PrestoCloud: Big Data Orchestration for Batch+Stream in Clouds & Fogs with ActiveEon

Denis Caromel, Vincent Kherbache, Sophie Song, Pascal Urso

ActiveEon
SCALE BEYOND LIMITS

Paris, Sophia Antipolis, London, San Jose USA

Company

Key information

- Founded in 2007 by Denis Caromel in Sophia-Antipolis, Spin-off of INRIA
- Addressing $80 Billion Hybrid Cloud Market with 27% CAGR
- Disruptive Patented Technology w/ Exceptional Business Outcomes
- 60% of the revenue from international

Global Locations

Sophia-Antipolis (France)
Paris (France)
London (United Kingdom)
San-Jose (United States)
Montreal (Canada)
Fribourg (Switzerland)
Dakar (Senegal)

ACTIVEEON
SCALE BEYOND LIMITS

ProActive Solution

Job Scheduling, Workload Automation
Orchestration & Meta-Scheduling

On-premises and on all clouds

Open Source

Management

Denis Caromel, CEO

François Tournesac, CSO

Brian Amedro, CTO

Partnerships

amazon web services | Partner Network
TECHNOLOGY PARTNER

Google Cloud Platform

redhat
BUSINESS PARTNER

Microsoft
Partner Network

docker
MEMBER

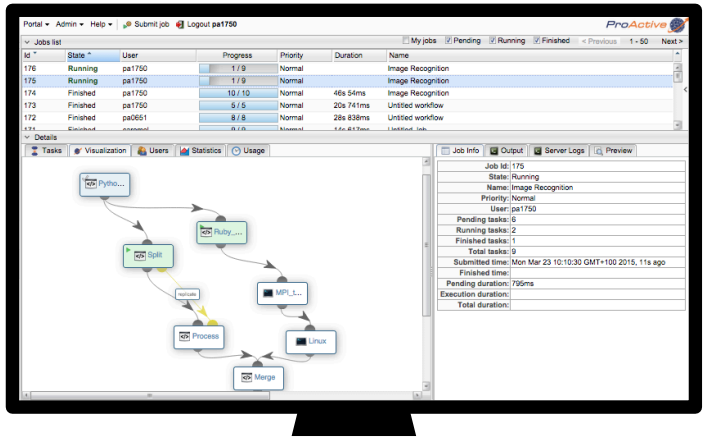
ProActive Workflows & Scheduling Core Product

