

ND280 Upgrade: Software

Offline, Online, Analysis and Computing

Yoshi Uchida

Neutrino Near Detectors Workshop, CERN

21 March 2017

- Overview and Status of ND280 Software
 - Offline, Online, Analysis and Computing
 - Software and Data Management
- Software Outlook (with and without ND280 Upgrade)
- Specific Opportunities

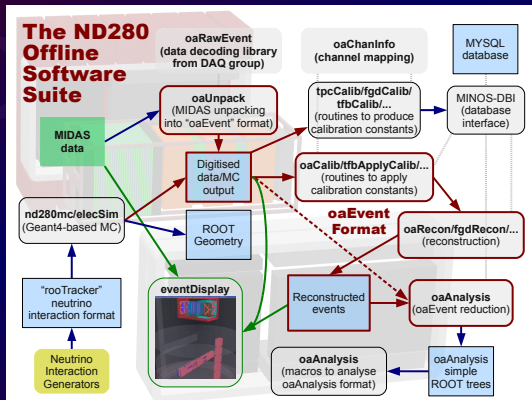


The ND280 Offline Software Suite



ND280 Software Offline

- Project commenced 2004
- Most framework design choices made by end of 2005
- Based on ROOT, and Geant4 and other open source packages
- Interfacing with NEUT, GENIE, NUANCE etc.
- Custom oaEvent format
- Main framework extremely stable by start of data-taking in 2009/10



The ND280 Offline Software Suite



ND280 Online Software

Interface with Detectors

- MIDAS DAQ writes raw .mid.gz binary files
- DAQ group provides “header” files
 - used to encode DAQ into MIDAS banks
 - same headers used by Offline Software for decoding
- Software representation of geometry formed within Geant4
- Exported to ROOT TGeometry format
- Official geometries saved on server
- Used in real-data decoding
- Offline Software can read raw data files directly
- Raw-data interface has been extremely reliable
- Very tightly-knit relationship between raw data, offline geometry and offline “digit” formats

ND280 Software Management

T2K-wide principle: **All tools and inputs that lead to a published result must be accessible for any collaboration member**

- Official CVS Repository for all code, with **ViewVC** and **LXR** interfaces
- **CMT**-based versioning and software deployment
- **TUT**-based unit testing of all low-level code
- **Bugzilla** for bug-reporting and tracking
- All software **provided as source code for compilation**; if they pass unit tests and validations, any hardware platform can be used (no binary distributions)
- Constant validation of high-level code (custom routines)
- ROOT and Geant4 distributed as part of software (currently 5.34.34 and 4.9.4)

More recently among some working groups:

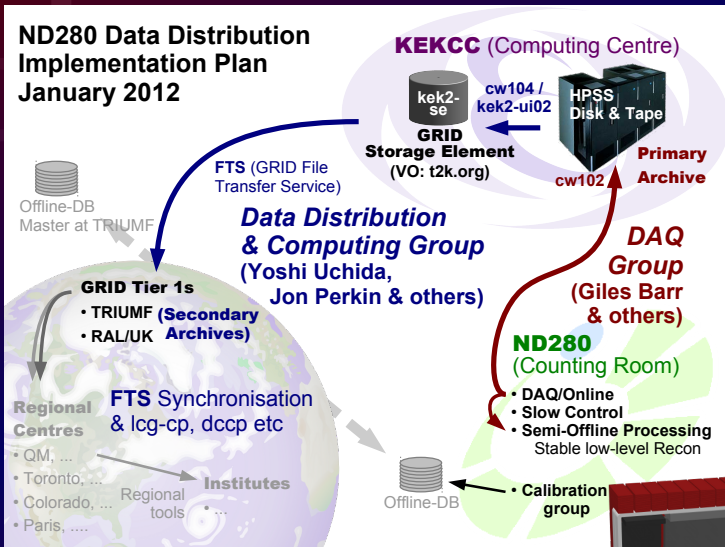
- Use of **Git** (frozen versions still must be committed to CVS)
- Unofficial use of **Flowdock** and **Slack** etc. for communication

