Prototype of a ^{58}Fe EMPIRE evaluation

Part II: GitLab with Continuous Integration demo and thoughts

M. Herman, M. Fleming

WPEC Subgroup 49 WebEx meeting
13 May 2020

Background and motivation

- Storing the ENDF-6 (or GNDS) files in a version control system is what library projects have been doing
- Supplemented with description in a detailed report or other peer-reviewed publication
- Further supplemented with detailed MF1 comments
- However, we often have some issues:
  - Reports don't provide all the data
  - MF1 can be very detailed but is often not – and in the best scenarios we have a lot of work to reproduce an evaluation
  - Ultimately, some evaluations cannot be reproduced today that we would like to build upon
- Question: can we at least version control all the inputs/data/codes/processes that went into making a file?

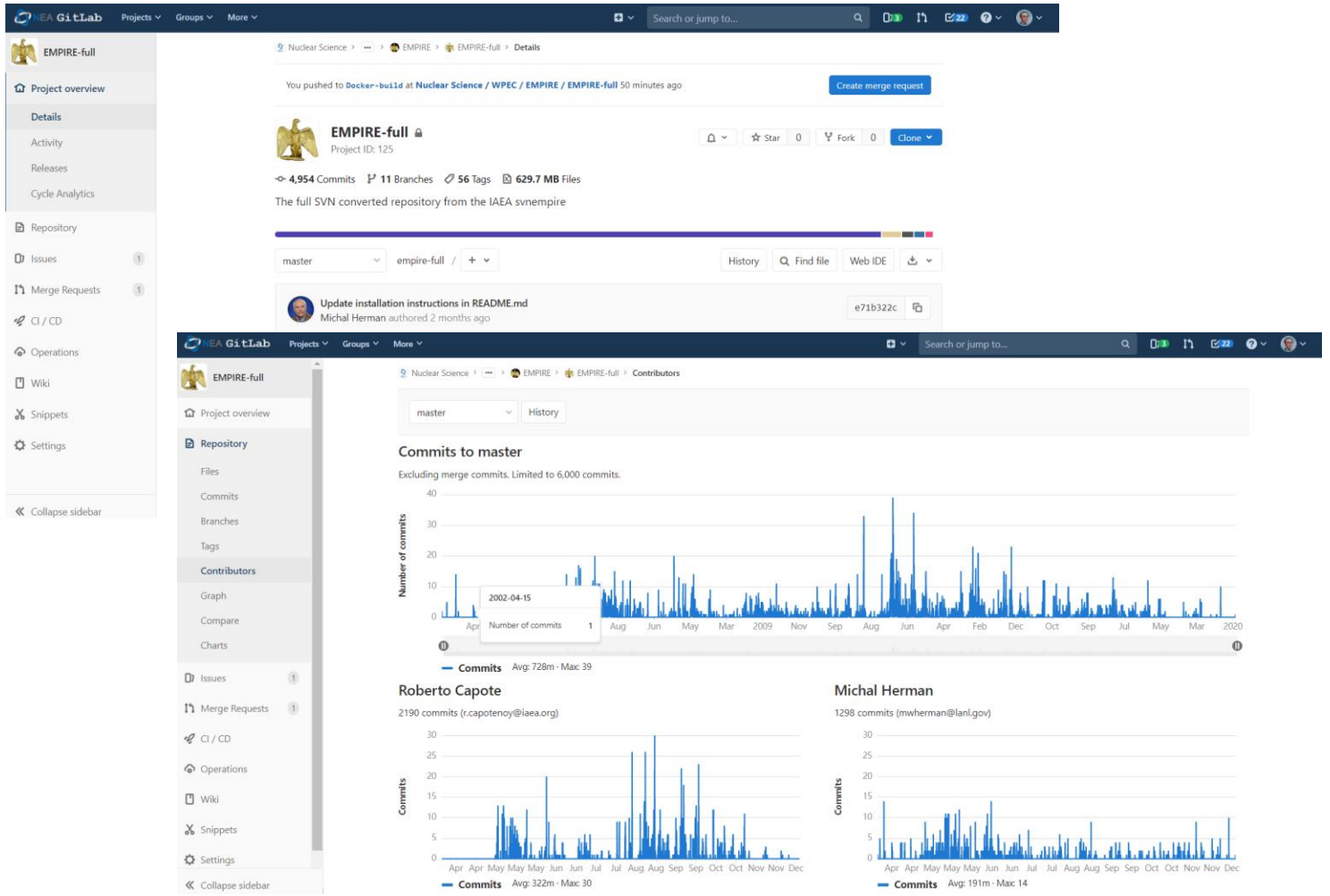
A version control system and collaborative platform

