Hyper-K and T2K
- Software and computing -

Sophie King

Kamiokande Series

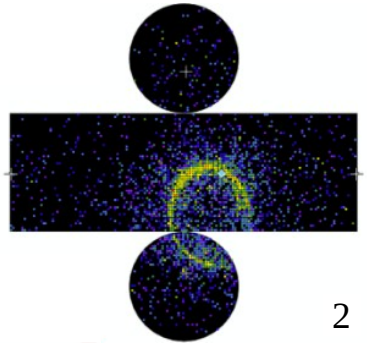
Hyper-Kamiokande

Tank of ultra pure water, Lined with PhotoMultiplier Tubes (PMTs)
→ Detect cherenkov radiation

* Natural neutrino sources:
  Atmospheric, solar, supernova, gamma-ray bursts
  → astrophysics, neutrino oscillations

* Accelerator beam neutrinos
  → neutrino oscilltions, **CP-Violation**

* Proton decay, new physics searches



T2K: Tokai-to-Kamioka

Long-baseline neutrino oscillation experiment
Super-Kaniokande as far detector
Comlex of near detectors, ND280

Hyper-Kamiokande

258 kton

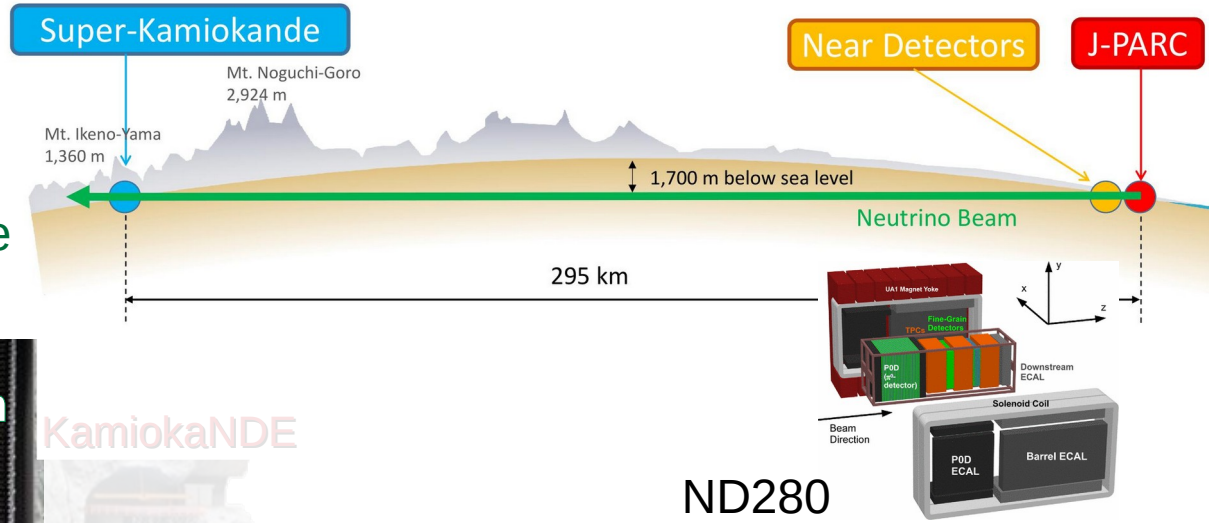
UNDER CONSTRUCTION!