No agreed collaboration-wide workflow which incorporates newer tools yet

ND280 Computing and Data Distribution

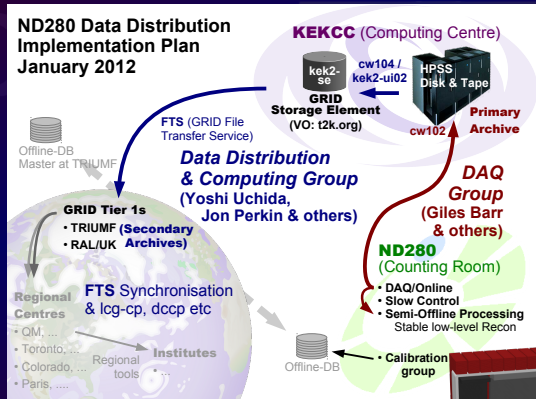
Overview

ND280 Data Distribution Implementation Plan January 2012



ND280 Computing and Data Distribution Overview

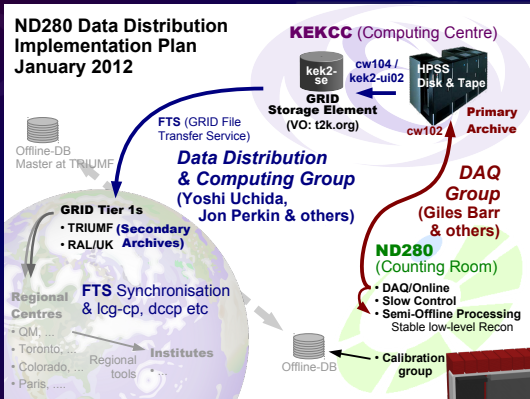
- Development since ca. 2008/09
- Partially Grid-based; but too early to borrow full framework directly from LHC experiments
- Using custom code built on low-level tools (e.g.: LCG, LSF) and FTS
- Tier 1s in Canada and UK
- MC Production in Canada, UK and US, France, Poland, ...



ND280 Online Software

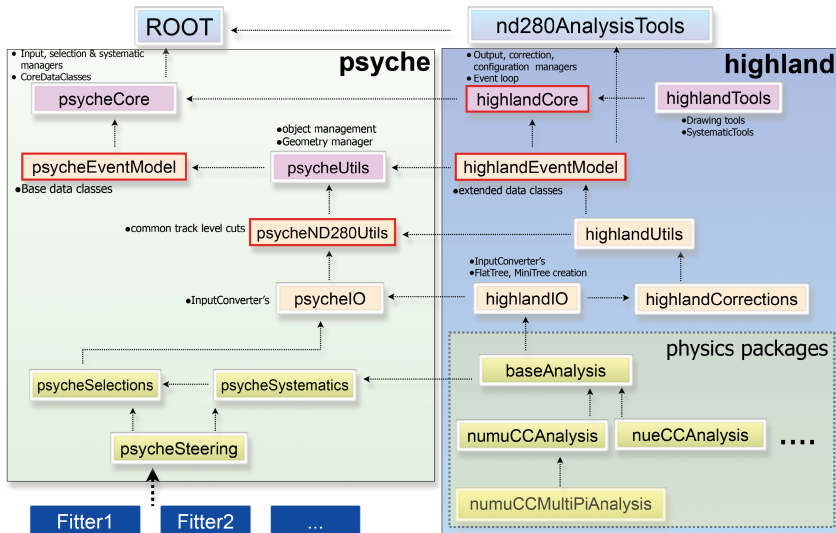
Interface with Database

- MySQL servers on-site at J-PARC and mirrored at TRIUMF
- Custom-designed database writing and look-up interface
- Proven to be stable and reliable