- We have to make some technology choices to start
  - The GNU git version control system has become a/the dominant tool for the kinds of VC requirements that nuclear data evaluation has
  - Extremely popular and mature platforms have been developed around git, with B$+ companies actively developing them
    - **GitHub** (acquired by MS) has the most popular cloud system with third-party CI
    - **GitLab** has open-source version for self-hosted solutions and integrated CI
  - NEA launched a GitLab server in late 2018 with early projects including GNDS, EXFOR compilation and VaNDaL
  - If you require access please me know

https://git.oecd-nea.org/

EMPIRE SVN-converted git and CI

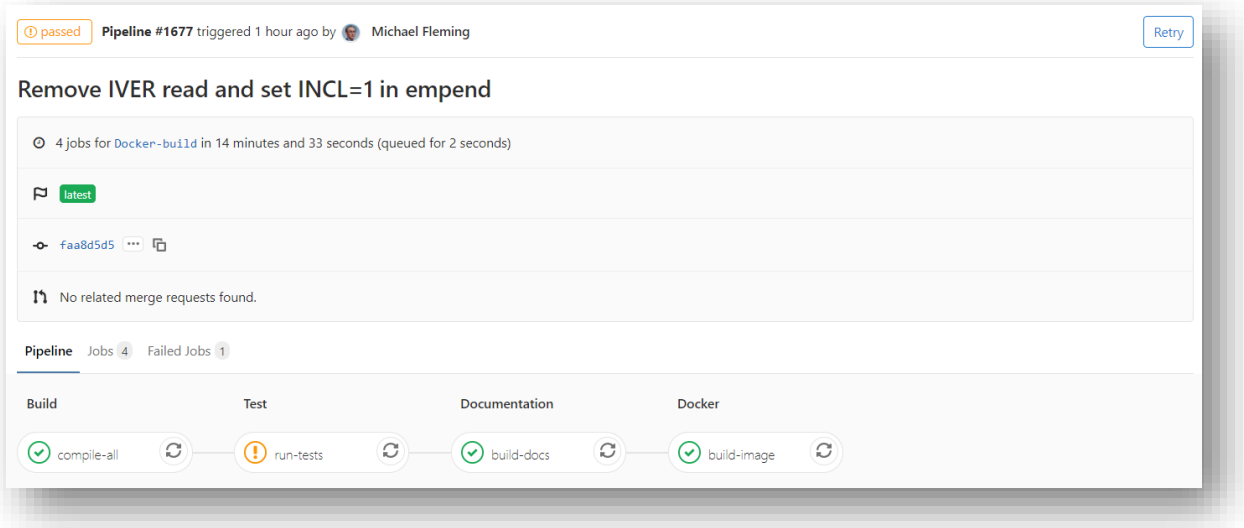
- The IAEA hosts the EMPIRE SVN repository and allowed a full repository conversion to git that is in the NEA GitLab (private)
- Contains some 2 decades of contributions with ~5000 commits
- Even with some databases in the VCS, it remains moderately sized at ~630 MB compressed

