MathCloud: From Software Toolkit to Cloud Platform for Building Computing Services

O.V. Sukhoroslov

Centre for Grid Technologies and Distributed Computing ISA RAS
Moscow Institute for Physics and Technology
MathCloud

- Research project started in 2008 at CGTDC ISA RAS
  - Investigate the use of distributed service-oriented environments for scientific research
  - Develop an open platform for building, deployment, discovery and integration of computing services
  - Create a testbed environment and use it for solving complex scientific problems

http://mathcloud.org/
Problem

```
[marllon.muniz@cromo ~]$ qstat -r

cromo.ufabc.edu.br:

<table>
<thead>
<tr>
<th>Job ID</th>
<th>Username</th>
<th>Queue</th>
<th>Jobname</th>
<th>SessID</th>
<th>NDS</th>
<th>TSK</th>
<th>Req'd Memory</th>
<th>Req'd Time</th>
<th>Elap Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>36060.cromo.ufab</td>
<td>fbattani</td>
<td>long</td>
<td>F_Zn_4sulf_frep</td>
<td>3808</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>144:0</td>
<td>R 38:22</td>
</tr>
<tr>
<td>3613L.cromo.ufab</td>
<td>cmaciel</td>
<td>medium</td>
<td>fe-dea-no3</td>
<td>24309</td>
<td>1</td>
<td>--</td>
<td>8gb</td>
<td>72:00</td>
<td>R 38:22</td>
</tr>
<tr>
<td>36355.cromo.ufab</td>
<td>reilya</td>
<td>long</td>
<td>ribbonAu0</td>
<td>4988</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>100:0</td>
<td>R 59:20</td>
</tr>
<tr>
<td>36357.cromo.ufab</td>
<td>reilya</td>
<td>long</td>
<td>ribbonAu0</td>
<td>24633</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>100:0</td>
<td>R 59:28</td>
</tr>
<tr>
<td>36358.cromo.ufab</td>
<td>josue</td>
<td>medium</td>
<td>svd-ap-sep-gauss</td>
<td>22561</td>
<td>1</td>
<td>--</td>
<td>4gb</td>
<td>72:00</td>
<td>R 58:02</td>
</tr>
<tr>
<td>36355.cromo.ufab</td>
<td>fbattani</td>
<td>long</td>
<td>F_4sulf2</td>
<td>15479</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>144:0</td>
<td>R 51:49</td>
</tr>
<tr>
<td>36395.cromo.ufab</td>
<td>lucas</td>
<td>long</td>
<td>SO05Mgpc-300</td>
<td>16494</td>
<td>16</td>
<td>--</td>
<td>4gb</td>
<td>144:0</td>
<td>R 29:11</td>
</tr>
<tr>
<td>36400.cromo.ufab</td>
<td>lucas</td>
<td>long</td>
<td>SO05Mgpc-350</td>
<td>11022</td>
<td>16</td>
<td>--</td>
<td>4gb</td>
<td>144:0</td>
<td>R 29:10</td>
</tr>
<tr>
<td>36401.cromo.ufab</td>
<td>lucas</td>
<td>long</td>
<td>SO10Mgpc-300</td>
<td>6533</td>
<td>16</td>
<td>--</td>
<td>4gb</td>
<td>144:0</td>
<td>R 23:54</td>
</tr>
<tr>
<td>36404.cromo.ufab</td>
<td>lucas</td>
<td>long</td>
<td>SO10Mgpc-350</td>
<td>11444</td>
<td>16</td>
<td>--</td>
<td>4gb</td>
<td>144:0</td>
<td>R 07:29</td>
</tr>
<tr>
<td>36436.cromo.ufab</td>
<td>cmaciel</td>
<td>medium</td>
<td>fe-dea-no3</td>
<td>8694</td>
<td>1</td>
<td>--</td>
<td>8gb</td>
<td>72:00</td>
<td>R 23:31</td>
</tr>
<tr>
<td>36444.cromo.ufab</td>
<td>cmaciel</td>
<td>medium</td>
<td>fe-dea-no3</td>
<td>25477</td>
<td>1</td>
<td>--</td>
<td>4gb</td>
<td>72:00</td>
<td>R 12:13</td>
</tr>
<tr>
<td>36445.cromo.ufab</td>
<td>cmaciel</td>
<td>medium</td>
<td>fe-dea-no3</td>
<td>25478</td>
<td>1</td>
<td>--</td>
<td>4gb</td>
<td>72:00</td>
<td>R 12:13</td>
</tr>
<tr>
<td>36448.cromo.ufab</td>
<td>ygor</td>
<td>medium</td>
<td>tut1</td>
<td>19353</td>
<td>1</td>
<td>--</td>
<td>9gb</td>
<td>72:00</td>
<td>R 05:56</td>
</tr>
<tr>
<td>36449.cromo.ufab</td>
<td>fbattani</td>
<td>long</td>
<td>F_AL_4sulf_freq</td>
<td>22049</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>144:0</td>
<td>R 05:03</td>
</tr>
<tr>
<td>36451.cromo.ufab</td>
<td>fbattani</td>
<td>long</td>
<td>F_4H_cris_freq</td>
<td>7409</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>144:0</td>
<td>R 04:34</td>
</tr>
<tr>
<td>36452.cromo.ufab</td>
<td>rgamorim</td>
<td>long</td>
<td>wig-Il-str67-a</td>
<td>14721</td>
<td>1</td>
<td>--</td>
<td>14gb</td>
<td>144:0</td>
<td>R 04:34</td>
</tr>
<tr>
<td>36453.cromo.ufab</td>
<td>rgamorim</td>
<td>long</td>
<td>wig-Il-str67-b</td>
<td>10848</td>
<td>1</td>
<td>--</td>
<td>12gb</td>
<td>144:0</td>
<td>R 04:34</td>
</tr>
<tr>
<td>36465.cromo.ufab</td>
<td>fbattani</td>
<td>long</td>
<td>F_AL_4H_freq2</td>
<td>8540</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>144:0</td>
<td>R 03:52</td>
</tr>
<tr>
<td>36467.cromo.ufab</td>
<td>reilya</td>
<td>long</td>
<td>ribbonAu0</td>
<td>24630</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>24:00</td>
<td>R 01:49</td>
</tr>
</tbody>
</table>
```
### Problem

