

Galaxy Admin workshop

Questions:

- Does Galaxy scale?
 - <https://docs.galaxyproject.org/en/latest/admin/scaling.html>
 - <https://docs.galaxyproject.org/en/latest/admin/cluster.html>
- Other frontends were mentioned. Is that going to be discussed? E.g. galaksio?
 - <https://github.com/fikipollo/galaksio>
 - <http://www.refinery-platform.org>
- This documentation for forwarding covers nginx and apache, but the nginx is from a custom ppa. How can this be achieved for e.g. RHEL7?
 - <https://depot.galaxyproject.org/yum/>
- There's a lot of chatter for securing the galaxy user from being able to write to code directories, but wouldn't this be safer using something like selinux or apparmor to specify only certain writable directories?
 - Having separate directories is also useful for backup only config and data
- Why does the 'shed_tool' installation need to be linked to a (SQL) database, which table are created and why? Aren't the dedicated .xml files enough for that kind of tasks?
 - Unfortunately not, for historical reasons tools from TS are registered in the database too.
 - Sometimes uninstalling a tool at a certain version, also removes other versions of this tool from shed_tool_conf.xml, so the database can help you recover from this bug
 - This can be a SQLite file, though (we use this on usegalaxy.org to install tools in CVMFS) see the `install_database_connection` option.
- How to determine the owner of the dataset based only on its absolute path:
/srv/galaxy/data/000/dataset_1.dat?
 - ```
SELECT hda.id, u.email FROM galaxy_user u JOIN history_dataset_association hda ON u.id = hda.user_id JOIN dataset d on hda.dataset_id = d.id WHERE d.id = 1;
```
- About compression of nginx static content: Is there any option to compress file reaching some (size) threshold instead of doing type filtering? Size of these static files can vary a lot and for the smallest ones, compress them would be less efficient (use cpu power without or with a barely noticeable size reduction).
  - [https://nginx.org/en/docs/http/nginx\\_http\\_gzip\\_module.html#gzip\\_min\\_length](https://nginx.org/en/docs/http/nginx_http_gzip_module.html#gzip_min_length) may do what you ask
- I've noticed a potentially nasty bug in forwarding/conda installation. If a conda dependency takes too long to install, the associated connection to galaxy may die, and subsequently render the entire conda environment invalid. Any ideas about workarounds?
  - It should continue installing even after the connection is interrupted, you can view installing repositories from the admin interface. But yes, this is a nuisance.
- Does ProFTPD support encrypted connections?
  - Yes, both FTPS <http://www.proftpd.org/docs/howto/TLS.html> and SFTP [http://www.proftpd.org/docs/contrib/mod\\_sftp.html](http://www.proftpd.org/docs/contrib/mod_sftp.html)
- Can you share admin scripts with us please?
  - Can you be more specific? Which admin scripts?
- Using both supervisord and systemd at once seems kind of redundant. Is there any kind of serious advantage to this? Would you like a copy of my systemd galaxy.service file?

