

Event Index Core *(Hadoop & HBase)*



- Import
- Shibboleth
- Important Missing Pieces

Andrea Favareto
Julius Hrivnac
Fedor Prokoshin
Grigorij Rybkin
Rainer Toebbicke
Ruijun Yuan



Import

- Runs smoothly
- Procedure:
 - All conf files go to **DatasetsConf**
 - checkConf.sh moves MC datasets to **DatasetsConf.mc**
 - And big MC datasets (>1000 files) to **DatasetsConf.big**
 - mergeConsumers.sh merges them into **DatasetsConf.merged**
 - Errors are copied by hand to **DatasetsConf.error**
 - cpFromHar.sh gets them from HAR to **DatasetsConf.har**
 - When everything fails, datasets go to **DatasetsConf.problem**
 - And are handled manually
- All new datasets are (immediately or within a day)
 - Checked for duplicated events
 - Checked for correctness wrt Ami
 - Checked for correctness of Trigger info
 - Fed into datasets overlap table
 - ? Should we fill Trigger overlap table, other Trigger statistics ?
- New datasets are uploaded to HBase 4x a day (Rainer)
 - ? Is status=bad upload working right ?

DATASETS

=====

EI	2009	526	
EI	2010	4518	
EI	2011	3944	
EI	2012	4400	
EI	2013	525	
EI	2014	492	
EI	2015	23631	(23630)
EI	2016	3535	(3162)
MC	2009	0	
MC	2010	0	
MC	2011	0	
MC	2012	3	
MC	2013	0	
MC	2014	0	
MC	2015	53040	(50929)
MC	2016	6	

previous WS





Shibboleth

- Shibboleth problem is solved
 - All commands can run with `-cookies cookies.txt` option on non-CERN machines
 - Those commands are now deployed (incl. CVMFS)
- Shibboleth cookies creation works fine on SLC machine
 - Will be included in CVMFS soon
- Two kind of cookies can be generated
 - Kerberos
 - Certificate
- Cookies are useful only on machine, there were created on
- Cookies generation sw can be installed on Fedora (24)
 - Quite complicated procedure requires root access



Shibboleth cookies creation

```
# a kerberos cookie: ← will go into setup
cern-get-sso-cookie --krb -r -u https://atlas-event-index.cern.ch -o
~/private/ssocookie.txt
# or a personal cookie:
# get myCert.p12 from http://cern.ch/ca
openssl pkcs12 -clcerts -nokeys -in myCert.p12 -out ~/private/myCert.pem
openssl pkcs12 -nocerts -in myCert.p12 -out ~/private/myCert.tmp.key
openssl rsa -in ~/private/myCert.tmp.key -out ~/private/myCert.key
rm ~/private/myCert.tmp.key
chmod 644 ~/private/myCert.pem
chmod 400 ~/private/myCert.key
cern-get-sso-cookie --cert ~/private/myCert.pem --key ~/private/myCert.key -r -u
https://atlas-event-index.cern.ch -o ~/private/ssocookie.txt
```



Shibollet cookies creation sw

```
cat >/etc/yum.repos.d/cern.repo << EOF
[cern]
name=cern
baseurl=http://linuxsoft.cern.ch/cern/centos/7/cern/x86_64
enabled=1
gpgcheck=0
EOF
#
dnf install perl-Module-Install
dnf install perl-Net-IP
#
wget
https://dl.fedoraproject.org/pub/fedora/linux/releases/23/Everything/source/SRPMS/o/ocsinventory-agent-2.1.1-6.fc
23.src.rpm
rpmbuild --rebuild ocsinventory-agent-2.1.1-6.fc23.src.rpm
cd ~/rpmbuild/RPMS;
dnf install -y noarch/perl-Ocsinventory-Agent-2.1.1-6.fc24.noarch.rpm
x86_64/ocsinventory-agent-2.1.1-6.fc24.x86_64.rpm
#
dnf install CERN-CA-certs cern-get-keytab cern-get-sso-cookie cern-wrappers lcm.noarch
#
addusercern afsid
```



Important Missing Pieces

- Commands with Shibboleth protection 
 - Working, back in CVMFS
 - Configuration should be more user-friendly
- Dataflow logistics
 - HBase upload is a black-box handled by Rainer
- Re-factoring of EL to make it more *flexible* and *user-friendly*  *Grigorij is looking at that*
 - It works, it's fast
 - Should be more modular (choosing strategy & implementation,...)
- Update TWikies
 - Up-to-date info is always available via CLI help
 - and TagConvertor Web Home
 - and Web Service GUI
- Exploratory work
 - Using Journal etc to make predictive optimisation
 - Interface with new Hadoop formats (Zbysek...)
 - ...



Info

Web Service: <https://atlas-event-index.cern.ch/ElHadoop>

Documentation & Distribution: <https://atlas-event-index.cern.ch/doc>

Sources: svn+ssh://svn.cern.ch/repos/atlasoff/Database/TAGHadoop/TagConvertor

AFS: /afs/cern.ch/project/jps/repos/atlas-eies/TagConvertor

CVMFS: \$ lsetup eiclient