DUNE Hardware DB: the requirements and an outlook

Maxim Potekhin
Brookhaven National Laboratory

DUNE DB Workshop@Colorado State U

12/03/2019

What's the Hardware DB?

- The term has been used in DUNE in various contexts
- For the purposes of this discussion it's a general term to refer to database(s) used in the manufacturing, QC/testing, asset management, logistics, installation and other aspects of the product life cycle

History and context (1)

- The FNAL Hardware DB was used in the construction of 2 subsystems in protoDUNE-SP (TPC, PD)
- BNL: a system for CE QC/testing (protoDUNE-SP)
- Proposals to review functional requirements for the Hardware DB
  - Marco's talk at the Jan'19 Collaboration meeting: https://indico.fnal.gov/event/16764/session/8/contribution/40/material/slides/0.pdf
  - Hajime's proposal and recent work (mobile, connected tech etc) https://indico.fnal.gov/event/21445/session/11/contribution/159/material/slides/0.pdf
- Input from Jim Stewart (see backup slides)
- Input from Brett

History and context (2)

- DocDB 16619: “Functional Requirements for the Hardware Database System in DUNE”
  - based on the considerations presented above and a study of the FNAL Hardware DB
- A survey recently launched in November 2019 with the purpose of information gathering and improved understanding of the requirements for interim and long-term solutions
  - ...will discuss some of the results of the survey
- Can the exsting situation be considered scalable/optimal for
  - protoDUNE-2
  - DUNE?

General observations (1)

- Each group/consortium “rolls its own” based on the principal needs of the activity they conduct and support
  - different focus in each group
  - cf. QC oriented work vs general asset management/logistics
- There isn't a single type of tools/products currently in use
  - commercial DB systems
  - in-house QC systems
  - ad-hoc records in an existing site registry (e.g. some of the protoDUNE DAQ hardware)
  - DB/spreadsheet combinations

General observations (2)

- The QC data generated during the production and testing is very diverse and may be extensive for some systems
  - can include the initial calibration-type data
  - overlap with configuration management and slow controls
- Does the Hardware DB need to contain survey (i.e. position) information?
- This is an important fact and a point of concern for the TPC electronics and Calibration consortiua
  - need a clean definition of the boundary between the Hardware DB and the Conditions and Calibrations DB
  - ...which overlaps/interacts with the Configuration DB
  - well designed interfaces between the above will be needed

The survey (November 2019) (1)

- HVS
  - Field Cage components for both SP and DP TPCs
  - Data component: mechanical properties and electrical continuity, resistance measurements of the HVS field cage components
  - QC done with iPads and bar code (or QR) scans - feed tables in the database (FNAL)
- SP-PD
  - QC travelers and were included in the database system for ProtoDUNE 1 (FNAL DB)
  - Data component: photosensor calibration test data and repeated detector module scans will be generated and stored
  - database will be used to generate a channel map during installation, allowing tracking from QC data to particular modules in the detector
  - “The primary lesson learned was that it was more difficult to maintain this level of record keeping during installation/integration”
  - “As parts are received underground bar codes should be scanned and the locations in the database updated. As they are installed into the detector (into the APAs and onto the CPAs in the case of the PD system) the specific location should be automatically updated to assist in generating a detector map. Cable/electronics channel maps might be automatically generated during installation. I think this should be a DUNE-wide implementation, with the same equipment used by all”

The survey (2)

- DP-PD (MySQL+spreadsheets)
  - Data component:
    - PMT performance data measured at production/assembly facilities and the Coating, Testing and Storage Facility
    - Results of continuity tests and resistivity/capacitance measurements of the HV/signal cables, results of transmission measurements of the optical fibers and connecting assemblies
    - The electronics test results of the high voltage distribution boxes and the high voltage modules and crates at the procurement/production facilities
    - Visual checks of the reflector/WLS panel assemblies and the installation hardware.
    - Complete mapping information starting from the PMT serial number and including the HV/signal cable ID, calibration fiber ID, feedthrough flange ID, feedthrough connector ID, warm cable ID, splitter box ID, signal cable ID, HV cable ID, HV module/channel ID, front-end module/channel ID, crate IDs, calibration box ID and light source ID
  - “All the components and the coupling between the components should be tracked precisely from the very beginning. This would allow tracing back any potential problem to its source and mitigation based on previous QC data”
  - “What is the wider scope of the DUNE Database Group? Is storing PMT calibration data continuously also within the scope? Is there going to be a common hardware database for all the consortia or separate databases for each consortia i.e. will the physical interfaces also be reflected in the database structure?”

