PRAGMA STUDENT EXPEDITION CHALLENGE

Challenge 1 – Build up Lifemapper Biodiversity Infrastructure

Aravindh and Ann

Agenda

- Setup time
- Experience
- Good to have info
- Suggestions
- Amazon EC2

MacOSX Setup Time

Component	Setup Time
Install VirtualBox **Install Lifemapper server roll**	~ 30 minutes
Seeding the data pipeline	~ 1 hour 26 minutes
Machine Specification:	Operating System: Mac OSX 10.9 (Mavericks) Base System Memory: 8 GB Virtual Box Version: 4.3.18 r96516 Virtual Machine Memory: 4 GB

Windows Setup Time

Component	Setup Time
Install VirtualBox **Install Lifemapper server roll**	~ 30 minutes
Seeding the data pipeline	~ 1 hour 18 minutes
Machine Specification:	Operating System: Windows 8.1 Base System Memory: 8 GB Virtual Box Version: 4.3.18 r96516 Virtual Machine Memory: 4 GB

Test Outputs – Check Jobserver

```
[lmwriter@fe ~]$ python2.7 /opt/lifemapper/LmWebServer/scripts/createTestUser.py
Successfully created user
```

```
[lmwriter@fe ~]$ python2.7 /opt/lifemapper/LmWebServer/scripts/checkJobServer.py
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     {'epsgcode': '4326', 'displayname': 'Test Chain56946.7033215', 'name': 'Test
points56946.7033215', 'pointstype': 'shapefile'}
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Test Chain56946.7033215
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Removing: /tmp/points-5045647.dbf, Success: None
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Removing: /tmp/points-5045647.shp, Success: None
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Removing: /tmp/points-5045647.shx, Success: None
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Removing: /tmp/points-5045647.prj, Success: None
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     Removing: /tmp/points-5045647.qix, Success: None
16 Oct 2014 09:52 MainThread.log.warning line 149 WARNING  Database connection is None! Trying to re-open ...
Closed/wrote dataset /share/lmserver/data/UserData/unitTest/4326/Layers/pt1.shp
creating index of new  LSB format
16 Oct 2014 09:52 MainThread.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 1 in MAL
Occurrence job id: 1
Model job id: 2
Projection job id: 3
The job has been deleted
The job has been deleted
The job has been deleted
16 Oct 2014 09:52 MainThread.log.info line 126 INFO      Removed 1 experiments for occurrence set 1
16 Oct 2014 09:52 MainThread.log.info line 126 INFO      Removed occurrence set 1
```

Test Outputs – Check Webserver

```
[lmwriter@fe ~]$ python2.7 /opt/lifemapper/LmWebServer/scripts/checkLmWeb.py
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/experiments
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/projections
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/rad/
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/rad/experiments
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/rad/layers
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios/4/atom
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios/4/eml
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios/4/html
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios/4/json
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/scenarios/4/xml
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/ascii
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/atom
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/eml
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/html
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/json
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/kml
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/raw
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/tiff
16 Oct 2014 09:54 MainThread.log.debug line 93 DEBUG    Url: http://fe.public/services/sdm/layers/64/xml
```

Test Outputs- Data pipeline

```
[lmwriter@fe ~]$ python2.7 /opt/lifemapper/LmDbServer/pipeline/localpipeline.py
Begin Pipeline with algorithms ['BIOCLIM', 'GARP_BS', 'ATT_MAXENT'], modelScenario CRU, projectionScenarios ['hadley_a1f', 'hadley_a2c', 'hadley_b1a']
16 Oct 2014 09:56 MainThread.log.debug line 93 DEBUG    Deleted /opt/lifemapper/pipeline.lm.die
16 Oct 2014 09:56 MainThread.log.info line 126 INFO     Pipeline starting at process id 13393, with kill file /opt/lifemapper/pipeline.lm.die
16 Oct 2014 09:56 MainThread.log.info line 126 INFO     processrunner_email opened databases
16 Oct 2014 09:56 MainThread.log.debug line 93 DEBUG    Reset 0 queued (90) jobs to status (1)
16 Oct 2014 09:56 MainThread.log.info line 126 INFO     chainer opened databases
16 Oct 2014 09:56 MainThread.log.info line 126 INFO     troubleshooter opened databases
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1 records for 1000053 (starting on line 1)
16 Oct 2014 09:56 processrunner_email.log.debug line 93 DEBUG    Pulled 0 SMTP Jobs
16 Oct 2014 09:56 processrunner_email.log.debug line 93 DEBUG    Pulled 0 local Jobs
16 Oct 2014 09:56 MainThread.log.debug line 93 DEBUG    ********* Main loop, sleeping 30s *********
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 2 in MAL
16 Oct 2014 09:56 troubleshooter.log.debug line 93 DEBUG    Moved 0 dependent SDM jobs and 0 dependent RAD jobs to ready (status=1)
16 Oct 2014 09:56 troubleshooter.log.info line 126 INFO     Moved 0 dependent SDM jobs; 0 dependent RAD jobs
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1 records for 1000225 (starting on line 2)
16 Oct 2014 09:56 chainer.log.info line 126 INFO     Revisit this: Unknown taxon for key 1000225
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1 records for 1000329 (starting on line 3)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 3 in MAL
16 Oct 2014 09:56 troubleshooter.log.info line 126 INFO     Deleting 0 experiments for anon
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1 records for 1000410 (starting on line 4)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 4 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 5 records for 1000431 (starting on line 5)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 5 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 41 records for 1000447 (starting on line 10)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 6 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1442 records for 1000461 (starting on line 51)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 7 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 2 records for 1000468 (starting on line 1493)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 8 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 2 records for 1000470 (starting on line 1495)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 9 in MAL
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG    Returning 1 records for 1000477 (starting on line 1497)
16 Oct 2014 09:56 chainer.log.debug line 93 DEBUG     inserted job to write points for occurrenceSet 10 in MAL
16 Oct 2014 09:56 MainThread.log.debug line 93 DEBUG    ********* Main loop, sleeping 30s *********
```

Installation Experience

Overall simple

But

......

Issues encountered

- BIOS Settings on Windows machine
- Compatible Virtual Box Version
- Corrupt rocks distribution image
- No idea on how long to wait for seeding to complete

Error messages

- /usr/bin/perl: error while loading shared libraries: libfreebl3.so: cannot open shared object file: No such file or directory

- /usr/bin/p11-kit: error while loading shared libraries: libfreebl3.so: cannot open shared object file: No such file or directory

Good to have Info

- Minimal requirements for hardware
- Software requirements for major OS (OSX, Windows, Ubuntu etc), for example, tested version of Virtual Box that works
- Methodology to setup user's own dataset
- Possibility of switching databases
- Dataformat expected by Lifemapper

Suggestions

- Ganglia shows what is the status of the cluster but does not warn you on threshold breaches
- Nagios monitoring tool is used to monitor subsystems and sent alert notifications to the user on threshold breaches
- Nagios available as a plugin for Ganglia

Amazon Machine Image