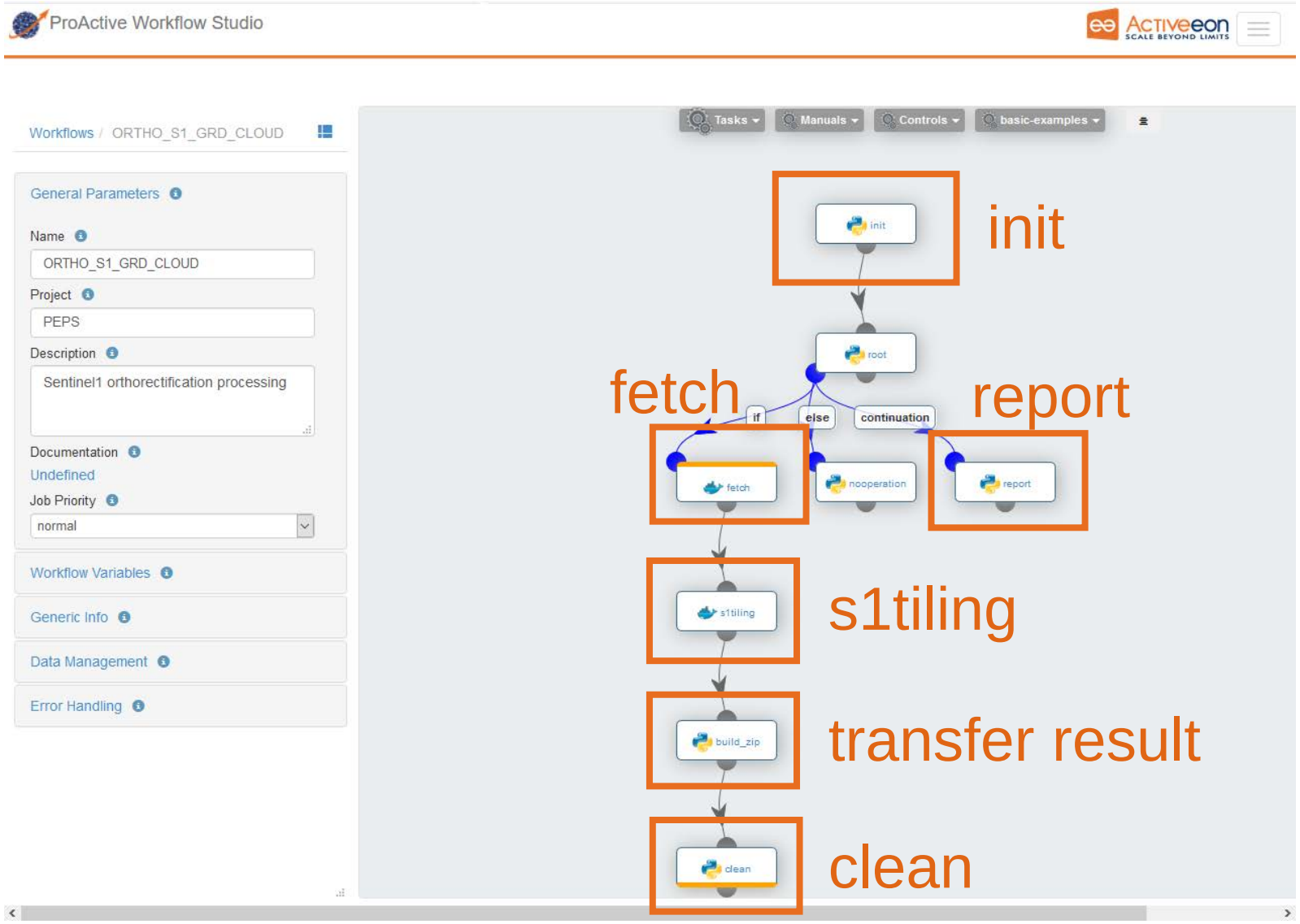
- Sentinel-1 ortho-rectification processing chain uses S1tiling algorithm developed to provide "Analysis Ready" time series for a very large number of applications
- Developed within the CNES radar service, in collaboration with CESBIO laboratory

Cloud bursting experiment: What?