The survey (3)

- DAQ
  - “Our understanding is that the Hardware DB aids tracking the HW life-cycle, and in case of HW failures it can help with the identification of reasons and possible batch problems”

The survey (4)

- TPC electronics
  - Interfaces: APA, HV, PDS, possibly SC
  - Data component: “The correct answer is a lot”
    - Calibration constants for all the ASICs (betweeen 6,000 and 54,000 per far detector module, depending on the choice), for all the FEMBs (3,000)
    - The exact amount of calibration data for each ASIC/FEMB has not been fully determined
    - At the moment we estimate something of the order of 25 Mbits is required to fully configure one far detector module during operation
    - During calibration the amount of data stored will be significantly larger
    - Additional information includes the configuration of the bias voltage supplies, and of power supplies
  - “The information in the calibration database will be a subset of the information stored in the hardware database”
  - “We will continue to store information in the hardware database throughout the lifetime of the experiment. We will store part of this information also in calibration database(s) that are required by the online to configure the detector during data taking and by the offline for data reconstruction. We do not think that it is appropriate to consider the hardware database separately from the calibration database”

The survey (5)

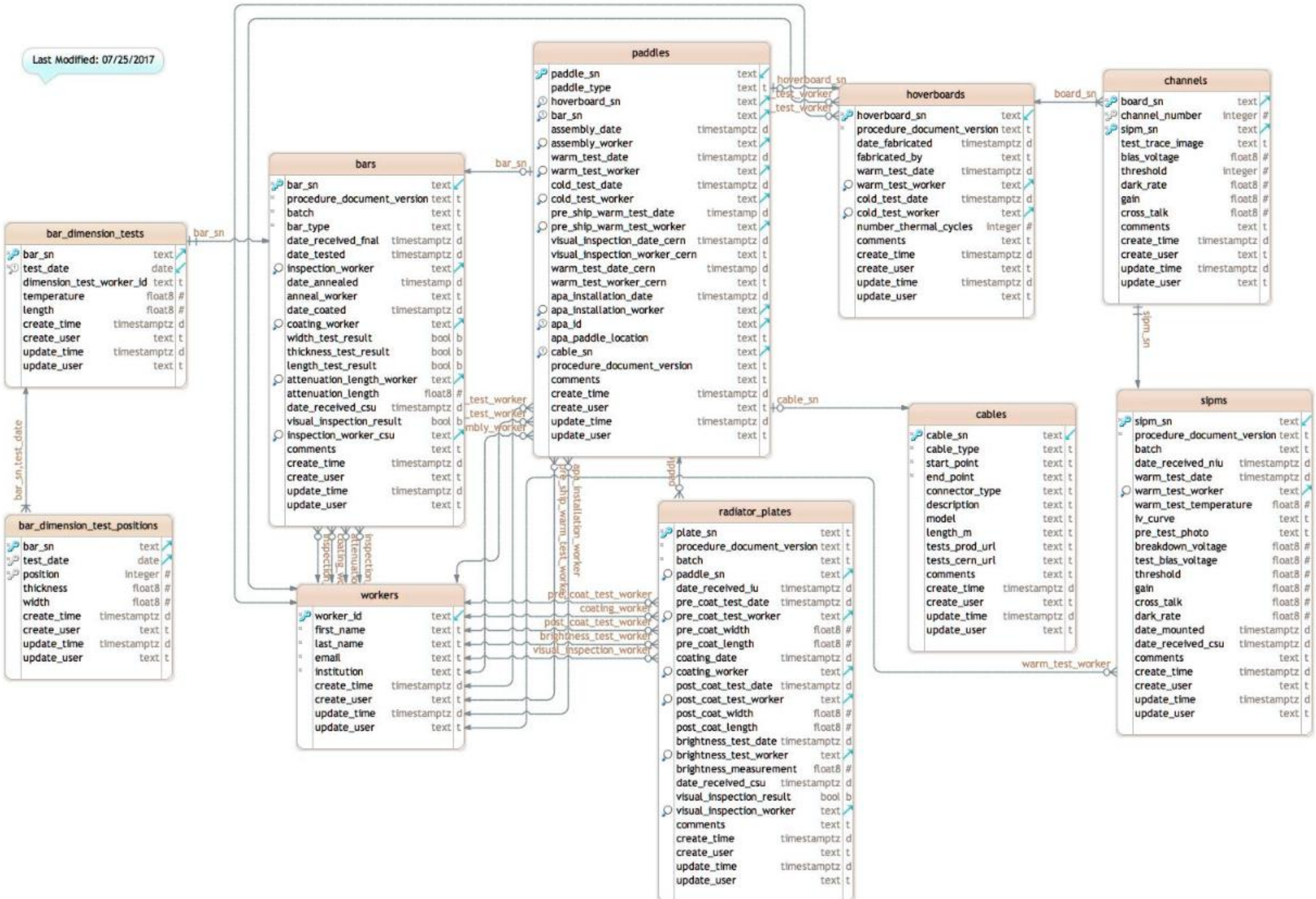
- Calibration (lots of info provided in the survey, thanks!)
  - At present, records are being kept mostly in shared documents and spreadsheets
  - "For QC, we will collect data at different stages: post-fabrication tests; tests at various test stands including protoDUNE-2; tests to be done after shipping and installation at SURF and underground; there will be system tests but also tests related to electronics and racks that we will be recording. "
  - "At present the hardware database records are intended to be used to obtain accurate installation information about the different components of the calibration systems, such as positions, and identification codes to different analysis algorithms. In case the hardware database permits versioning, it would be the intention of the calibration consortium to use that feature to maintain up to date records, of modifications to the hardware installation"

