Live Demonstration on a Production Platform

ProActive PACA:
1 368 Cores, 30 TB, 480 Cuda GPUs
Production Platform operated by:

Total:
- 1 368 Cores (soon 1 428)
- 480 GPU CUDA Cores
- 30TB Storage (soon 150 TB)

Publically Available Today
Conclusion
Typical Use Cases:
- Stress Tests
- Valuation of Exotic
- Algorithmic Trading
- Accelerate learning process

Flexibility:
- Multi-Appli Orchestration
- Physical Machines + Virtualization

In Production and For Developments:
- Continuous Integration
- Non Regression Tests
- Scalability Tests
- Load Injection

Partners:
- Hardware Vendors
- Integrators

Redistribution:
- OEMs with ISVs
Orange Labs Paris Use Case
A Community Cloud show case: Benchmarking as a Service

- CCs enable more realistic workload thanks to widely distributed load injection sources
- External and internal clients can get a benchmarking infrastructure (hardware and software) on demand, instead of owning expensive and underused resources
- Use data centers’ free resources, as well as CCs (unused employees’ ordinary PCs, e.g. at night, weekend, holidays…) instead of dedicated load injection servers
- Contributors: MAPS/MEP, in liaison w/ DDSI, ActiveEon

What is a “Community Cloud”?
- Members of a "Community" (SMEs, NGOs, individuals) agree to share CPUs, storage, network to setup a secured Cloud
- Same uses as classical clouds, but cheaper!

Opportunities for Orange
- Business market: new or existing offers harnessing edge resources
  - a trusted operator shall provide a valuable support to Communities
  - security, availability and QoS, complementary resources...
- Group CAPEX and OPEX reduction
  - Orange's employee PCs may form a Community Cloud: see the demo
- Green cloud by using existing resources rather than new big datacenters
- Emergent countries: an incentive for the development of broadband access

Next steps
- R&D field user tests of the Benchmarking aaS
- Orange Hub: prototyping a deployment of Community Cloud agent in SME premises
ProActive MapReduce (CO, SP2, Task 2.1)

- Same APIs as Hadoop
  (Easy switch from Hadoop to ProActive)

- Does not require an HDFS File System

- Runs on general purpose, Multi-tenant, Multi-Applications Grids and Clouds

- Available as PaaS in Java
Workflow ProActive MapReduce
### ProActive MapReduce vs. Hadoop+HDFS

<table>
<thead>
<tr>
<th>File Size</th>
<th>Sequential</th>
<th>Hadoop</th>
<th>PA MapReduce</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7 GB</td>
<td>5m 04s</td>
<td>1m 17s</td>
<td>1m 05s</td>
<td>4.6</td>
</tr>
<tr>
<td>4.3 GB</td>
<td>25m 31s</td>
<td>2m 30s</td>
<td>2m 20s</td>
<td>10.9</td>
</tr>
<tr>
<td>7.3 GB</td>
<td>46m 00s</td>
<td>3m 31s</td>
<td>3m 30s</td>
<td>13.1</td>
</tr>
<tr>
<td>20 GB</td>
<td>2h 07m 00s</td>
<td>8m 30s</td>
<td>7m 09s</td>
<td>17.8</td>
</tr>
<tr>
<td>50 GB</td>
<td>5h 19m 00s</td>
<td>21m 05s</td>
<td>25m 11s</td>
<td>12.7</td>
</tr>
<tr>
<td>100 GB</td>
<td>10h 38m 00s</td>
<td>43m 23s</td>
<td>58m 42s</td>
<td>10.9</td>
</tr>
</tbody>
</table>

- Data available in a NAS (General purpose storage)
- Transfer to HDFS for Hadoop
- Used directly without copy for ProActive
- Use Case of Map/Reduce on fresh data
- Different ProActive Map/Reduce configuration for recurrent MR on in place Data (e.g. ProActive HDFS interface)
ProActive MapReduce with GlusterFS

<table>
<thead>
<tr>
<th>File Size</th>
<th>Sequential</th>
<th>Hadoop (HDFS)</th>
<th>PA MapReduce (glusterfs)</th>
<th>PA MapReduce (ceph)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 GB</td>
<td>2h 07m 00s</td>
<td>3m 25s</td>
<td>4m 30s</td>
<td>5m 50s</td>
</tr>
<tr>
<td>50 GB</td>
<td>5h 19m 00s</td>
<td>8m 30s</td>
<td>8m 30s</td>
<td>11m 30s</td>
</tr>
<tr>
<td>100 GB</td>
<td>10h 38m 00s</td>
<td>15m 30s</td>
<td>15m 00s</td>
<td>19m 00s</td>
</tr>
</tbody>
</table>