ACTIVEEON SCALE BEYOND LIMITS

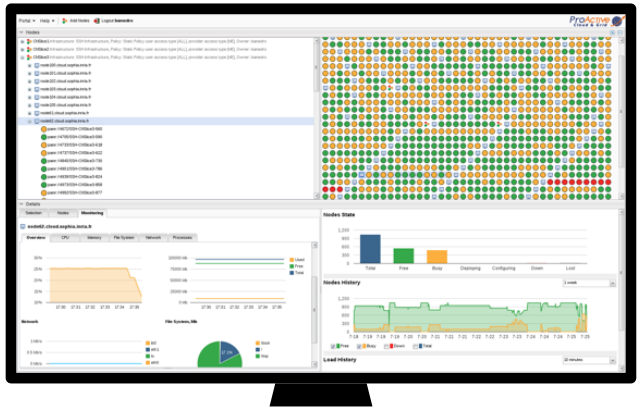
Processing & Automation Workflows

Workflow Studio

Scheduling & Orchestration



Job Console



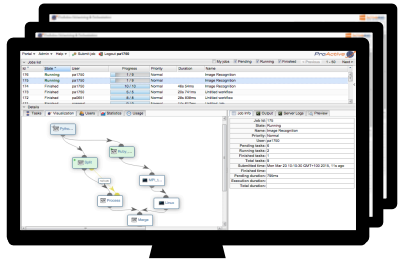
Resource Management & Monitoring

Resource Manager

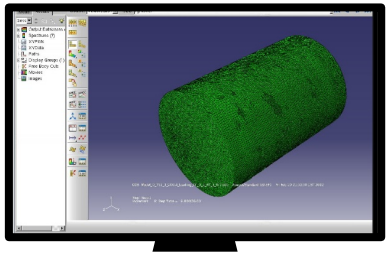
Big Compute for Everyone

ACTIVEEON
SCALE BEYOND LIMITS

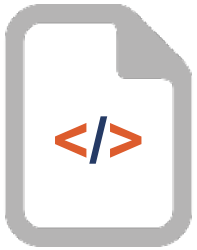
Application-specific Interfaces



Integrated Web Portals



Remote Visualization



API

ProActive Workflows

Fault Tolerance
Cloud bursting
Resource agnostic
Micro-service

Scheduler & Meta-Scheduler

Resource Manager

LSF
PBS Works
GRID ENGINE
slurm workload manager
hadoop
TIBCO Spotfire
R
MATLAB
GREENPLUM DATABASE
ANACONDA
sas

Clusters
kubernetes
docker
Azure
amazon web services
Google

PrEstoCloud

Proactive Cloud Resources Management At The Edge For Efficient Real-time Big Data Processing

PrEstoCloud Factsheet

Partners

- 36 months
  - Starts on January 2017
  - Ends on December 2019
- ActiveEon:
  - Core tech Provider
  - 60 PM
  - 501K€
- Provides Cloud Automation, Workflows and Scheduling for Real-Time Big Data Processing

