A JavaScript based CernVM-FS Monitor

Marcin Mokrzan
Organic unit: EP-SFT
Supervisors: Jakob Blomer, Gerardo Ganis
About me

Gdynia Maritime University
  faculty of Electrical Engineering,
  Major Computer Control Systems.
Graduate Work:
  Construction of a mobile robot performing
  the task of detecting and following a line.
University of Gdańsk,
  Master of Science in Computer Science
Software developer
What is CernVM-FS?

Provides:
- Scalable
- Reliable
- Low-maintenance
software distribution service

It was developed to assist High Energy Physics (Hep) to deploy software on the worldwide distributed computing infrastructure. CernVM-FS moves the scientific codes to where the data is.
Replica Server (Stratum 1)

New software versions are published at the Stratum 0 at CERN

The data is replicated to multiple Stratum 1 servers to:
- Improve reliability
- Reduce latency

Stratum 1 replica servers are located in:
- Europe
- the U.S
- Asia

The figure, shows the situation for the repositories hosted in the cern.ch domain.
Replica Server (Stratum 1)

New software versions are published at the Stratum 0 at CERN.

The data is replicated to multiple Stratum 1 servers to:
- Improve reliability
- Reduce latency

Stratum 1 replica servers are located in:
- Europe
- the U.S
- Asia

The figure shows the situation for the repositories hosted in the cern.ch domain.
Registered Stratum 0 Repositories

- MICE (mice.egi.eu)
- WENMR (wenmr.egi.eu)
- Physics IBERGRID (phys-ibergrid.egi.eu)
- ATLAS nightlies (atlas-nightlies.cern.ch)
- ILC (ilc.desy.de)
- Blue Brain Project (bbp.epfl.ch)
- CernVM 3 (cernvm-prod.cern.ch)
- Biomed (biomed.egi.eu)
- T2K (t2k.egi.eu)
- ALICE Conditions (alice-ocdcb.cern.ch)
- CALICE (calice.desy.de)
- HERMES (hermes.desy.de)
- H1 (hone.desy.de)
- OLYMPUS (olympus.desy.de)
- XFEL (xfel.desy.de)
- ZEUS (zeus.desy.de)
- ALEPH (aleph.cern.ch)
- CvmFS Configuration (cvmfs-config.cern.ch)
- CERN@School (cernatschool.egi.eu)
Goals of my project

Issue for current monitor:
- Manual: all repositories and replica servers have to be added by hand
- The service requires a dedicated database
- The service runs on a dedicated VM
- Legacy codebase

How to address it?
1) Convert legacy Python code base to an existing JavaScript client library
   - JS client library was designed for the use in a web browser
   - Needs to be adjusted to run on a web server in node.js

2) Query the repositories themselves for monitoring information so that we can drop the need for a database

3) Create new website cvmfs-monitor

→ More accurate and easier to maintain monitoring information
Update CernVM-FS Emscripten Backend

Where I started:
- CernVM-FS Emscripten Backend
- JS client library is designed for the use in a web browser
- Reminder: the emscripten backend allows for running C++-to-JS compiled applications directly from cvmfs

What did I do:
- Change code structure
- Define and reduce external dependencies
- Add units tests
- Create build scripts
- Add modern javascript like ES6 syntax
- Fix bugs (related to crypto, decompression and data encoding)
- Create json file with data
- Create web server for backend (express.js)
- Create new web application (react)
New: cvmfs/js unit tests with mocha.js

```
[marcin@geantbuild cvmfs-emscripten]$ npm run test
> @ test /home/marcin/projects/cvmfs-emscripten
> mocha --require @babel/register ./test/test.js
```

```
masterkeys
   getMasterKeys()
      ✓ MasterKeys is an Array
   addMasterKey()
      ✓ MasterKeys increases internal masterkey list

ts.rsasign
   doPublic
      ✓ deepStrictEqual - doPublic function returns correct value
   bitLength
      ✓ bitLength of signature as BigInteger is less or equal to the bit length of the key
   calculateHex
      ✓ calculateHex - verifyRawWithMessageHex works as expected with the given input
```
Main improvement for the new monitor web application

• Remove dependency on manually curated database of repositories and stratum servers
  • Instead: rely on stratum addresses provided by repository owners in cvmfs itself
    → scales much better for large number of repositories
• Provide json health summary to feed 3rd party monitoring systems

• Reuse CernVM-FS Emscripten code, thereby addressing missing functionality of the current monitor’s Python legacy
  • For instance: currently we cannot monitor repositories that use SHAKE-128 algorithm