A few use cases/examples

FNAL Hardware DB example - PDS paddles - schemas

proto_dune.jpg

Stephen White, 08/23/2017 01:18 PM



BROOKHAVEN NATIONAL LABORATORY

DUNE

FNAL Hardware DB example - PDS paddles - Web UI

protoDune Hardware Database

DEEP UNDERGROUND NEUTRINO EXPERIMENT

Home Permission Mgr Admin Mgr Log out

Table paddles

dump to CSV plot... search data config data <prev next>

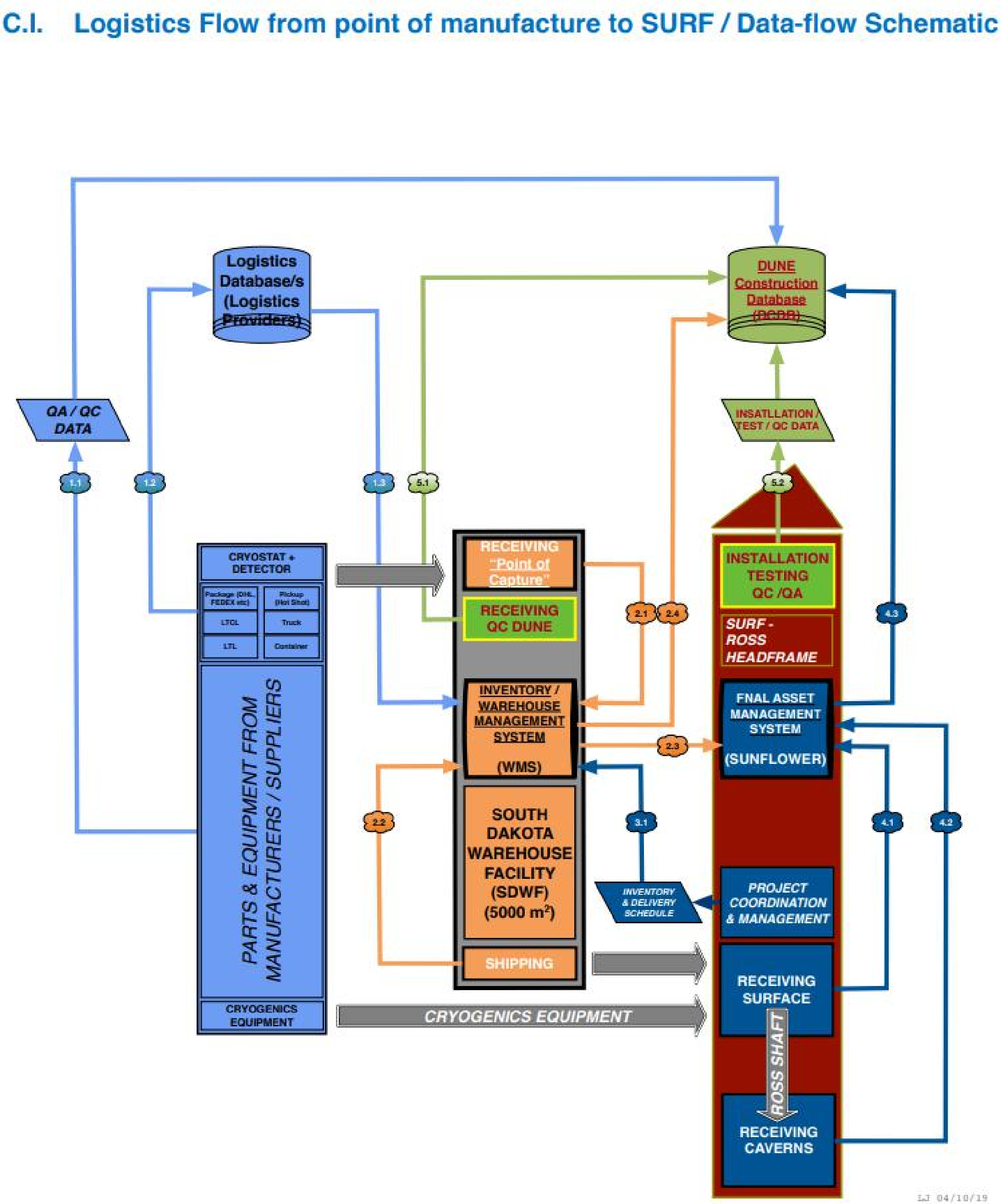
Sorting by: paddle_sn desc

	Paddle Sn	Paddle Type	Hoverboard Sn	Bar Sn	Assembly Date	Assembly Worker	Warm Test Date	Warm Test Worker	Cold Test Date	Cold Test Worker	Pre Ship Warm Test Date	Pre Ship Warm Test Worker	Visual Inspection Date Cern	Visual Inspection Worker Cern	Warm Test Date Cern	Warm Test Worker Cern	Apa Installation Date	Apa Installation Worker	Apa Id
view/edit	IU9	IU	HAM-002-0020	IU9	03/16/2018 09:00:00-0500	ap_csu01	03/20/2018 13:23:00-0500	ac_csu01											
view/edit	IU55	IU	HAM-002-0056	IU55	02/14/2018 14:30:00-0600	zr_csu01	03/12/2018 09:07:00-0500	ap_csu01											
view/edit	IU54	IU	HAM-002-0008	IU54	02/28/2018 13:50:00-0600	zr_csu01	03/09/2018 09:54:00-0600	ac_csu01											
view/edit	IU53	IU	403-002-0023	IU53	12/19/2017 13:45:00-0600	zr_csu01	12/20/2017 17:02:00-0600	ap_csu01											
view/edit	IU52	IU	403-002-0042	IU52	11/21/2017 14:00:00-0600	zr_csu01	11/29/2017 13:38:00-0600	ap_csu01	12/05/2017 00:00:00-0600	ap_csu01	12/07/2017 11:21:00	ap_csu01							