- The tool benefits from the opensource Orfeo Tool Box developed by CNES
- Resulting images are registered to Sentinel-2 optical images, using the same MGRS geographic reference. It promotes joint use of both missions



Cloud bursting experiment: How?



init

Initialize workflow folders and variables

fetch Docker task

Fetch image data

s1tiling Docker task

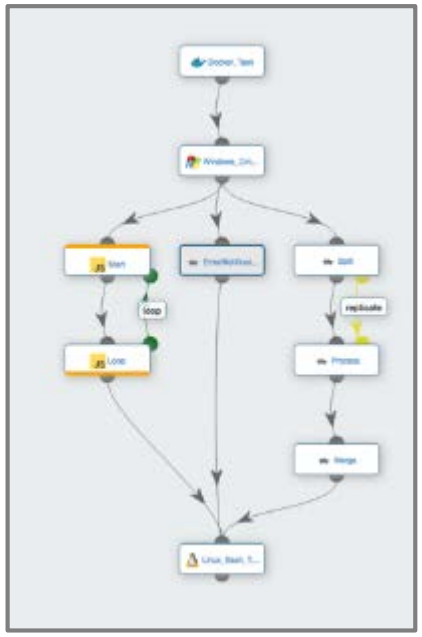
Image processing over 6 CPU cores

Transfer result

Clean

Report

Transparent hybrid cloud bursting?

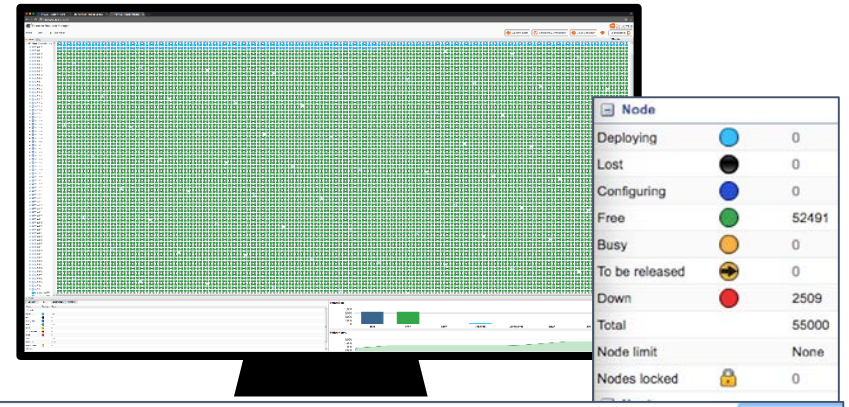


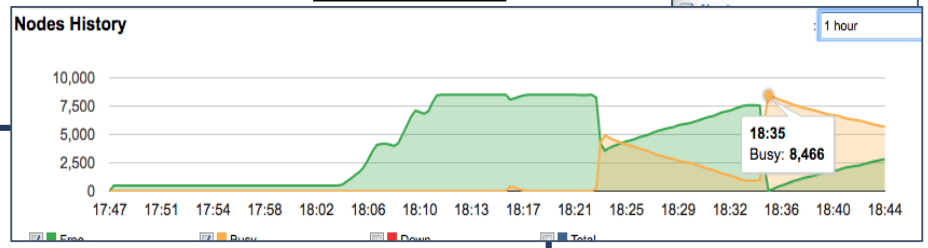
hadoop · kubernetes · docker · Java · Spark · python · sas · GREENPLUM DATABASE · ANACONDA · kafka