Super-Kamiokande

KamiokaNDE

3 kton

Neutrino beam from JPARC



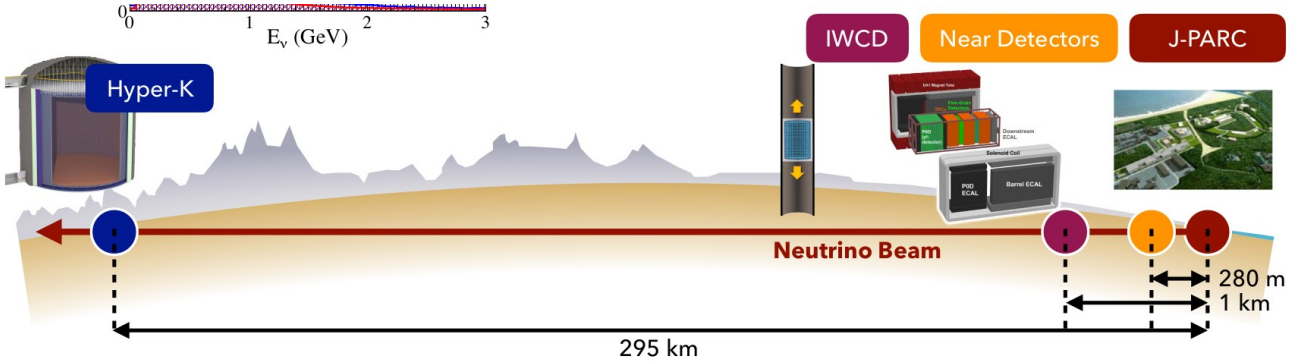
ND280
near detector to constrain systematics at SK

Hyper-Kamiokande (2027 - 2047)

Long-baseline neutrino oscillation experiment
Hyperk-Kaniokande as far detector
Comlex of near detectors, ND280, IWCD

IWCD:
Intermediate water
cherenkov detector

Hyper-Kamiokande



50 kton

3 kton → **T2K soft/comp transitions into HK**

T2K and HK share detectors and have same far detector technology

ND280 Computing status

Raw data, data processing, MC production, calibration

GridPP Instance of DIRAC
- → This serves all of ND280
- → Great user support from Imperial, Gridpp-Dirac mailing list, GGUS ticket system

- → DIRAC for job submission
  - → LHC Grid CEs in UK and France
- → Singularity container sandboxes distributed on CVMFS
  - → CVMFS stratum-0 hosted by RAL

- → Storage at Grid sites in UK, France, Canada, Japan
  - → DIRAC File Catolgue

ND280 Software

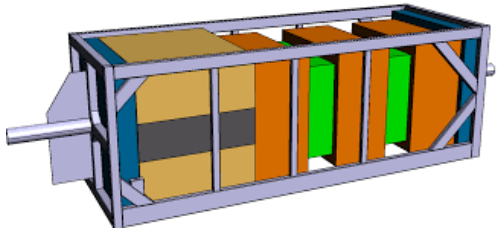
ND280 software framework (package structure, file formats) well established for many years.

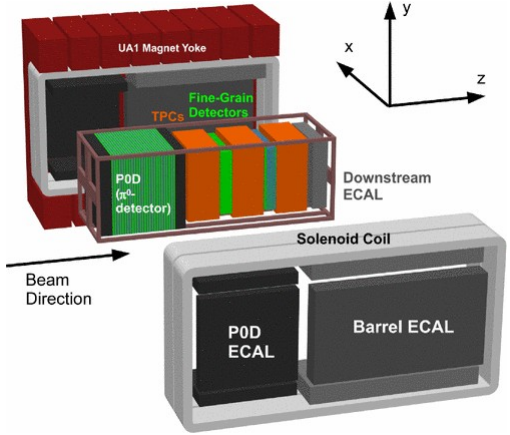
Recent overhaul

* cvs, cmt → **gitLab**, **cmake**
* Continuous integration using **gitLab** and **Docker**

ND280upgrade next year

→ **new sub-detectors**





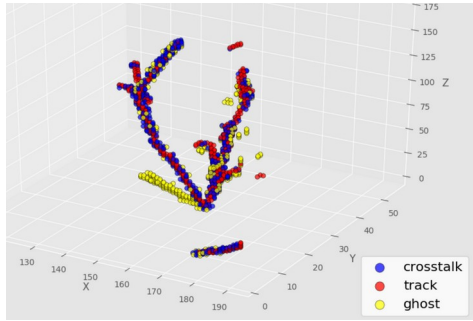
→ new reconstruction development

→ **Machine learning** techniques being explored

→ CNN, GNN

→ A lot of this being carried out at **CERN**

→ T2K members **collaborating with computing scientists**



HK Computing

Hyper-Kamiokande

Grid set up similar to T2K/ND280 (these resources will transfer to HK eventually)