view/edit	IU51	IU	403-001-0039	IU51	11/21/2017 16:00:00-0600	zr_csu01	11/29/2017 13:43:00-0600	ap_csu01	12/05/2017 00:00:00-0600	ap_csu01	12/08/2017 12:04:00	ap_csu01							
view/edit	IU50	IU	403-002-0044	IU50	12/01/2017 14:00:00-0600	zr_csu01	12/20/2017 17:07:00-0600	ap_csu01											
view/edit	IU49	IU	403-001-0006	IU49	06/01/2017 00:00:00-0500	jd_csu01	06/12/2017 00:00:00-0500	jd_csu01	07/09/2017 00:00:00-0500	jd_csu01		jd_csu01	08/18/2017 00:00:00-0500	en_fnal01	08/18/2017 00:00:00	en_fnal01	08/26/2017 00:00:00-0500	tr_umn01	1
view/edit	IU48	IU	403-003-0004	IU48	02/12/2018 15:00:00-0600	zr_csu01	03/14/2018 18:01:00-0500	ac_csu01											
view/edit	IU47	IU		IU47	02/26/2018 13:15:00-0600	zr_csu01													
view/edit	IU37	IU	403-003-0011	IU37	02/07/2018 17:30:00-0600	zr_csu01	02/08/2018 14:07:00-0600	ac_csu01											
view/edit	IU35	IU	403-002-0038	IU35		zr_csu01	02/08/2018 15:22:00-0600	ac_csu01											
view/edit	IU31	IU	403-002-0020	IU31		zr_csu01	10/27/2017 10:53:00-0500	ap_csu01	11/01/2017 00:00:00-0500	ac_csu01	11/06/2017 09:12:00	ap_csu01							
view/edit	IU28	IU	403-003-0063	IU28	06/02/2017 00:00:00-0500	jd_csu01	06/12/2017 12:50:00-0500	jd_csu01	07/08/2017 00:00:00-0500	jd_csu01	07/12/2017 00:00:00	jd_csu01	08/22/2017 00:00:00-0500	en_fnal01	08/22/2017 00:00:00	en_fnal01	08/26/2017 00:00:00-0500	tr_umn01	1
view/edit	IU27	IU	403-003-0002	IU27	02/07/2018 13:00:00-0600	zr_csu01	02/08/2018 16:40:00-0600	ac_csu01											

