ECCE Computing
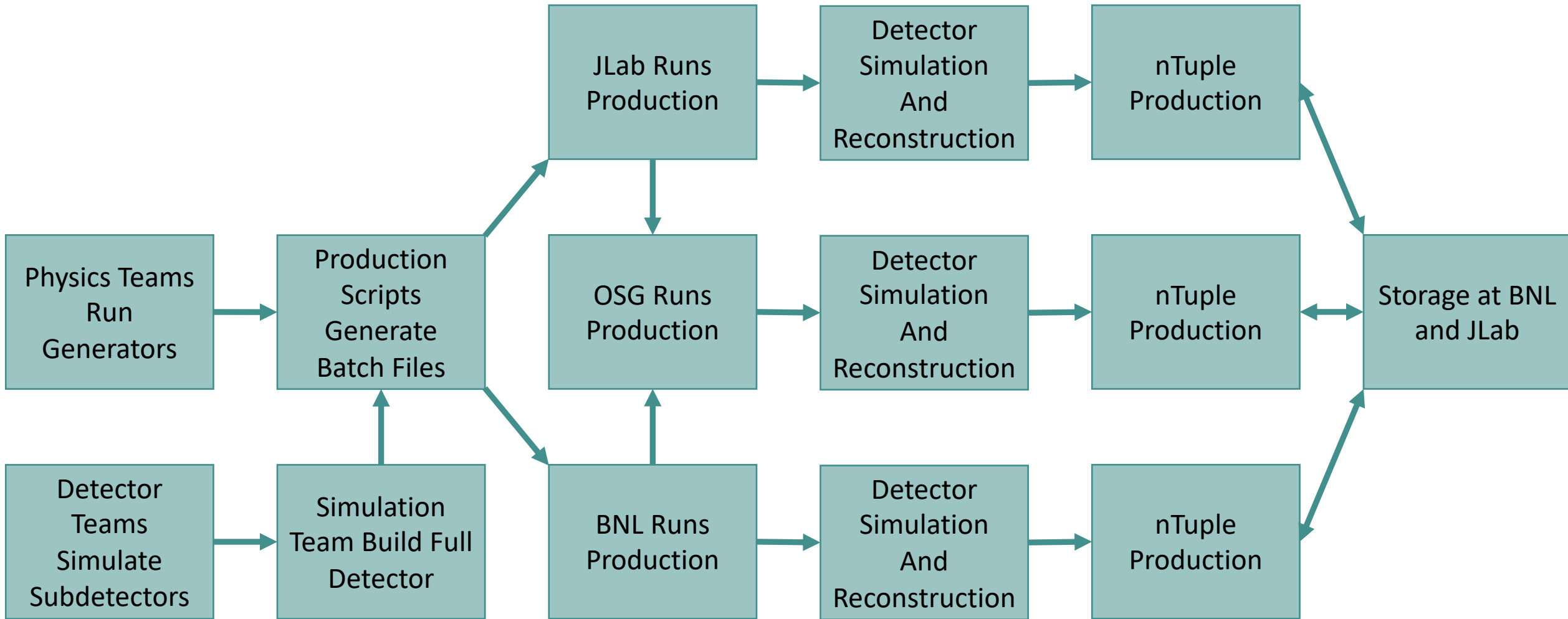
Current Status & Predictions
Cameron Dean, on behalf of ECCE

07/28/2021
EIC Software & Computing Meeting
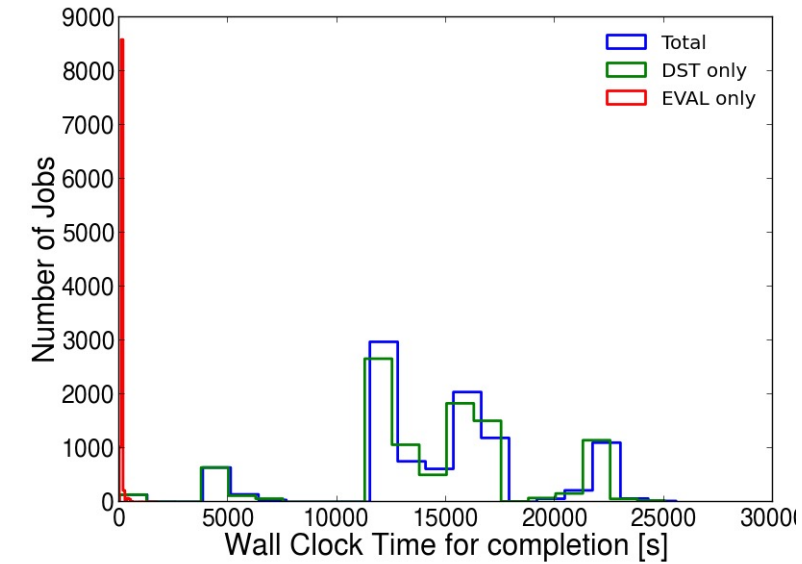
Production Workflow

Production Statistics and Notes

Planned simulation campaigns	2
Predicted events per campaign	120M - 160M
Typical event size	200kB
Typical event generation time	7s
Total storage per campaign*	30 TB
Typical job memory size	< 1.5 GB
Current events simulated in campaign one	101.3M

*We have not decided if we want to keep the first campaign raw data when we run the second campaign

- First simulation campaign is under way, "June 2021 Concept"
- Second detector design is being finalized today, "July 2021 Concept"
- Production sites show > 90% job success rate (OSG > 98% success rate)
- Decided to separate simulation and reconstruction from nTuple production (former is time/resource intensive, latter is in current development)
- nTuple production automatically produces revisions and runs quickly (~2.5ms/event)
- To-do: add job monitoring and resubmissions (in beta-testing)

User Work Methods

- Many users across the world, not everyone joins with active BNL membership
  - Not feasible to get everyone accounts in proposal timescale
- Singularity (and VirtualBox) is used to distribute daily software stacks (and simulations)
- S3 (BNL) and xrootd (JLab) protocols used to distribute data
  - minIO client and read-access keys are distributed from ECCE stack
- Mattermost and Discourse used to rapidly communicate (non-BNL users are issued special access keys for manual authentication)
- Users are encouraged to use low-volume nTuples over DSTs
  - Keeps bandwidth to a minimum
  - Many physics plots can be made without large data processing