*** GridPP DIRAC**

* **UK, France.** Expanding to also include: **Italy, Poland, Canada, Japan**

* Currently run single core CPU jobs for MC production

* Testing to multi-thread some of the software packages

* Reconstruction is the bottleneck in MC production → needs development

Large resources in Japan (KEK CC, Kamioka)

Non-Grid storage: **irods** (Japan), **nextCloud** (UK)

HK Computing

JENIFFER 2 Project (MSCA-RISE, EU funding)

→ Collaboration between **T2K/HK and Belle II**

* Computing work package looking at **cloud computing with DIRAC and VCYCLE**

→ Proof of concept tests: INFN ReCaS Napoli Cloud (Italy), LPNHE (France)

→ Exploring ways to set up joint cloud resources between HK and Belle II

GPU

T2K analysis level: Oscillation analysis group, use **GPU for MCMC fitting** software

→ Compute canada, UK (RAL/Oxford and Imperial batch systems)

→ also require access to machines for code development

→ Need more GPU time as we collect more data.

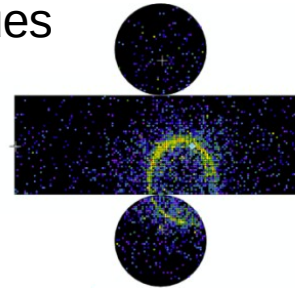
HK machine learning: Led by Canadian group and uses Compute Canada resources

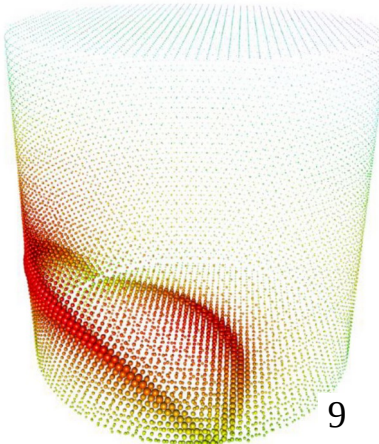
- current focus on **PID**, starting to look at **reconstruction, simulation, calibration** methods.
  - point cloud, CNN, GNN
- current work focuses on the intermediate detector (smaller version of HK)
- applying similar techniques to HK can run into computing issues

→ larger data set (fast access storage for training)

→ increased requirement for GPU memory

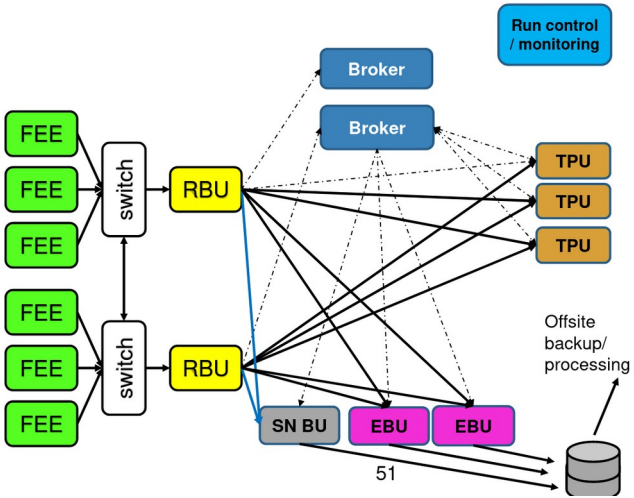
→ need to maximize efficient use of resources





DAQ

The Hyper-K DAQ system uses the ToolDAQ framework
- developed in the UK
- modular, scalable, fault tolerant



raw hit rate: 5GB/s (mostly dark noise)