- With supervisor you can restart just a single or a group of web/job handlers, is that possible with systemd?
  - Why would you want to do this in supervisord, when uWSGI can do this itself?
    - <https://uwsgi.readthedocs.io/en/latest/MasterFIFO.html>
      - The various job and workflow handlers are (for now) separated processes not managed through uWSGI
        - Systemd is really good at handling lots of separately grouped processes.
  - Systemd wasn't widely available back when we adopted supervisor as a best practice but it's not a bad choice to use only systemd. Supervisord provides a couple other things that might be beneficial - unprivileged installation if needed, and granting unprivileged users control (systemd can do this too but in the version of systemd currently in EL 7, you can't grant it on a per-service level).
- There doesn't appear to be a reports.ini.example file shipped in current galaxy. Looks like there is from the 16.01 release - is this an old example?
  - Right, the file is now in config/reports.yml.sample
- It is good to know so many possibilities to deploy and admin Galaxy. In my case, I just use Galaxy as a place for non-CS users to learn to use the established bioinformatics workflows. Therefore I focus on the containerisation of Galaxy. Does Galaxy work well with Kubernetes? Does the existing Docker image work well with Kubernetes? Is it stable and reliable enough now? Are there any limitations (like data management, etc.)?
  - There is a K8s job runner documented in job\_conf.xml.sample\_advanced. I am not sure how much use it has seen but it's worth checking out.
- We have seen that supervisord,systemctl and nginx can be use to run others (than Galaxy) applications. Is it possible to share uwsgi to serve others website the same way ? What do we need to change in the configs files then ?
  - I don't think that's the case for uWSGI
- During first Ansible exercise we have use a static plain text API key set in the role/default/main.yaml file, how can we dynamically generate this API key, in a encrypted/hashed way with Ansible instead ? Same question for password.
  - Look at Ansible Vault: <https://docs.ansible.com/ansible/2.4/vault.html>
- Looks like the GalaxyKickStart still uses the old reports\_uwsgi.ini file, but that doesn't exist at all in the modern galaxy config (<https://github.com/ARTbio/GalaxyKickStart/pull/292>).
- Can you add which is the use case for the CLI job runner in the "Using heterogeneous compute resources" slides?
  - If your Galaxy server has a shared filesystem with a cluster but does not have permission to interact directly with the scheduler. Pulsar can also be used for this purpose.
- Are there any resources available (with working example and details commands) about the slides recommended way of doing a git merge for a Galaxy version update in case of local modifications ?
  - Any git documentation on merging (or if you prefer, on rebasing) would apply
- Will it be hard to identify which process correspond to each uWSGI web handler or mule in 18.01?
  - By default, the log message format now includes: [p:<PID>, w:<WORKER\_ID>, m:<MULE\_ID>]
  - You can separate the log files using the custom filename\_template parameter of FileHandler handlers
  - This is documented and will be under the "Logging Configuration" section of <https://docs.galaxyproject.org/en/latest/> once #5105 is merged
  - I guess also uwsgitop can help

## Missing/Interesting topics:

- What are mules for uwsgi?
  - <https://docs.galaxyproject.org/en/master/admin/scaling.html> and the newest version (not built yet) at <https://github.com/galaxyproject/galaxy/pull/5105/commits/36ecca47e7c859a980c4e775db0ebedbb1c02cef>
  - <https://galaxyproject.github.io/dagobah-training/2018-oslo/whatsnew/18.01.html>
- DRMAA
  - This will be covered during the “Connecting Galaxy to a compute cluster” session
- Uwsgi zerg mode is nice  
<https://github.com/galaxyproject/dagobah-training/blob/2018-oslo/sessions/10-uwsgi/ex2-zerg-mode.md>
- Documentation for setting up Galaxy production server on Red hat please?
- General operational stuff - backups, logfile trimming/rotation
  - I feel that's a bit out of scope for this course, unfortunately
  - I mean specifically wrt. Galaxy backups (ie. what components of it need to be backed up). Logfiles - true, that is a general problem.
- Security overview/policy? Recap maybe on approach to ensure user's data is not shared by default
- Webhooks to customize Galaxy GUI
- Galaxy Tours
- Different Docker setup where docker is not run as root, and where sysadmins have to use static NAT from the docker to the host machine using iptables :)
  - <https://fralef.me/docker-and-iptables.html>
- How much work to make new GIE from existing web app?
- Galaxy pages?

[https://raw.githubusercontent.com/bgruening/galaxy-universe/master/images/main\\_scheme.png](https://raw.githubusercontent.com/bgruening/galaxy-universe/master/images/main_scheme.png)

## Lightning talks (Thursday afternoon)

- Kjell Petersen (Elixir Norway) - Demo of NeLS - Norwegian e-Infrastructure for Life Sciences.
- Nikolay Vazov (USIT, University of Oslo) - Lifeportal in a nutshell (Friday, after 9.15)
- Gary Robertson (HZI) - Tool demonstration usecase / non-shared filesystem alternative approach

## Dinner Suggestions

- [Matkjelleren on Paleet](#) (Ate there on Monday)
- [Istanbul](#) (Good food, enough to feel full, cheap)
- [VIPPA](#) (reservation for large groups)
- [Punjab Tandoori](#) (If you like Indian spices, prices are ok)