Role of Partners

Software AG, Germany

Industrial partners

CVS Mobile, Slovenia

ADITESS, Cyprus

LiveU, Israel

N.Amram Technologies, Israel

Technology providers

Nissatech, Serbia

ActiveEon, France

UBITECH, Greece

Software AG, Germany

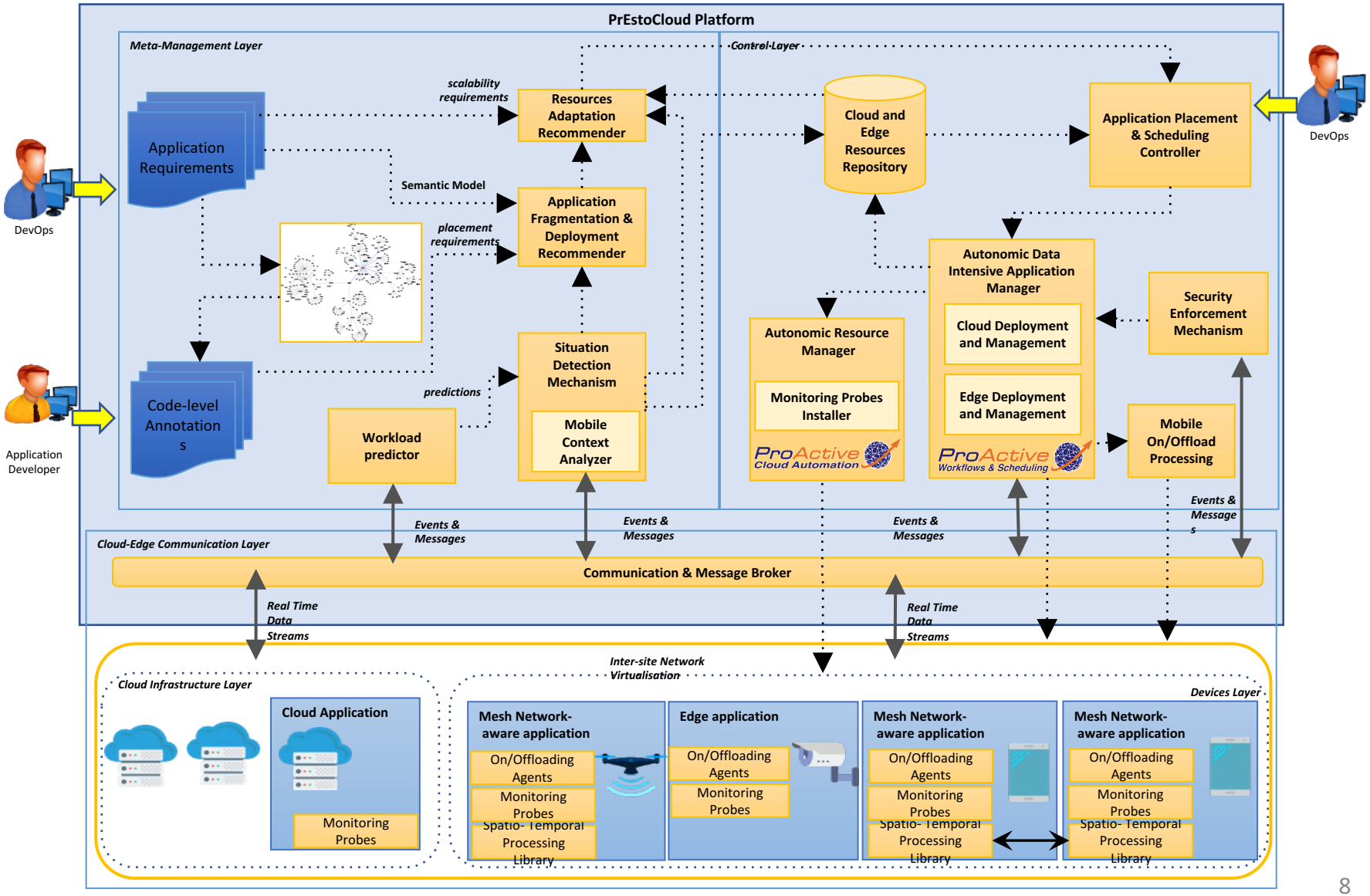
Research institutes

ICCS, Greece

IJS, Slovenia

CNRS, France

PrEstoCloud



Resources Management

Activeeon
SCALE BEYOND LIMITS

- Extended TOSCA specification
  - On public & private clouds: Virtual machines, Containers
  - On edge devices: Containers, Mobile app

ProActive
Cloud Automation

- Inter-site, multi-layer network virtualization
  - Route-based VPN
  - Security enforcement mechanism (RBAC, NFV)

Real-time big data processing

- Resource provisioning
  - Scheduling and orchestration of (de)commissioning workflows
  - Real-time events for the monitoring of Cloud/Edge resources and applications

Message broker (AMQP, MQTT, etc.)

  - Decision-making in the resources placement

Recommendation system (Constraints solver)

- Data-intensive applications management
  - Predictive Workload management
  - Scheduling and orchestration of deployment workflows
- (Re)configuration

Kafka Storm Visdom Streaming

ACTIVEEON
SCALE BEYOND LIMITS

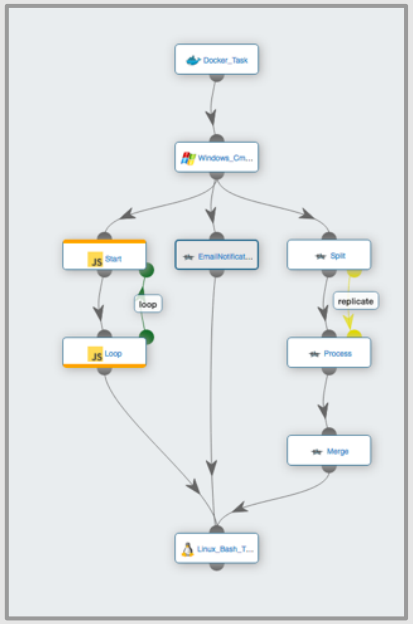
- Real-Time Data Streaming using ActiveEon Workflows:
  - Apache Kafka (Data streaming)
  - Apache Storm (Real-time stream processing)
  - Visdom (Stream Visualization from Facebook)