ProActive Workflows & Scheduling

Workflow Scheduler

Resource Manager





Azure Node Sources

VM · VM

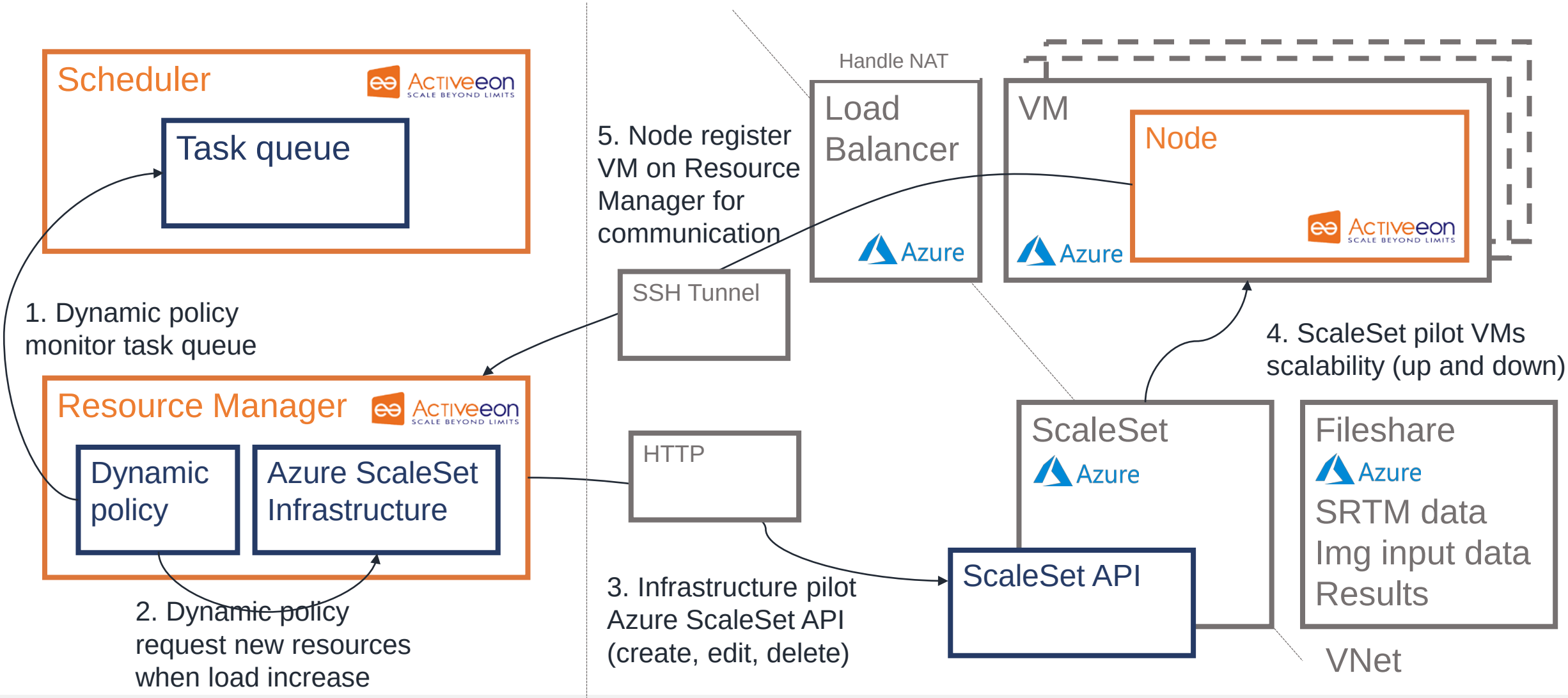
aws AWS Node Sources

Amazon S3 · Autoscale · EC2 · EC2

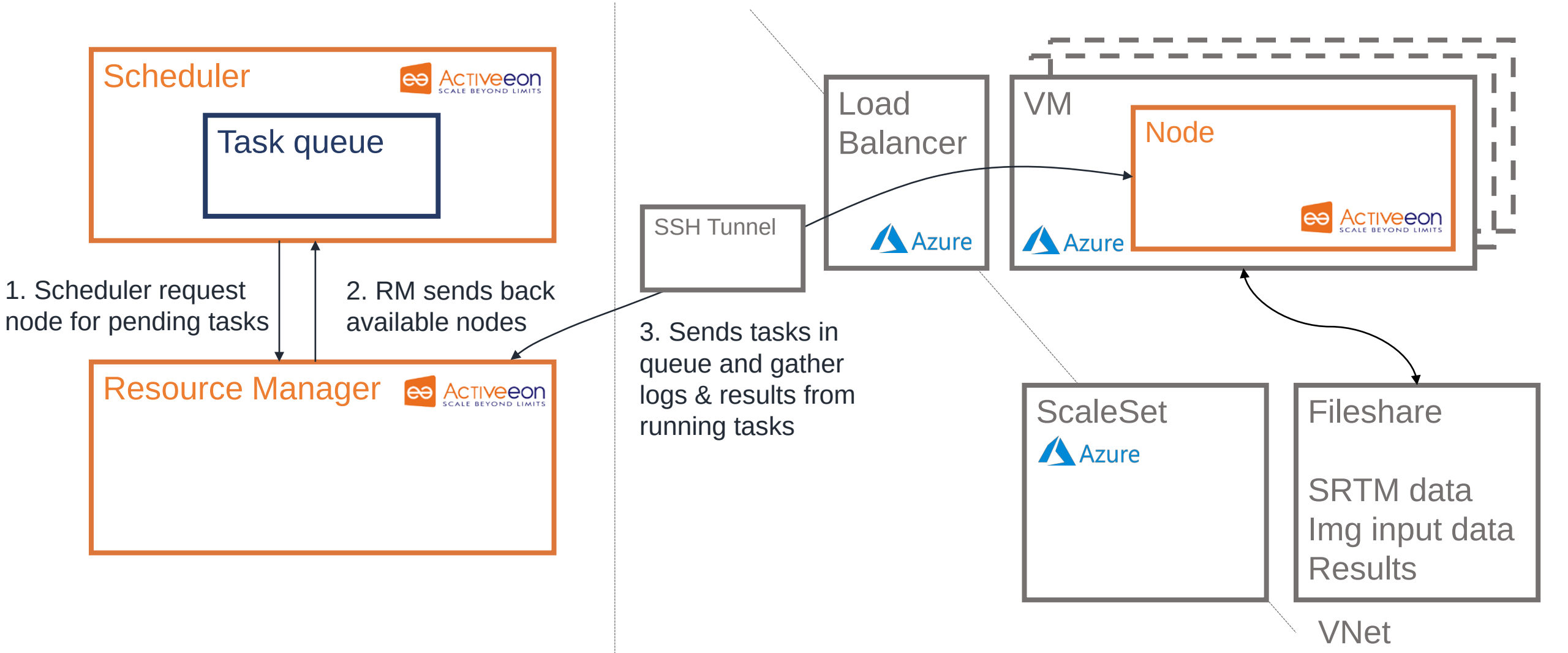