#### maq-0.7.1

<table>
<thead>
<tr>
<th>Option</th>
<th>Type</th>
<th>Value</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-&lt;</td>
<td></td>
<td>INFILE</td>
<td>Chromosome reference (FASTA) (required)</td>
</tr>
<tr>
<td>-reads_1 bq</td>
<td></td>
<td>INFILE</td>
<td>First read (BFC) (required)</td>
</tr>
<tr>
<td>-reads_2 bq</td>
<td></td>
<td>INFILE</td>
<td>Second read (BFC) (optional)</td>
</tr>
<tr>
<td>-1</td>
<td>INT</td>
<td>0</td>
<td>length of the first read (&lt;=127)</td>
</tr>
<tr>
<td>-2</td>
<td>INT</td>
<td>0</td>
<td>length of the second read (&lt;=127)</td>
</tr>
<tr>
<td>-m</td>
<td>FLOAT</td>
<td>0.001</td>
<td>rate of difference between reads and references</td>
</tr>
<tr>
<td>-e</td>
<td>INT</td>
<td>0</td>
<td>maximum allowed sum of qualities of mismatches</td>
</tr>
<tr>
<td>-d</td>
<td>OUTFILE</td>
<td>□</td>
<td>adapter sequence file</td>
</tr>
<tr>
<td>-a</td>
<td>INT</td>
<td>250</td>
<td>max distance between two paired reads</td>
</tr>
<tr>
<td>-A</td>
<td>INT</td>
<td>0</td>
<td>max distance between two RF paired reads</td>
</tr>
<tr>
<td>-n</td>
<td>INT</td>
<td>2</td>
<td>number of mismatches in the first 24bp</td>
</tr>
<tr>
<td>-M</td>
<td>SELECT</td>
<td>□</td>
<td>methylation alignment mode</td>
</tr>
<tr>
<td>-u</td>
<td>OUTFILE</td>
<td>□</td>
<td>dump unmapped and poorly aligned reads to FILE</td>
</tr>
<tr>
<td>-H</td>
<td>OUTFILE</td>
<td>□</td>
<td>dump multiple/all 01-mismatch hits to FILE</td>
</tr>
<tr>
<td>-C</td>
<td>INT</td>
<td>250</td>
<td>max number of hits to output, &gt;512 for all 01 hits.</td>
</tr>
<tr>
<td>-s</td>
<td>INT</td>
<td>0</td>
<td>seed for random number generator (Default: random)</td>
</tr>
<tr>
<td>-w</td>
<td>FLAG</td>
<td>□</td>
<td>disable Smith-Waterman alignment</td>
</tr>
<tr>
<td>-t</td>
<td>FLAG</td>
<td>□</td>
<td>turn all reads (usually not recommended)</td>
</tr>
<tr>
<td>-c</td>
<td>FLAG</td>
<td>□</td>
<td>match in the colorspace</td>
</tr>
</tbody>
</table>