→ reduce this using triggers

→ real-time processing of hits needs to be fast

→ GPUs are used to increase the speed

→ Machine learning techniques also being explored

The most challenging part of the DAQ system is the need to be ready for a Supernova
- stress of data from a near supernova
- temporarily store data long enough that we cover the time period for a supernova that is identified by an external experiment

Summary of role of UK

Internally

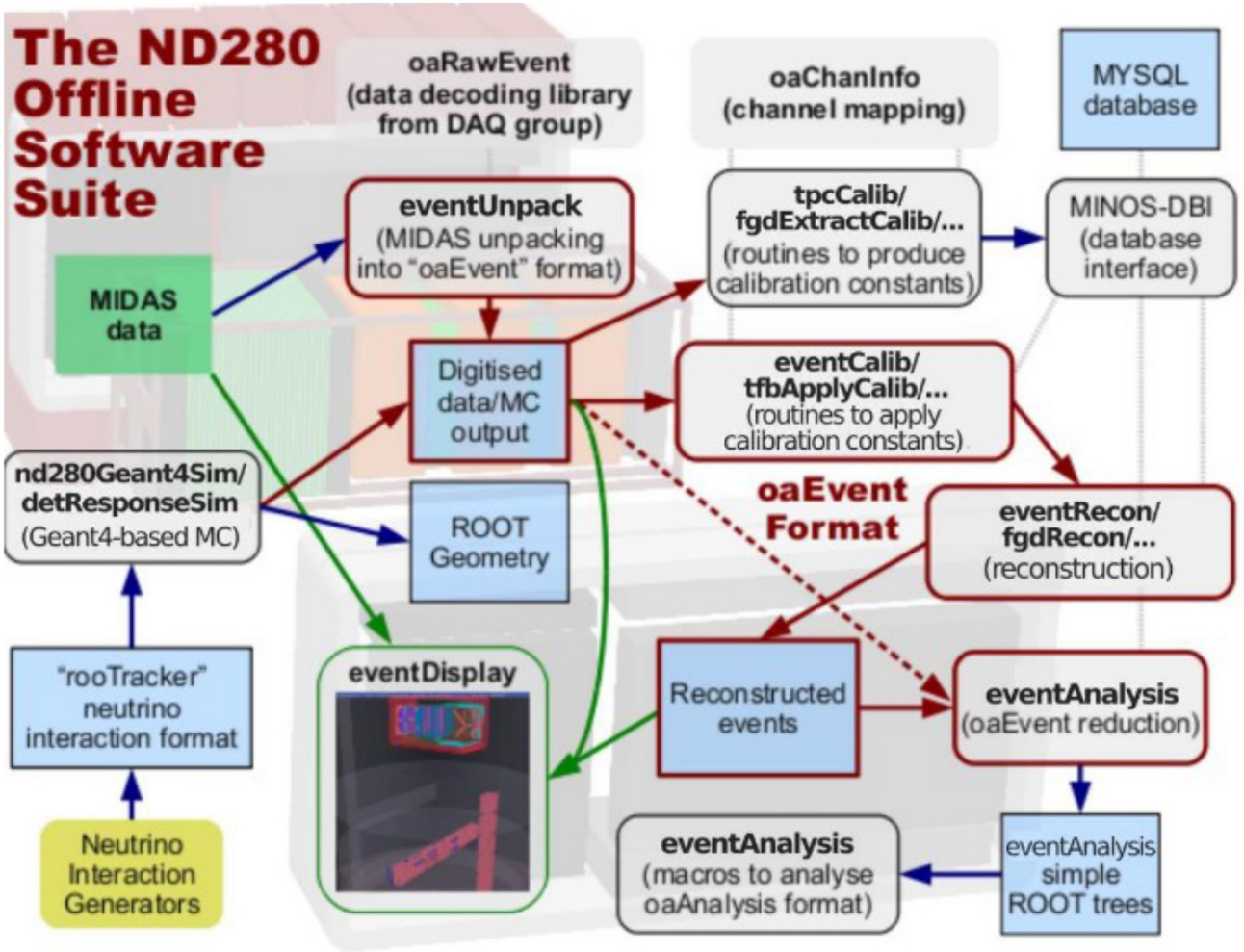
- UK group lead Computing in ND280 and HK
- UK group lead DAQ in ND280 and HK
- UK group is currently leading software consultation for HK experiment
  - to defne what we need from a software framework
    - ensure we efficently meet needs of comp/soft/calib/daq/physics and multiple detectors
    - building on highly successful ND280 software consultation done many years ago