FNAL Hardware DB: observations

- Worked well for two subsystems in protoDUNE-SP
- Denormalized schemas
- Custom schemas for each subsystem
  - down to the explicit mention of CERN i.e. hardcoded logistcs
  - manufacturing and QC procedures are effectively hardcoded in the schemas
- Unclear how more complex QC cases can be handled
- Lack of clear interface to the (future) Conditions DB/Calibrations DB
- Aggregation of components implemented by using foreign keys
  - hierarchical structure of most DUNE component is recognized and addressed

Example:
SURF logistics



Example: “QC App”

- Credits: Hajime et al
- Mobile computing: an iPad application helping users to feed and access QC data stored in the DUNE hardware DB
- Need to recognize the importance of this use case since the sheer scale of DUNE components won't allow for all QC and other procedures to be done on a workbench
- Valuable experience that needs to be leveraged

Based on what we learned, what's missing?

- **Integration**
  - current lack of unified view, multitude of databases and solutions with combinatorially large number of potential interfaces and exchange formats
  - need to properly address component aggregation/association
- Coverage of all aspects of the product **life cycle** (some items do exist in isolation)
  - manufacturing
  - QC + other testing (possibly in multiple locations)
  - assembly
  - logistics
  - installation
  - documentation + traceability (i.e. keeping references to relevant docs, technical drawings etc)
  - testing “in situ”
  - maintenance and component replacement
  - interfaces with Conditions/Calibrations
  - cabling
  - ...other suggestions?

The risks

- Can we afford to do nothing at all? A few points to consider
  - duplication of effort
  - lost opportunity to share expertise and the development cost
  - need to properly address component aggregation/association
  - looking at the survey and other input, some requirements are not currently met
- Do we understand the operational risks for DUNE with the current set of tools and systems?
  - both in the construction phase and then into the operations
  - what worked for protoDUNE(s) will not likely scale well for DUNE
  - implications for efficacy of QA
  - in case of proprietary DB products being used, should we give considerations to
    - integrated licensing cost over the lifetime of the experiment
    - what to do if the provider goes out of business?