- https://www.youtube.com/watch?v=i6ZOyoXSq-A&t=5s

Testing the New Azure Data Centers with Microsoft & ActiveEon

Azure PoC in the Box

Activeeon
SCALE BEYOND LIMITS



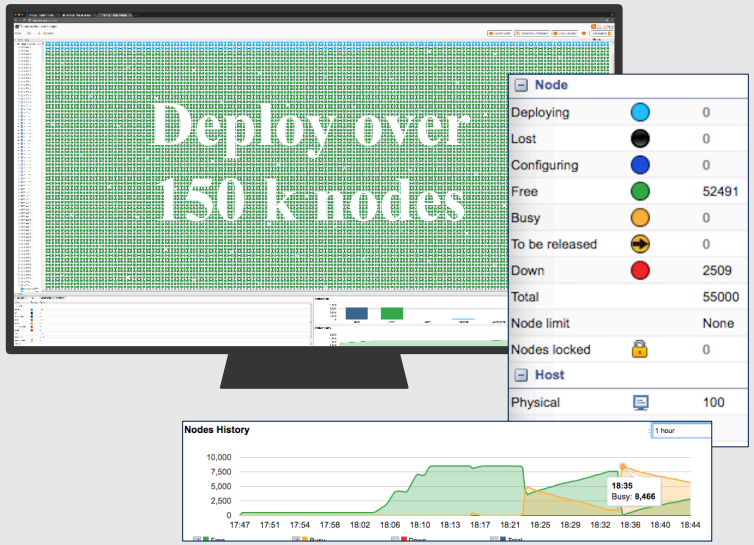
hadoop · kubernetes · docker · Java · Spark · python · sas · GREENPLUM DATABASE · ANACONDA · kafka