ND280 Offline Software

High-Level Analysis

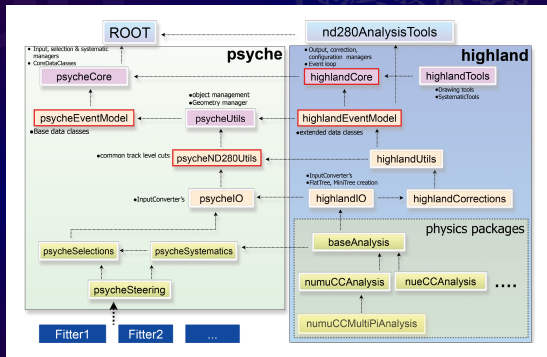


ND280 Offline Software

High-Level Analysis

User-level final analysis tools lies downstream of overall software framework fixed prior to data-taking

- Several ideas arose from physics working groups
- Current main analysis framework in development since ca. 2012
 - High-Level Analysis at the Near Detector
 - Propagation of SYstematics and CHaracterization of Events



Anselmo Cervera

Development, innovation and improvement ongoing

ND280 Software Outlook

Continued running of existing detector at high rates

(Strictly YU's personal thoughts)

- **MIDAS-raw data encoding/decoding-geometry framework** is very solid and does not need touching
- **Calibration** work for detectors is mature; labour-intensive but coping
- **Detector alignment** is done "by hand"; can cope if no significant changes
- **Computing (data and MC production)** has been keeping up, but larger data volumes and sophisticated analyses are now pushing the limits
- Maturing simulations and analyses may allow for **reduction of information** that is persisted
- **High-level analysis framework** developing nicely, perhaps short in manpower
- **Software management/communication etc.** functioning well, but are all a bit "noughties" (CVS and CMT and mailing lists)
- Keeping up with external package evolution e.g., **ROOT 6**, is proving difficult

ND280 Software Outlook

With new detectors

(Strictly YU's personal thoughts)

- **New online and offline code** needed (geometries, interface with DAQ, detector and electronics response simulation, implementation of calibration procedures)
 - especially if DAQ systems are different to existing MIDAS set-up; this would introduce a significant amount of work
 - low-level data storage classes (raw, digits and recon) may need revamping
- **Additional computing and manpower resources** needed for new-detector commissioning, calibration, reconstruction
- May need more active work on **detector alignment**
- **DAQ data rates** (from detectors to KEK primary archive) are designed for existing detectors; need to revise

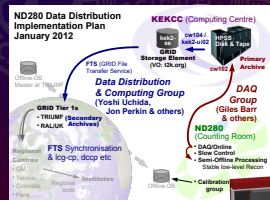
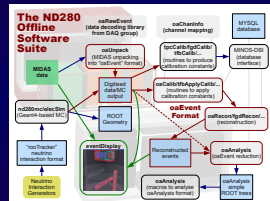
Opportunities

Expertise in and/or contributions to the following would be valuable (or critical) for an upgraded ND280:

- Incorporation of new detectors into online and offline software (obviously)
 - must happen at many levels—not a trivial set of tasks
- Update how detector alignment is handled
- Transitioning to more modern methods of software management (Git and GitLab, CMake etc.)
- Closer coordination with external software (ROOT, Geant4 etc.)
- Streamlining/automating data and MC production tasks (expertise from large experiments?)
- Optimising CPU and storage use by the software
- Increasing CPU and storage, file transfer rates
- High-level analysis frameworks
- ...

Conclusions

- ND280 Software Framework is working well
- Low-level code is very stable and reliable
- But now approaching the end of its expected lifetime
- Some modernising and re-organising of management methods and high-level code could be useful
- A new ND280 detector would require additional work
 - geometries/channel mapping etc.
 - interface with DAQ
- Computing set-up functions well, but age is showing
 - manpower-intensive
 - reaching resource limitations
- Opportunities for significant and valuable contributions across all these areas



Thanks To

Current experts:

- Alex Finch (Lancaster)
- Anselmo Cervera (Valencia)
- Clark McGrew (Stony Brook)
- Jeremy Lopez (Colorado)
- Phill Litchfield (Imperial)
- Thomas Lindner (TRIUMF)
- Yoshi Uchida (Imperial)

and the many T2K collaborators who have contributed to ND280 Software and Computing (but in particular the past members Nick West and Jon Perkin)