Today's discussion

- Let's take a step back and see whether we
  - understand operational risks if extrapolating usage of current solutions
  - can identify functional requirements for the support the DUNE hardware life cycle
- Some basic requirements have been summarized in DocDB 16619
  - by no means complete or all-inclusive but mainly to start a discussion
  - expect iteration and broader contributions
- Can we agree on some or all of the requirements?
- See what we can do in terms of cross-consortia cooperation
- What steps forward can be taken

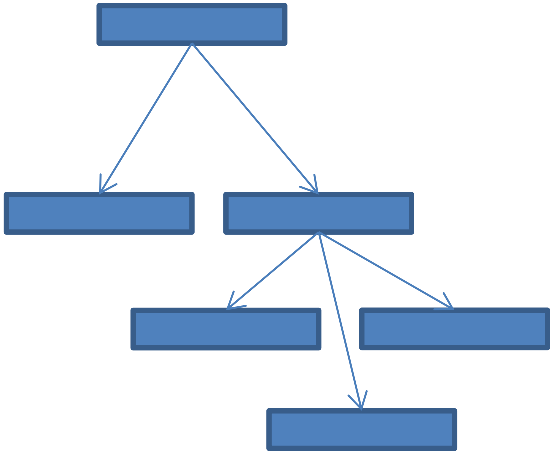
Generalizing the use cases: a model

A Generalized Model

- The model includes
  - Components and Aggregations
  - States and Transitions
- In fact, no new content: based on the existing use cases and systems
- Hopefully demonstrates commonality across subsystems

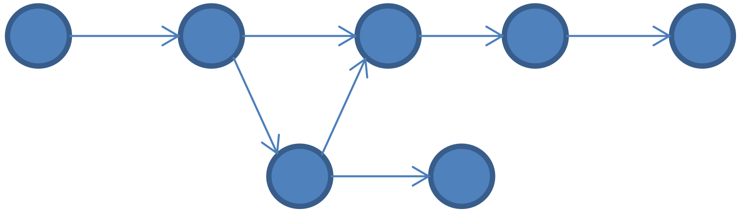
Components and Aggregations

- Consider the apparatus as an hierarchy of elements
- cf. a FEMB can be modeled as an aggregation which contains various electronics components e.g. ASICS
- Most components are aggregations themselves, although in graph terms, the leaves of the tree are atomic by definition
- In the following we assume that a component may be either atomic or an aggregation of other components subject to the same assumption



States and Transitions

- Each component can exist in a variety of states
- The chain of events starting from design/ordering/manufacture (or procurement) and then on to QC, shipping, installation, testing “in situ” etc is reflected by the changing state of the component as recorded in the database
- The sequence of transitions depends on the type of the specific component
- Each state transition may optionally result in a data product specific to the component and the specific transition
  - *can be recorded as an external reference (cf. interface with ConditioDB)*
- Each state transition may optionally require a record in the audit trail e.g.the time stamp, the reason for the event + any comments required in the process + the entity responsible for changing the state and updating the record



Evolution of Aggregate Components

- Experience shows that the hardware may need to be swapped, repaired, updated etc
- Consider the use case where a particular component within an aggregation needs to be replaced, such as a failed ASIC installed on a board. The existing database ID can be preserved and the version number updated, or the the object can be retired and recreated under new ID
- The choice will depend on the situation or resolved later in the design stage
- This approach allows to handle an important use case whereby a component is redesigned in terms of its construction and composition while retaining its functionality and place in the overall hierarchy.

Component Location and Ownership

- We will use the concept of the *functional position* introduced in ATLAS, which refers to a physical position of a component specific to its function and which cannot be moved
- It is complementary to the component's *geographic location*
- Ownership refers to the entity who is assigned custody of the item and is authorized to perform specific actions relate to the item
- The geographical location, functional position and ownership of an item are an integral part of its state

Functional Requirements (abridged, see DocDB 16619)

- The Unique ID Requirement (not identical to the manufacturer's and other ad hoc IDs)
- Traceability - identify component in technical documentation/drawings etc
- Completeness - support the full life cycle
- Record of state transitions
  - History
  - Periods of vailidity
  - Evolving documentation and metrics after deployment
- Record of Connectivity
  - Physical and electrical connections
  - History
- Interfaces
  - Network and Web
  - Component life cycle management e.g. links to the Conditions/Slow Controls etc
  - Role-based Auth/Auth
  - Industrial ID Systems and Mobile Platforms