Private Node Sources

openstack · vmware · PBS · slurm · docker · Node · Node

Transparent hybrid cloud bursting: How?



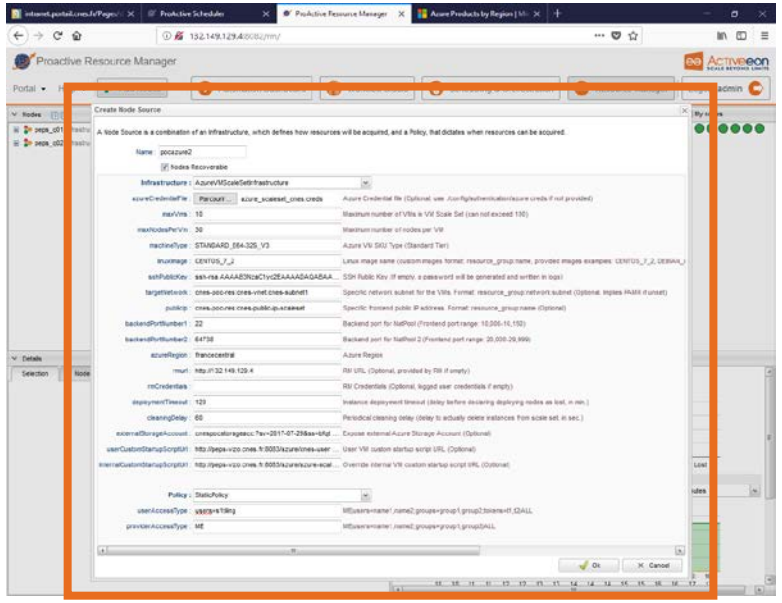
Transparent hybrid cloud bursting: How?