**Powered by sana**
Problem
Problem

Computing Services

- Solve Specific Classes of Problems
- Hide Complex Computing Infrastructure
- Reusable and Composable
Requirements

1. Use computing environment as a service for solving a problem
2. Access services via convenient and familiar interface
3. Easily combine services with each other
4. Easily create services and share them with colleagues
Service-Oriented Environment

Users

Applications

Services

Computing Resources
Approach

Service
Approach

Service

Unified Programming Interface
Approach

Service

Unified Programming Interface

Web Interface

Application

Service
Approach

- Service-Oriented Architecture
- Service-Oriented Science (Foster, 2005)
- Software as a Service
- Web 2.0
- Representational State Transfer (REST)
Computing Service as a Web Function

Service
Identified by URI
Computing Service as a Web Function

Input parameters
(Problem Description)
JSON Schema

Service
Identified by URI
Computing Service as a Web Function

Input parameters
(Problem Description)
JSON Schema

Service
Identified by URI

Output parameters
(Problem Solution)
JSON Schema
Computing Service as a Web Function

Input parameters
(Problem Description)
JSON Schema

Service
Identified by URI

Output parameters
(Problem Solution)
JSON Schema

Input JSON
HTTP

Output JSON
HTTP
Computing Service as a Web Function

Input parameters
(Problem Description)
JSON Schema

Output parameters
(Problem Solution)
JSON Schema

Required Features
- Service Introspection
- Long Computations
- Passing Large Data

Core Technologies
- HTTP + REST
- JSON
- JSON Schema

Service
Identified by URI

Input JSON

Output JSON

HTTP
# Unified REST API

<table>
<thead>
<tr>
<th>URL Type</th>
<th>GET</th>
<th>POST</th>
<th>DELETE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERVICE_URI</td>
<td>Get service description</td>
<td>Submit new request → Get JOB_URI</td>
<td></td>
</tr>
<tr>
<td>JOB_URI</td>
<td>Get job status and results</td>
<td></td>
<td>Cancel job</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Delete job data</td>
</tr>
<tr>
<td>FILE_URI</td>
<td>Get file data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SERVER_URI</td>
<td>Get list of services provided by server</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MathCloud Platform

- A set of tools for building, deployment, discovery and integration of computing services implementing the unified REST API
  - Service Container (Everest)
  - Service Catalogue
  - Workflow Management System (WfMS)
  - Client Interfaces
  - Security
- Some components are publicly available
- Ultimate goal is to open source the whole platform
Service Container (Everest)
"name": "POV-Ray",
"description": "Runs POV-Ray raytracer",

"inputs": {
  "scene": {
    "title": "Scene file",
    "type": "string",
    "format": "uri",
    "required": true
  },
  "params": {
    "title": "Rendering params",
    "type": "string",
    "description": "Example: +W320 +H240"
  }
},

