ProtoDUNE Run II Database Challenges

Norm Buchanan

Colorado State University

Colorado State University

DUNE Database Workshop (Virtual)
July 2, 2021

Goal of ProtoDUNE DBs

Really two goals:

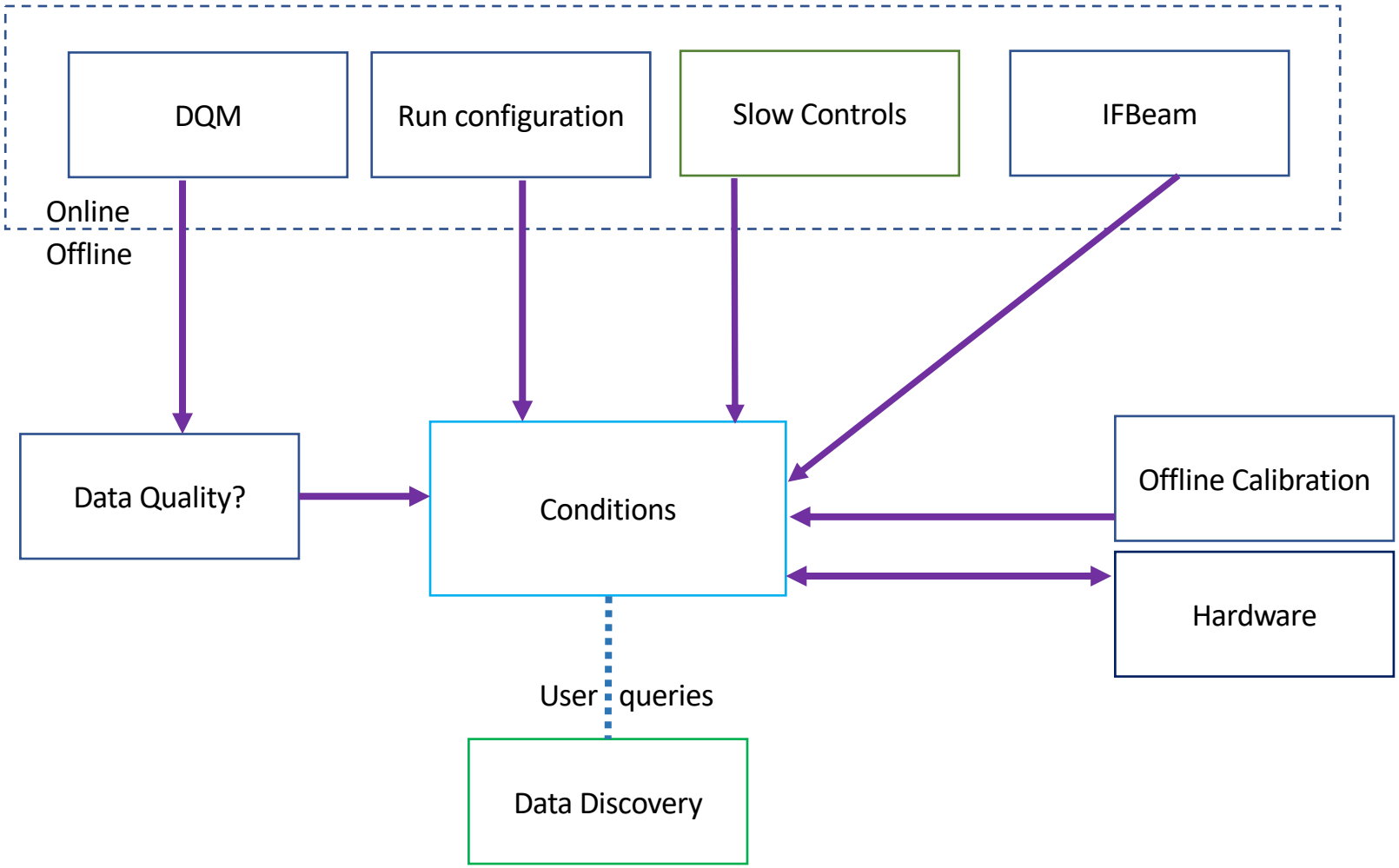
1. Provide metadata for offline analysis and related tasks
2. Test the DB systems in preparation for DUNE.

The first goal must be achieved in full even if this means initially exercising parts of the full DB system.