Externally

- resources we receive from GridPP make UK one of the biggest contributors to HK and ND280
- GridPP Dirac service underpins our model for raw data and MC storage and processing
  - we receive great user support

BACKUP

ND280 Software



ND280 Software

Hyper-Kamiokande

Software packages within nd280 software suite are split into 6 ‘meta-packages’

Analysis

analysisMaster	1.82
tfbExtractCalib	1.44
mppcExtractCalib	1.26
ecalApplyCalib	2.24
soffTasks	1.57
alignTools	0.15.1
fgdHighLevelCalib	0.3.1
eventDisplay	3.89
eventAnalysis	6.12
ecalTestBeamAnalysis	0.9.1
selectControlSample	1.9.1
analysisTools	1.20

Base

baseMaster	1.70
testBase	1.17.1
oaEvent	9.2
oaUtility	5.14
oaRawEvent	3.31.1
oaRuntimeParams	0.10
oaSlowControlDatabase	1.15.1
oaOfflineDatabase	2.16
nd280SoftwareControl	2.2
oaCalibTables	1.27.1
oaIngridUtils	1.7.2
oaGeomInfo	5.10
oaGeometryLookup	0.7.1
oaChanInfo	3.18
oaApplyAlign	2.6
oaMagnetCalib	4.18

Simulation

simMaster	2.20
nd280Geant4Sim	6.11
selectEventSim	0.15.1
cosmicTriggerGeant4Sim	1.9.1
atmPitSim	1.9.1
sandGeant4Sim	0.9.2
neutGeant4CascadeInterface	1.5.1

Calibration

calibMaster	1.64
eventUnpack	3.9.1
fgdRawDataApplyCalib	2.7.1
tfbApplyCalib	3.51.1
dataQuality	1.31
beamData	0.33
eventCalib	4.12
smrdApplyCalib	1.36
fgdExtractCalib	0.16
detResponseSim	6.29.1
tpcExtractCalib	0.8