POV-Ray
 Runs POV-Ray raytracer

Scene file

Rendering params
Example: +W320 +H240

Submit
"outputs": {
    "image": {
        "title": "Rendered image",
        "type": "string",
        "format": "uri",
        "required": true
    },
    "log": {
        "title": "Runtime log",
        "type": "string",
        "format": "uri",
        "description": "This is the raw output of POV-Ray",
        "required": true
    }
},

"config": {
    "input-files": {
        "scene": "scene.pov",
    },
    "adapter": {
        "type": "command",
        "command": "povray +Iscene.pov +Oimage.png +FN $params$",
    },
    "output-files": {
        "image": "image.png",
        "log": "stderr"
    }
}
## Service Catalogue

### Math Cloud

**conv_hull**
- **convex hulls (using Qhull)**

**convex java**

### echo.java

**echo-java**
- **A simple example of service implemented in Java**

**echo java**

### povray

**povray**
- **POV-Ray raytracer. An advanced example for command-line service plugin.**

**povray**

### johnson

**johnson**
- **Johnson polyhedra**

**convex java**

<table>
<thead>
<tr>
<th>ID</th>
<th>Название</th>
<th>Тип</th>
<th>Описание</th>
</tr>
</thead>
<tbody>
<tr>
<td>input_file</td>
<td>file</td>
<td></td>
<td>input file</td>
</tr>
<tr>
<td>append</td>
<td>boolean</td>
<td></td>
<td>append the convex hull to the input file</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>Название</th>
<th>Тип</th>
<th>Описание</th>
</tr>
</thead>
<tbody>
<tr>
<td>output_file</td>
<td>file</td>
<td></td>
<td>Output file</td>
</tr>
<tr>
<td>error_file</td>
<td>file</td>
<td></td>
<td>Error file</td>
</tr>
</tbody>
</table>

### Tags

- antiprism(3)
- bonmin(2)
- cluster(2)
- command(1)
- convex java(1)
- echo(1)
- echo java(1)
- glpk(2)
- ipopt(2)
- lpsolve(2)
- maxima(1)
- NeosClient(1)
- optimization(3)
- polyhedra(1)
- povray(1)
- test(2)
- weave(1)
- грац(1)
- кластер(1)
- сервис(1)
Workflow Management System
Client Interfaces

- Accessing MathCloud services from applications
- Building complex workflows

- Java API
- Python API
- Command-line interface

- Any HTTP library or client (eg. curl)
- Easy access from JavaScript (eg. jQuery)
Security

Service
Security

https any client Service
Security

Service

HTTPS
any client

HTTPS
browser only

Loginza.API
Security

HTTPS
any client

Service
allow  deny

HTTPS
browser only

Loginza.API
Security

- HTTPS to any client
- HTTPS browser only
- Request on behalf of user X
- Service
  - allow
  - deny
  - proxy.allow
  - allow
  - deny

Loginza.API
Demo and Downloads

- Demo Testbed
  - Everest + demo services
  - WfMS
  - Service Catalogue
  - [http://mathcloud.org/ru/testbed/](http://mathcloud.org/ru/testbed/)

- Downloads
  - [http://mathcloud.org/dist/](http://mathcloud.org/dist/)
MathCloud Platform as a Service

Core Services

- Service Catalogue
- WfMS
- Security
MathCloud Platform as a Service

Core Services
- Service Catalogue
- WfMS
- Security

Service Hosting
- Everest NG
  (multiuser, dynamic deploy, clustered)
MathCloud Platform as a Service

Core Services
- Service Catalogue
- WfMS
- Security

Service Hosting
- Everest NG
  (multiuser, dynamic deploy, clustered)

Computing Infrastructure
- CGTDC ISA RAS
- Partners
- Service Providers
- Clouds

Resource Management
- Accounting
We invite all interested parties to cooperate and join the project.

http://mathcloud.org
oleg.sukhoroslov@gmail.com