Hadoop:
- 20 machines, 8 cores each
- HDFS - split size 64M
- Upload times Transfer to HDFS not taken into account for Hadoop

ProActive:
- 16 nodes, 8 cores each
- GlusterFS, Ceph: 4 data servers + 1 frontend server
**Cloud & Grid IaaS**

Physical Machines (Servers, Clusters, Desktops) and Virtual (*Hyper-V*, *VMware*, *KVM*, *Xen*; *OpenStack*)

Dynamic Policies, Full accounting of resource usage

Public Cloud Burst (EC2, Azure, Data Centers)

---

**Orchestration & Scheduling**

Multi-Application & Multi-Tenant Portal with Data Management, Remote Visualization

APIs (Java, REST, CLI)

Integration in existing Applications and Web Portals

---

**HPC Workflow & Parallelization**

Studio for HPC, Workflow Visualization

Native Tasks (MPI, OpenMP, Multi-thread, GPU)

Java APIs for Parallelization & Distribution

Matlab & Scilab Distributed Computing

---

**Technology & Solutions**
**ProActive Parallel Suite** is an innovative Open Source solution (OW2) for acceleration and orchestration of applications, seamlessly integrated with the management of high-performance Clouds (Private and Public with bursting capabilities).

**ProActive** platform features High-Performance Workflows and Application Parallelization, together with enterprise Scheduling and Orchestration coupled with the dynamic management of private Heterogeneous Grids and Clouds.

With a ProActive platform, our users tackle at once the acceleration and orchestration of all demanding enterprise applications, and the management of their own Enterprise Cloud.

**Workflow & Parallelization**

*HPC Workflows, Application Parallelization*
Accelerate & Scale your most demanding applications, execute in the Clouds:
- Studio for HPC, Workflow Visualization
- Native Tasks (MPI, OpenMP, Multi-thread, GPU)
- Java APIs for Parallelization & Distribution, MapReduce
- Matlab & Scilab Integration

**Orchestration & Scheduling**

*Multi-Platform Job & Workflow Scheduler*
Balance HPC loads on all company resources and on external clouds:
- Multi-Application & Multi-Tenant
- Portal with Data Management, Remote Visualization
- Rich APIs: Java, REST, CLI
- Integration in existing Applications and Web Portals

**Clouds & Grids**

*Heterogeneous Resource Manager*
Leverage existing infrastructures:
- Physical Machines (Servers, Clusters, Desktops, GPU)
- Virtualization (Hyper-V, VMware, KVM, Xen, QEMU, OpenStack)
- Dynamic Policies, Green IT
- Accounting of resource usage
- Cloud Burst (EC2, Azure, private Data Centers)

ProActive unique features are full support for Linux, Windows and Mac, the management of Physical and Virtual machines, a rich set of APIs (Java, REST, CLI), and a graphical Studio for the design and execution of HPC workflows in the Cloud.
**Products**

- Programming
- Scheduling
- Resourcing

---

**ProActive Parallel Suite**

**Workflow & Parallelization**

- HPC Workflows, Application Parallelization

**Orchestration & Scheduling**

- Multi-Platform Job & Workflow Scheduler

**Cloud & Grids**

- Heterogeneous Resource Manager

---

**Enterprise Edition**

ALL-IN-ONE
Certified Package
Free evaluation with One Month Support

---

Accelerate & Scale your most demanding applications, execute in the Clouds:

- Studio for HPC, Workflow Visualization
- Native Tasks (MPI, OpenMP, Multi-thread, GPU)
- Java APIs for Parallelization & Distribution
- Matlab & Scilab Integration

Schedule and manage parallel jobs on all your company’s resources and on external clouds:

- Multi-Application & Multi-Tenant
- Portal with Data Management, Remote Visualization
- Rich APIs: Java, REST, CLI
- Integration in existing Applications and Web Portals

Leverage existing infrastructures, connect with external Clouds:

- Physical Machines (Servers, Clusters, Desktops, GPU)
- Virtualization (Hyper-V, VMware, KVM, Xen, QEMU, OpenStack)
- Dynamic Policies, Green IT
- Accounting of resource usage
- Cloud Burst (EC2, Azure, private Data Centers)

---

Download **Enterprise Edition**

Download **Community Edition**

Download **Community Edition**
PPS: ProActive Parallel Suite

**Workflow & Parallelization**
for Application Developer, End-User

**Orchestration & Scheduling**
for SysAdmin

**Cloud & Grid IaaS**
for SysAdmin

- Studio
  - APIs: Java, CLI, REST
  - Matlab Scilab
  - MapReduce

- Web Portals
- Java Active Object: API, GUI

- Orchestration Server
- REST Server
- Web Server
- Scheduler Worker

- Cloud Grid Server
- Linux Agent
- Windows Agent
- Virtual Machine Management