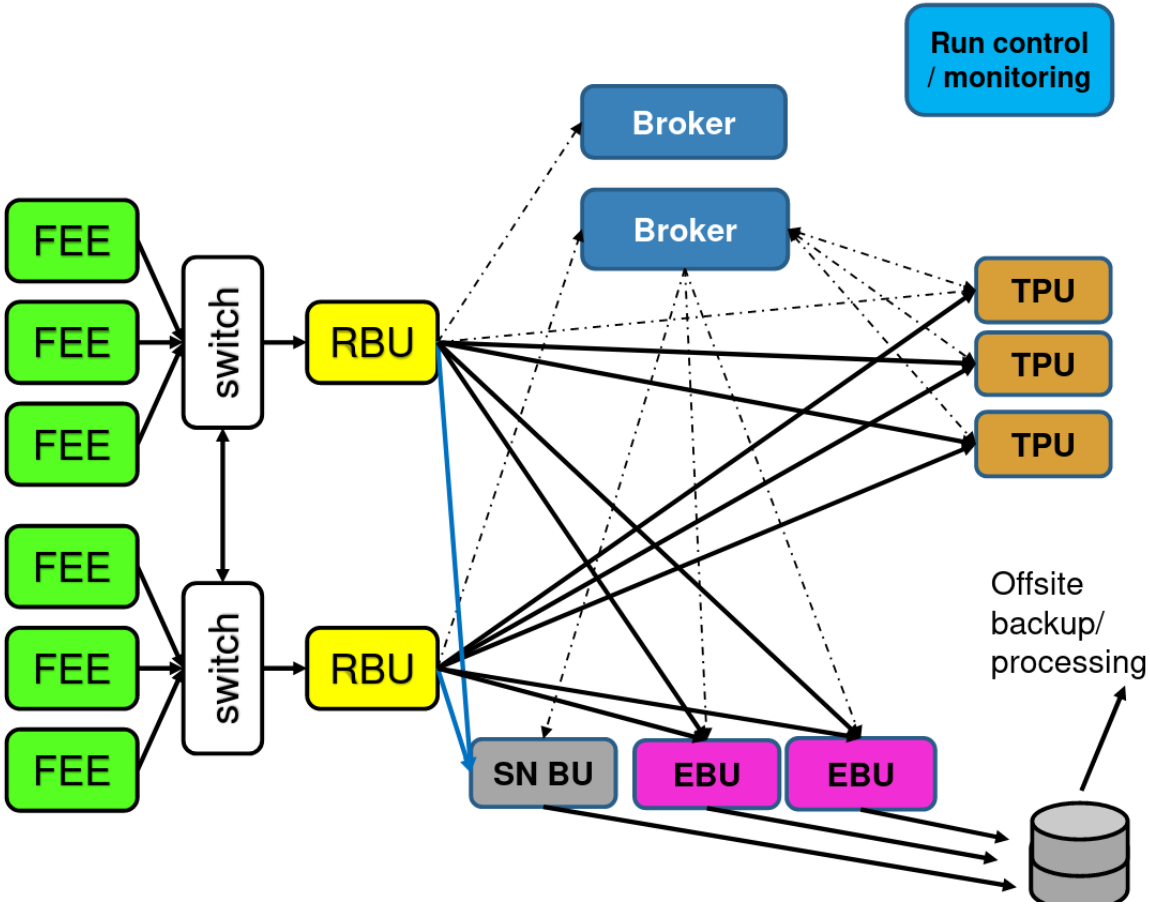
Reconstructio

reconMaster	4.50
reconUtils	1.35.1
RECPACK	4.17.1
recPackRecon	8.57
sbcatRecon	5.5.1
p0dRecon	9.12
tpcRecon	6.34
trexRecon	2.36
fgdRecon	6.11
smrdRecon	5.14
trackerRecon	3.5.1
ecalRecon	11.10
p0decalRecon	2.10
eventRecon	5.12
ingridRecon	2.6

Externals

externalsMaster	1.83
nd280SoftwarePolicy	3.8
CLHEP	2.1.1.0
Geant4	10.1.03.01
MYSQL	5.6.20.01
ROOT	6.20.04.03
GSL	1.15.0.00
CERNLIB	2005.9
NEUT	5.4.0.00

DAQ



Simplified schematic of the DAQ framework

→ work in progress

FEE: front end electronics

RBU: Readout buffer unit

EBU: Event builder unit

SNBU: SN builder unit

TPU: Trigger processing unit

GPU T2K Compute canada

2020/04 – 2021/03: 7.45 GPU years

ND280 CPU

e.g. Jan 2019 - Dec 2019
(haven't updated these numbers for 2020, but it was similar)

Sum CPUT Time
- total: 914,928
- busiest month: Sept: 217,806

Sum CPU Work HS06
- total: 9,238,823
- busiest month: Sept: 2,137,187

Very spikey CPU usage
* 2x2 months of high usage
* 2x2 months of medium usage
* low usage rest of the time

Usually we can get around 2,000 single core jobs running at once
(though this starts to decrease if we are running for a week or so)

With the help of GridPP we have increased this to ~4,000 during the odd emergency :)