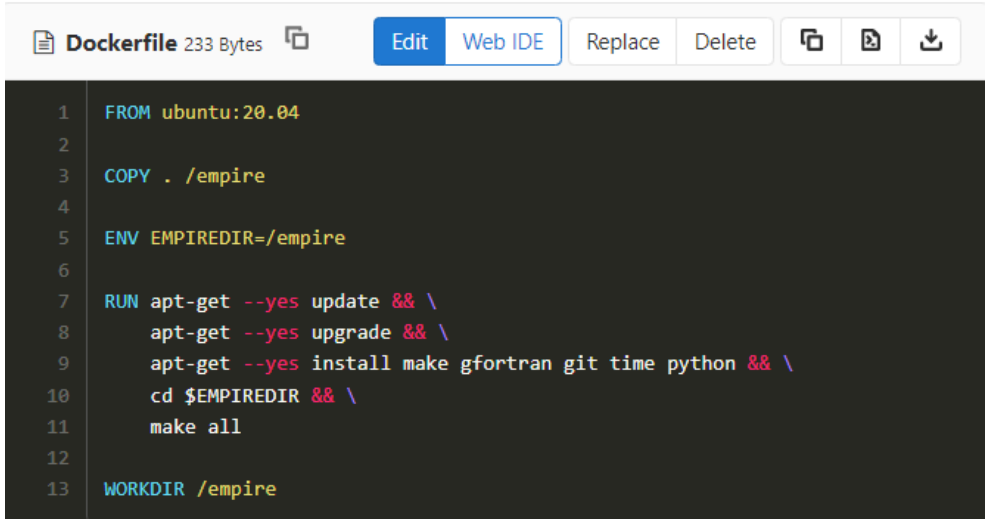


https://git.oecd-nea.org/science/wpec/empire/empire-full

EMPIRE CI system

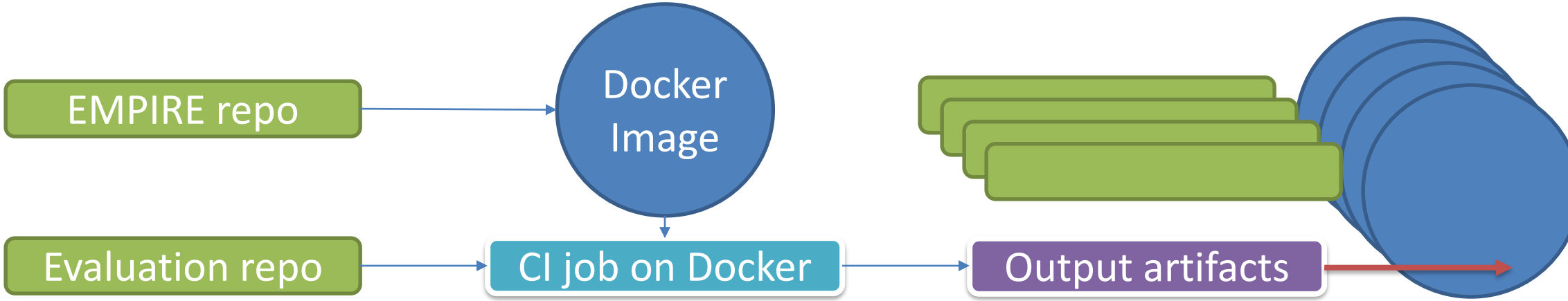
- GitLab provides a very convenient and powerful continuous integration system to automatically perform tasks on any change
  - Typically we compile, run unit/integration/system/regression/etc tests and perform any deployment tasks
  - The standard EMPIRE tasks have been integrated into a CI within Docker containers
  - A Docker build is integrated as a final step after tests to provide images for evaluation workflows – also suitable for general distribution – with very simple requirements that can be replicated on many OSs





Version control and CI of evaluations

- Apart from the code repository, the inputs for an evaluation of ^{58}Fe are stored
- A CI was added to automatically run these inputs with the current EMPIRE version (although any tagged version could be chosen)
- Alternatively evaluations could be linked via **submodules**, but Docker provides a complete run environment that has been tested