ProActive Workflows & Scheduling

Workflow Scheduler

Resource Manager

Scale automatically
Leverage Azure services



Azure Node Source Infrastructure

VM · VM

Using Azure Scale Sets

Existing Resources

Local & Network Resources

Private Cloud, HPC & Others

slurm workload manager · vmware · openstack · docker · kubernetes · LSF

150 K Nodes Benchmarks on Azure

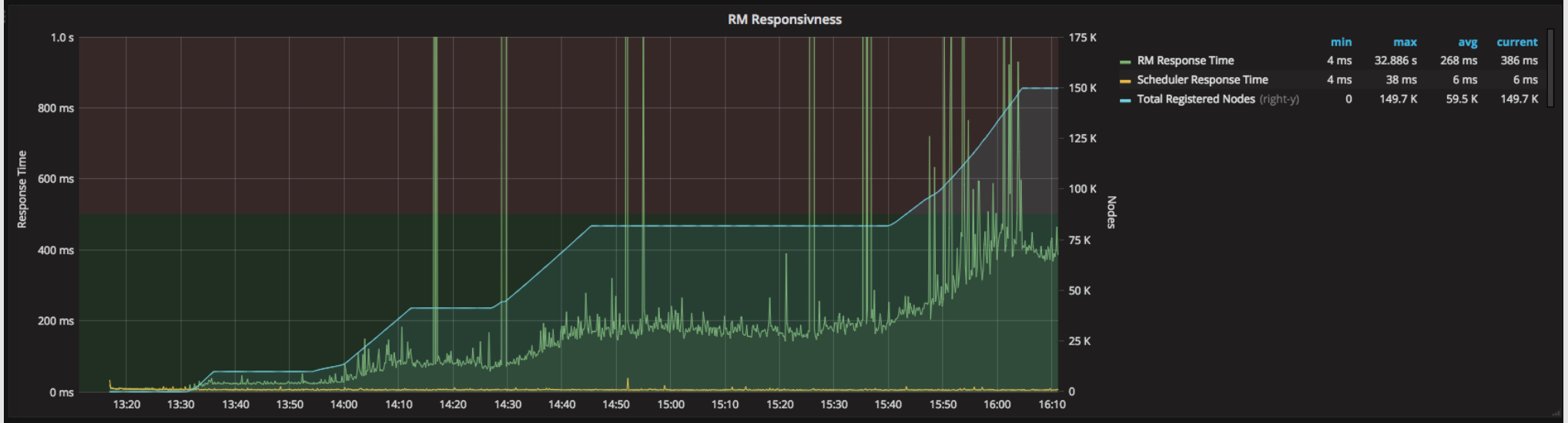
Activeeon
SCALE BEYOND LIMITS

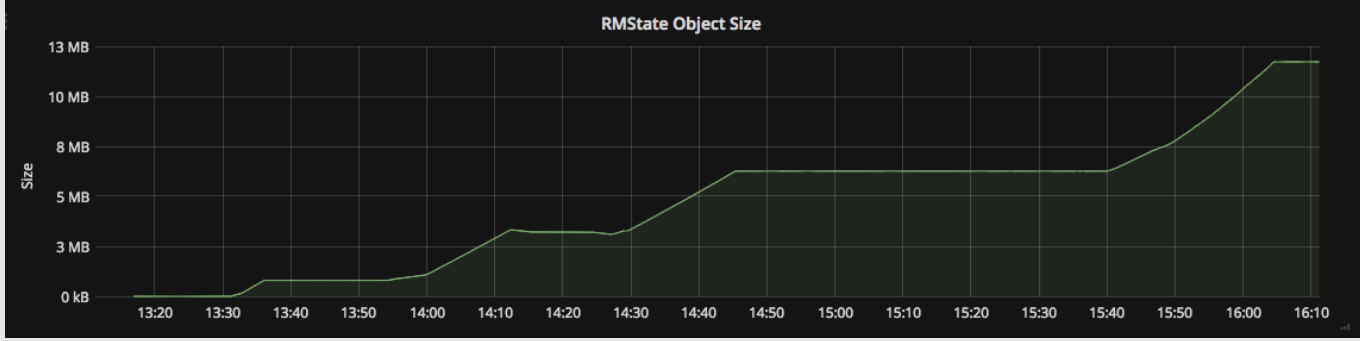
Setup

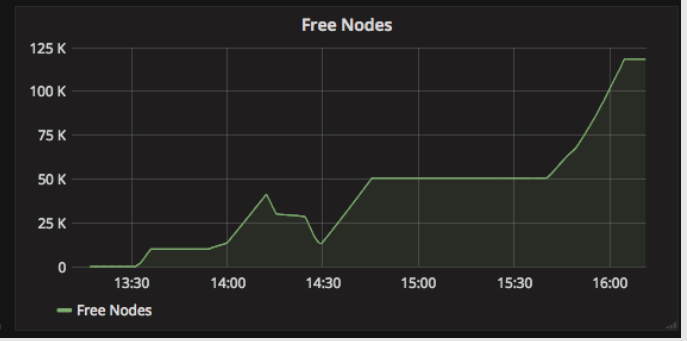
- RM isolated on a 16 cores, 64 GB machine
- Scheduler isolated on a 16 cores, 64 GB machine
- No REST Server
- Native ProActive client
- Azure VMScaleSet Infrastructure (Over Provisioning disabled)

Actions

- Retrieve rmState and schedulerState every 10sec (apiCall)
- At 13:23, created a Static NodeSource AzureVMScaleSet with 5 VMs, 2 000 Nodes per VM => + 10K Nodes
- At 13:40, created a Static NodeSource AzureVMScaleSet with 10 VMs, 5 000 Nodes per VM => + 50K Nodes
- At 14:17, created a Static NodeSource AzureVMScaleSet with 100 VMs, 500 Nodes per VM => + 50K Nodes
- At 14:54, created a Static NodeSource AzureVMScaleSet with 250 VMs, 500 Nodes per VM => + 125K Nodes ==> Failed because of quota limit (300 cores)
- At 15:30, created a Static NodeSource AzureVMScaleSet with 240 VMs, 500 Nodes per VM => + 120K Nodes







Automate Accelerate & Scale
10K Nodes, 20K Tasks, 1M Jobs

Activeeon
SCALE BEYOND LIMITS

+33 988 777 660

contact@activeeon.com

@activeeon

Paris, Sophia Antipolis, London, San Jose USA