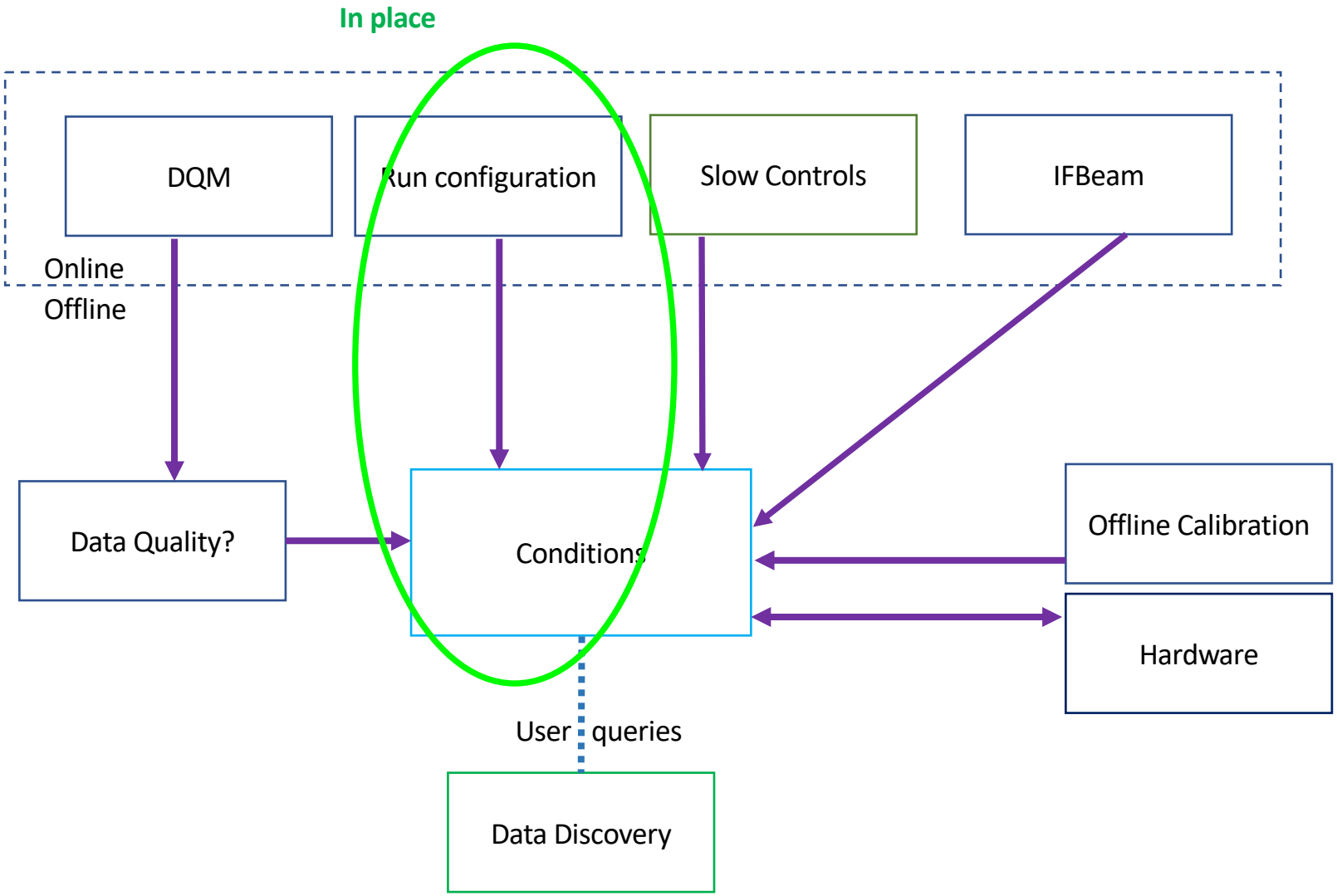
We will aim to have the full database system implemented for physics running in the summer of 2022.

Clearly, there will be some evolution to the DB system for DUNE but what we do for ProtoDUNE will be the starting platform.