• Ready for central OpenShift deployment
  → We remove dependency of custom virtual machine
What json file include

```
{
  "administrator": "ATLAS CernVM-FS Writers",
  "email": "lxcvms-atlas@cern.ch",
  "organisation": "CERN",
  "description": "ATLAS Software",
  "url": "https://cern.ch/atlas",
  "recommendedStratums": {
    "url": "http://cvmfs-stratum-zero.cern.ch/cvmfs/atlas.cern.ch",
    "revision": 53998,
    "publishedTimestamp": 1568041805
  },
  "recommendedStratums": {
    "url": "http://cvmfs-stratum-one.cern.ch/cvmfs",
    "revision": 53998,
    "health": "green",
    "id": 1,
    "publishedTimestamp": 1568041805,
    "name": "CERN",
    "location": {
      "range": {
        3166175232,
        3166240767
      },
      "country": "CH",
      "region": "GE",
      "eu": "0",
      "timezone": "Europe/Surich",
      "city": "Geneva",
      "lt": {
        46.2022,
        6.1457
      },
      "metro": 0,
      "area": 100
    }
  },
  "custom": {
    "_comment": ""
  },
  "health": "green",
  "oldestRevisionStratumOne": 53998,
  "whitelistExpiryDate": "2019-09-27T08:00:00.000Z",
  "download": {
    "catalog": "http://cvmfs-stratum-one.cern.ch/cvmfs/atlas.cern.ch/data/ca/3eb20e10c3e51dbc9f7e5248fc031588c5fff6C",
    "certificate": "http://cvmfs-stratum-one.cern.ch/cvmfs/atlas.cern.ch/data/0b/457ac12225018e0a15330364c20529e15012abX",
    "metainfo": "http://cvmfs-stratum-one.cern.ch/cvmfs/atlas.cern.ch/data/1b/ab16ce33ed1579ce0409cbb31a2e3938ffa86M"
  },
  "rootHash": "ca3eb20e10c3e51dbc9f7e5248fc031588c5fff6",
  "hashAlgorithm": "sha1"
}
```
Structure of the new monitor web app

Registered Repositories

- ALEPH (aleph.cern.ch)
- ALICE (alice.cern.ch)
- ALICE Conditions (alice-ocdb.cern.ch)
- ALICE nightlies (alice-nightlies.cern.ch)
- AMS (ams.cern.ch)
- ATLAS (atlas.cern.ch)
- ATLAS conditions data (atlas-condb.cern.ch)
- ATLAS nightlies (atlas-nightlies.cern.ch)
- AUGER (auger-egi.eu)
- Belle (belle.cern.ch)
- Biomed (biomed.egi.eu)
- Blue Brain Project (bbp.epfl.ch)
- Boss (boss.cern.ch)
What I use to make web application

To create new project: `npx create-react-app cvm-monitor`
Registered Repositories

- ALEPH (aleph.cern.ch)
- ALICE (alice.cern.ch)
- ALICE Conditions (alice-ocdb.cern.ch)
- ALICE nightlies (alice-nightlies.cern.ch)
- AMS (ams.cern.ch)
- ATLAS (atlas.cern.ch)
- ATLAS conditions data (atlas-condb.cern.ch)
- ATLAS nightlies (atlas-nightlies.cern.ch)
- AUGER (auger.egi.eu)
- Belle (belle.cern.ch)
- Biomed (biomed.egi.eu)
- Blue Brain Project (bbp.epfl.ch)
- Boss (boss.cern.ch)
- CALICE (calice.desy.de)
- CERN@School (cernatschool.egi.eu)
- CernVM 3 (cernvm-prod.cern.ch)
- Chipster (chipster.egi.eu)
- CLICdp (clicdp.cern.ch)
- CMS (cms.cern.ch)
- CMS Nightlies (cms-ib.cern.ch)
# ATLAS (atlas.cern.ch)

## Stratum 0

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Revision</td>
<td>54039</td>
</tr>
<tr>
<td>Oldest Stratum 1 Revision</td>
<td>54037</td>
</tr>
<tr>
<td>Last Modified</td>
<td>10th September 2019 4:04:20 pm</td>
</tr>
<tr>
<td>Whitelist Expiry Date</td>
<td>2019-09-27T08:00:00.000Z</td>
</tr>
</tbody>
</table>

## Stratum 1

### CERN

- **Revision:** 54039
- **Last Modified:** 10th September 2019 4:04:20 pm
- [http://cvmfs-stratum-one.cern.ch/cvmfs](http://cvmfs-stratum-one.cern.ch/cvmfs)

### Fermilab

- **Revision:** 54037
- **Last Modified:** 10th September 2019 3:01:22 pm
- [http://cvmfs.fnal.gov/cvmfs](http://cvmfs.fnal.gov/cvmfs)

### SDCC by RACF at BNL

- **Revision:** 54039
- **Last Modified:** 10th September 2019 4:04:20 pm
- [http://cvmfs.raclbnl.gov/cvmfs](http://cvmfs.raclbnl.gov/cvmfs)

### STFC Rutherford Appleton Lab

- **Revision:** 54037
- **Last Modified:** 10th September 2019 3:01:22 pm
- [http://cvmfs-wlcg.gridpp.rl.ac.uk/cvmfs](http://cvmfs-wlcg.gridpp.rl.ac.uk/cvmfs)
How I get geographic coordinates
Web application conclusion

Automatically download data for repository:
- Using metainfo: the information snippet maintained by repository owners
- Using Public keys
- Monitoring only owner-recommended stratum ones

We need cooperation from repository owner:
- Require (valid!) metainfo published
- Require set of public keys
- Require access to the stratum ones

But when any of this problems occur an error message is displayed
- “Error: metainfoHash is undefined; Without metainfo we cannot proceed”
- “Error: Unable to verify whitelist”
Responsive Web Design
Status and Next Steps

All updated server code:
https://github.com/cvmfs-contrib/cvmfs-emscripten/tree/feature-node

Web application code:
https://github.com/mormar/cvm-monitor

Previous code:
https://github.com/cvmfs-contrib/cvmfs-emscripten/tree/master

Move deployment of the web application on CERN OpenShift
Thank you

Jakob Blomer
Gerardo Ganis
Johannes Heinz
Simone Mosciatti
Andrea Stano