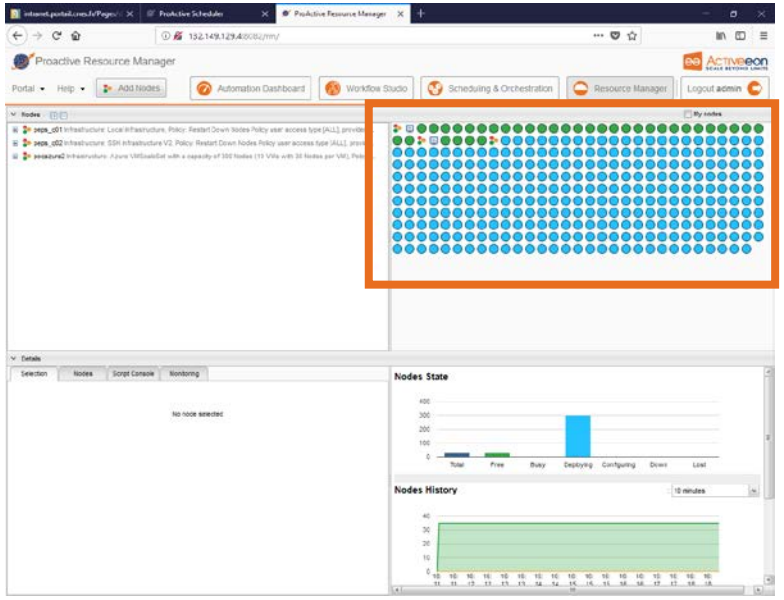


Cloud bursting experiment POC1

1

Resource allocation from Resource Manager



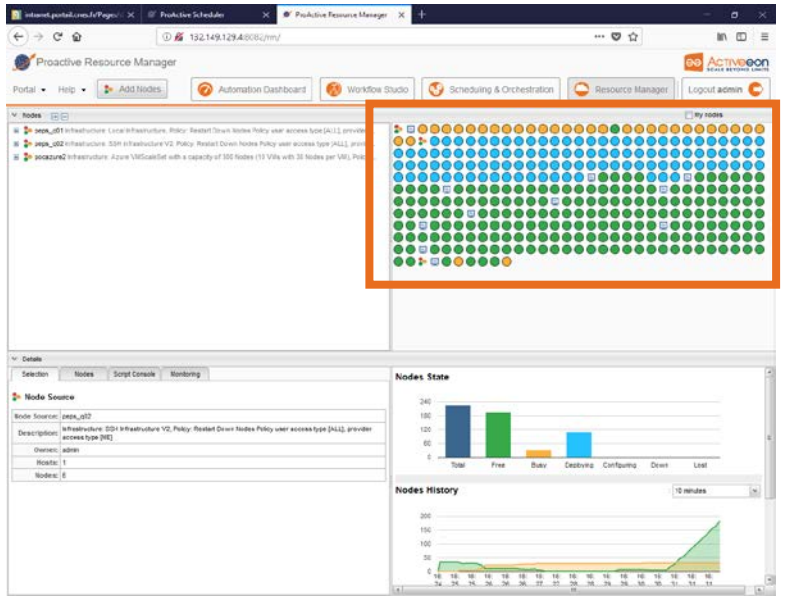


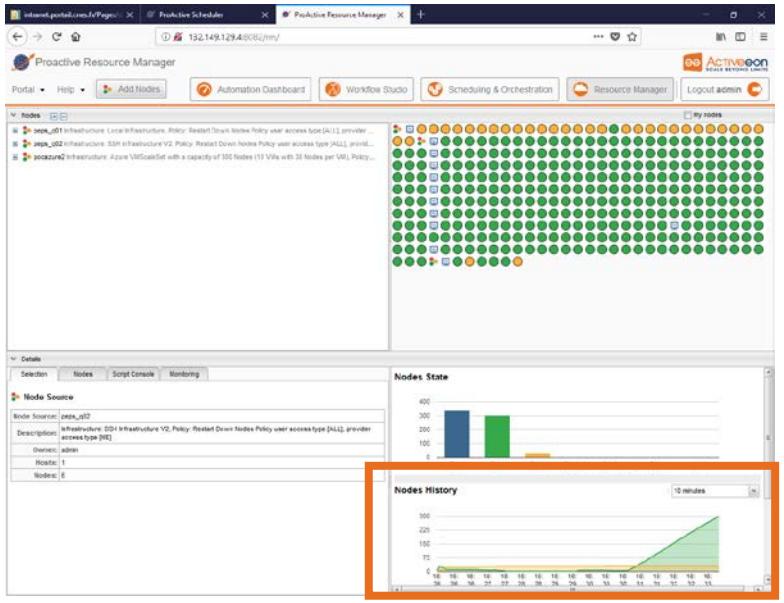
2

Waiting for VMs to start

3

First VMs Connected

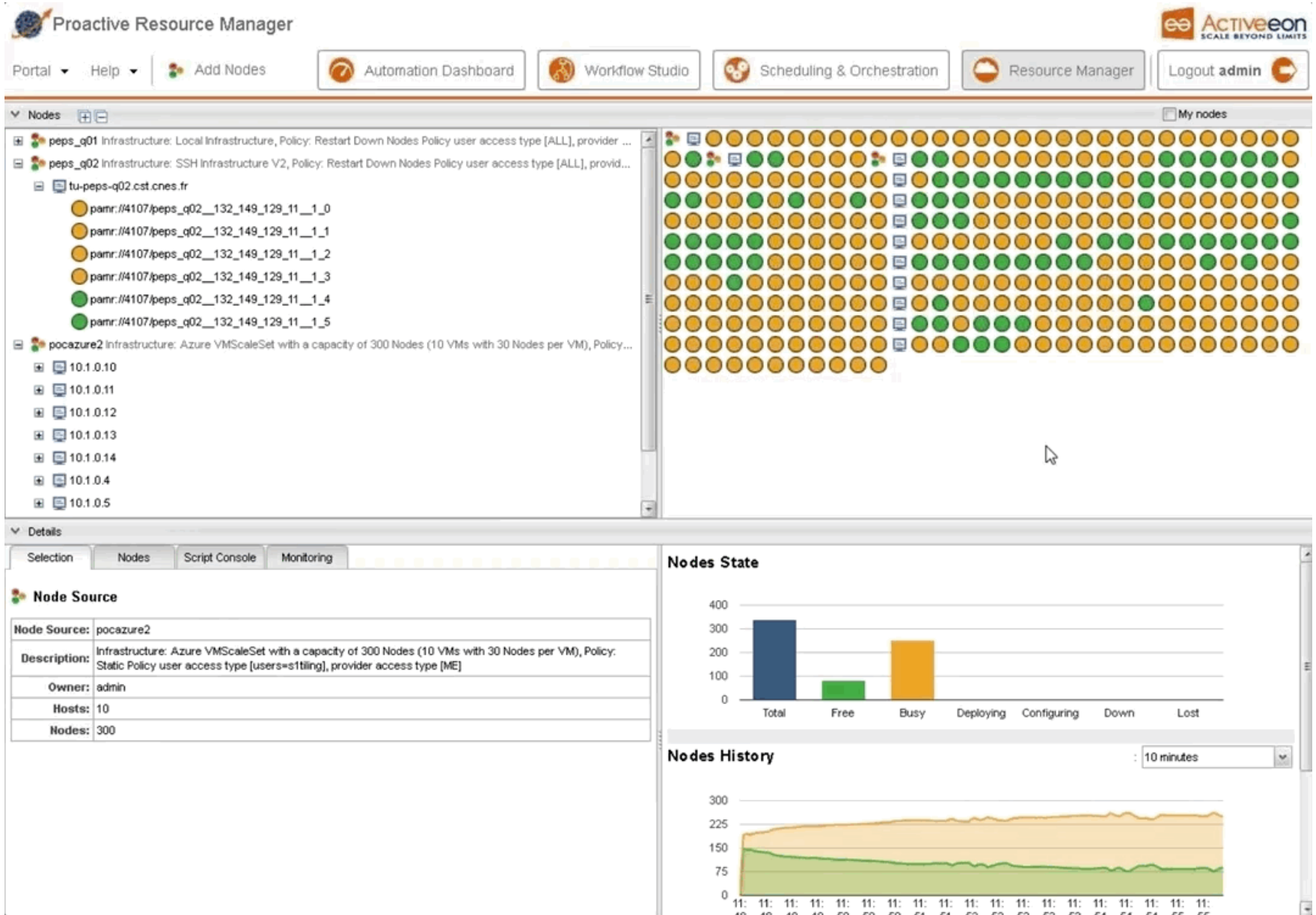




4

Visualize CPU cores available through time

Cloud bursting experiment POC1



> Up to 640 cores

- VM monitoring
- Resource utilization status

Results

Technical results:

- Successful transparent hybrid cloud bursting in 2018 from CNES datacenter into Microsoft Azure public cloud with 640 cores running the ortho-rectification chain

- Comparable performances obtained with:
  STANDARD_E64-32S_V3 Azure VM
  and
  on premise resources (HPC & dedicated Nodes)

- Issues identified related to bandwidth:
  a direct link to the datacenter used is necessary

Results

Costs and cloud benefits:

One of the main element of costs gives:

- 0.061€ CPU/hour for the STANDARD_E64-32S_V3 Azure VM

And

- 0,05€ CPU/hour for a CNES cluster HPC node all included (computation, local storage, full support, electricity consumption, investment)

- Benefits: Data processing architecture comparison, Performance & Cost optimization, Requirement maturity

Perspectives

- Computing distribution has been addressed
- To be done: data interconnection avoiding data duplication

=> Goal: Facilitate data centers cooperation

Thank you for your attention