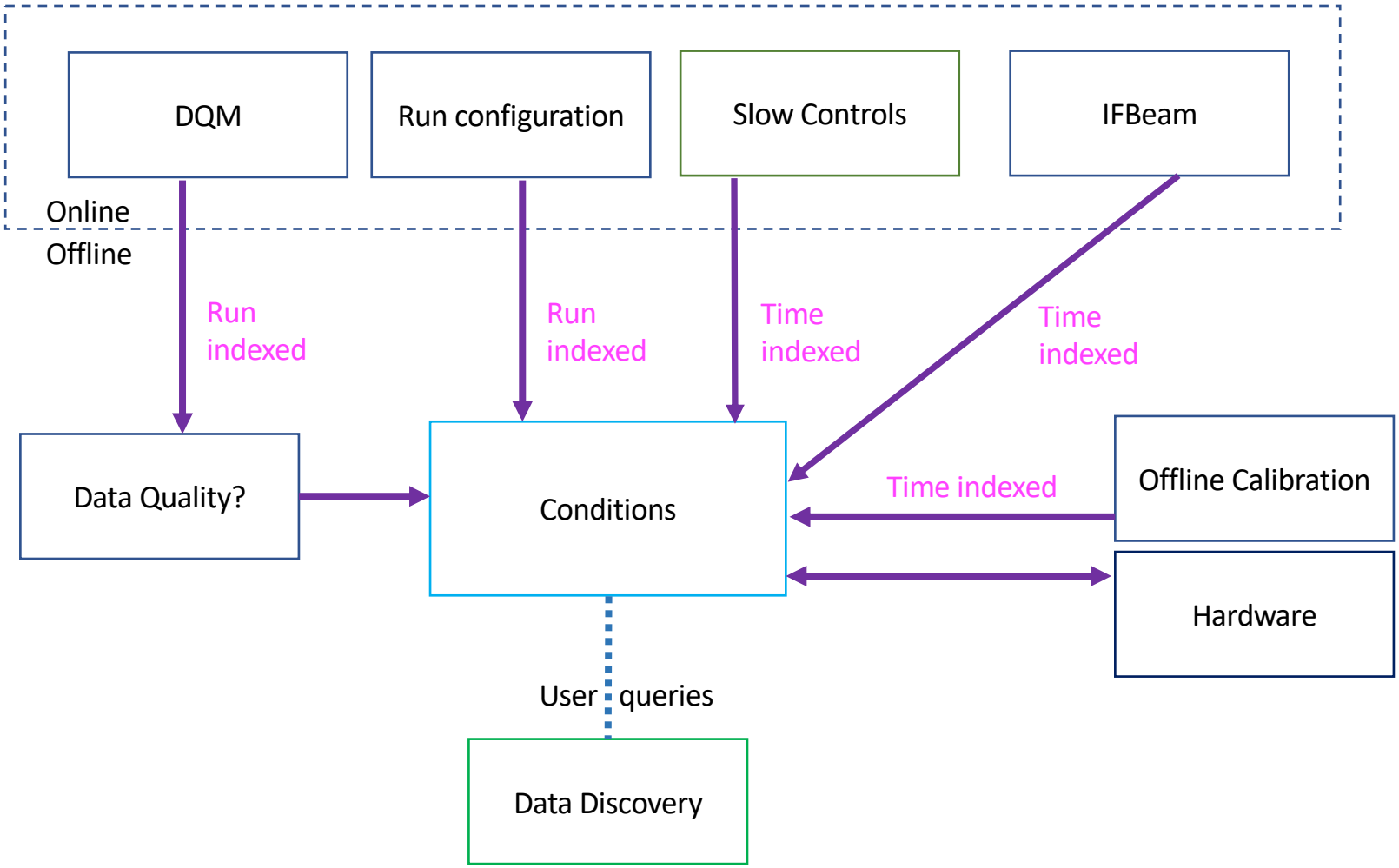
Target DB System for ProtoDUNE



Target DB System for ProtoDUNE



Target DB System for ProtoDUNE



Challenges

- Complete the various unfinished databases
  - Calibration constant database hasn't been considered yet
  - Need input from offline users to ensure needs will be met
  - David suggested that we use Run I data to help guide us for Run II – especially regarding SC filtering
- Aggregating Metadata into Conditions DB (uconDB)
  - I had assumed that the conditions DB would be indexed by run (subrun) but maybe this won't be the case. We need to make a decision on this.
  - APIs are needed to get the metadata from various DBs and insert it into the uconDB
  - Again we need input as to what gets collected for the Conditions DB and ultimately into the data discovery DB

Challenges continued

- Additional challenges
  - We would like to test as much of the long-term DB infrastructure as possible so interfacing with MetaCat would be good
  - We haven't discussed a geometry database and if there is a desire to have one or perhaps this will be handled in another manner
  - What interface is needed between the HWDB and the Conditions DB?
  - I have almost certainly missed things as well.

Aggressive timeline

We would like to have the databases for the run in place for detector and DAQ commissioning in early 2022.

Where do things stand right now?

- IFBeam is good to go.
- DAQ configuration DB is partially in place but there will be some amount of work needed to update the existing API and add in ”run-generated” information.
- Slow Controls DB is in early stages of development but we have support for some of the effort – eg. filtering.
- Offline calibration database (derived calibration constants) – nothing has been done for this – at least by the DB group.
- Run history (or data discovery) DB is in place but API between Conditions DB and Run History needs effort.
- We heard yesterday about the parts ID DB, which is needed for the HWDB.

The good news

We have new DOE funds focused on the Computing Consortium and the Databases were listed as the highest priority.

Both BNL and CSU are currently each searching for a postdoc to assist with addressing these challenges. Given that we can build on work done for Run I we should be in much better shape for Run II.

Discussion