The Timeline

- Near and Medium Term
  - Preparations for the protoDUNE-2 experiment, with manufacturing of components to commence in 2020, and installation scheduled for 2021
  - Early manufacturing of certain components of the DUNE experiment
- Longer term, including manufacturing and installation of the DUNE detector, commissioning and operation
- We can't create an optimal solution complete with life cycle management right away
- Possible measures to "bridge the gap":
  - Survey and catalog the interim systems (e.g. existing and in use at participating sites) which are likely to be used in near and medium term
  - Identify additional attributes (if any) to be added to object models in these systems and modest changes in functionality to allow for eventual migration to the final system
  - Define formats for data export and ingestion necessary for such migration as well as future I/O and data exchange protocols
- Start a discussion: can we envisage a common solution for DUNE?
  - Substantial effort
  - Substantial benefits

Backup slides

Input from Jim Stewart (1)

- All test information can be accessed based on PBS ID (like a model number) and a unique ID (like a serial number).
- An operator must be able to find the test data easily. This means efficient mapping between physics, installed hardware, and parts.
- Part history needs tracked. If a part is repaired then the before and after test data needs to accessible. If a sub-component is repaired or replaced the related information needs archived (diagnostic test, results leading to the repair).
- Location of the parts needs captured during fabrication, shipping, storage and either deployment or long-term storage.
- Who performed work, when, using what equipment must be captured for each test.
- The data must be hierarchical - which chips are on which boards that are connected to which APA needs captured.
- Need to fix the numbering and coordinates in a well documented and clear way.
- Need to be able to easily assign parts to detector locations at the time of installation. The installation team need to be able to assign APA 67 to row 1 north easily at the time of installation.

Input from Jim Stewart (2)

- All assignments need to be easily modified in case something needs to change.
- Need to be able to easily transfer information between stand alone tests and tests with final daq.
- All data must be archived and backed up. A plan for maintaining the data even if institutions leave DUNE is needed.
- The QC plan and procedures related to each part need to be accessible.
- Given the PBS and UID all related drawings, engineering analysis notes, procedures, safety documents, and interface documents must be accessible.
- The PBS must go on the drawings as part of the BOM.
- The SDS for each part need to be available from the database.
- For COT equipment the manuals should be easily retrieve form the PBS
- If equipment needs to be configured then the configuration description and files if needed are to be linked to the PBS. The UID needs to record the configuration of any configurable parts.
- All cables need a link to the documents defining the cables (conductors, pins, pin definition, etc.)

Input from Jim Stewart (3)

- All circuit boards need links to the fabrication package (Gerber files, QC tests, layout).
- For electronics the location needs to include the position in racks and so on. This requires a rack numbering scheme.
- A method to connect survey information to parts needs to be available.
- Any part inside the cryostat needs to be connected to a grounding document that is approved.
- A map or index of all data related to a part is needed in an easily understandable format.

Input from Brett Viren (an excerpt)

- A CE FEMB develops excessive noise at some point. DAQ runs with it for a period and then re-configures to ignroe it in the self-trigger system. Later, the source of noise is understood and removed and the FEMB is allowed to once again contribute to self-trigger. This class of problem occurs throughout the FD in different FEMBs over the life of the experiment.
- How does an analyzer know the self-trigger configuration over the entire period of the analysis (say many years) in order to calculate total trigger exposure?
- How does an analyzer know to treat as special any candidate reconstructed vertices that have contribution from the problematic FEMB channels?
- How does the exact times of detection and resolution (which precede the subrun changes) get stored and communicated to the offline?

Backup slides - existing solutions

FNAL Hardware DB

- Pros:
  - Existing, proven system, experience in DUNE
  - Supports many of the existing requirements as stated above
- Cons:
  - A rigid schema not well suited for evolution
  - Doesn't meet all requirements, does not support the component life cycle
- Can be the basis for future development and added features

ATLAS Equipment Database

- Pros:
  - A full featured system covering many complex use cases
  - Likely supports most and perhaps all of the DUNE requirements
  - Proven during construction, installation and operation of a very complex detector
  - Integration of a number of separate databases (using the Glance system)
- Cons:
  - Substantially based on a proprietory commercial system, entailing the licensing and life cycle issues
  - Has many hooks into the vast CERN information infrastructure which would be hard to disentangle in the reuse scenario
  - Considerable amount of work was required to tool the database for specific ATLAS needs, this part won't be immediately reusable in any case
- Need to take a further look

Sunflower asset management system

- Need in-depth review of how it fits with the requirements