ECCE Software “Lessons Learned”

Joe Osborn

What Worked

- Diversifying to run at multiple locations, federated approach to job submission
    - BNL, JLab, Bates, ORNL, OSG (others?)
        - OSG became difficult to use at times, but didn’t slow us down because we had diversity of sites
            - One site going down doesn’t slow down job production
- Single system that was responsible for running jobs (git repo with production scripts)
- Office hours were very useful - several people regularly called in to get support/help
- Did not invest time to build software infrastructure, could get started immediately without overhead and with limited person power
    - Got software up and running quickly in March-May 2021, with focus on distributing knowledge to users through workshops

What Didn’t Work

- Limited person power to provide user support and help users develop their own analysis modules
- Single production system (git) also had downsides vs. a more robust system e.g. Panda
- Event evaluator was a catch all and had lot of overhead, often not specific enough for some analyses and contained a lot for others
    - Common flat tree output? How do we define it? Is it really useful as a general tool?
        - Event Evaluator tried to fit this role but required many patches/updates and was often a lot of overhead
- Software repository organization was really bad. Supporting three builds (sPHENIX, EIC-general, ECCE) was very confusing to users and made CI builds very difficult
    - ECCE specific repos should be broken off and maintained separately in future
    - EIC-general is used by entire EIC community
- Lack of CI punished experts and caused human hours to be put into maintaining builds/checking compilation etc.
    - Note - coresoftware main repo is CI integrated. Many EIC related PRs were submitted here
- Time frame was really limited during proposal process - 2 year process stuffed into 9 months

Proposed Survey Questions

1. Did you use ECCE software to develop physics analysis or detector development?
2. Were there adequate resources/tutorials/workshops to help you get started using the software?
    1. If the answer is no, what resources did you use and what would you have liked to have seen?
3. Did you use Mattermost or Discourse for communication with colleagues/experts?
4. Did you analyze centralized production data from the first simulation campaign?
5. Did you analyze centralized production data from the second simulation campaign?
    1. If yes to either of these questions, were you slowed down at all by the data arriving or having to learn the framework?
6. Did you develop your own analysis modules or use one of the central “evaluator” modules?
    1. If evaluator, which one(s) did you use?
7. (Open ended) What did you find easy to use with the ECCE software?
8. (Open ended) What did you find difficult to use with the ECCE software?
9. Based on the previous question, what would you like to see implemented in the software to make it easier to use?