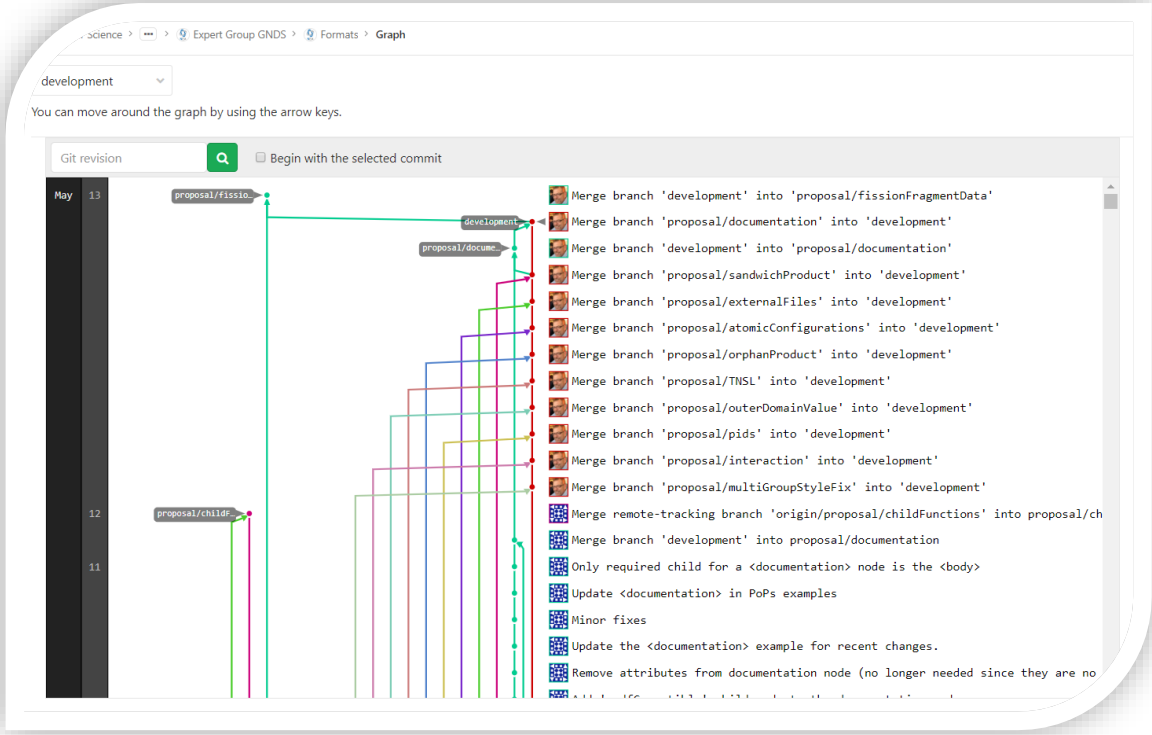
https://git.oecd-nea.org/science/wpec/sg49/Evaluation-inputs

Some comments

- EMPIRE is relatively small at <700 MB compressed, but it contains multiple separate projects:
  - RIPL
  - PREPRO
  - ENDF utilities
  - Zerkin programs
  - Various other tools common to other projects…
- These create a parallel maintenance effort, fixes must be translated back and forth (and often aren't)
- Recommendations:
  - **Databases need rigorous version control systems that we can draw from directly within automatic tools**
  - **Utilities and shared codes (particularly open ones) should be in accessible VCS for automatic use**
  - **Once done, submodules and APIs can be used to greatly improve processes**

Parting observations

- Docker containerisation trivialises the deployment to multiple operating systems/compilers and gives users more so they can address their own local system issues
- GitLab/Hub requires very little effort to engage with and makes use of code systems tremendously better than no VCS
- Collaboration is encouraged and easy, without relinquishing any control over a software project
- As a distributed VCS, git repositories can be re-hosted at any location that the owners wish – private spaces can be used to prepare before release



GNDS repository showing a host of contributions managed by the Chair

https://git.oecd-nea.org/science/wpec/gnds/formats