Discovery of θ_{13} at Daya Bay using NERSC & ESNet

Craig E. Tull, PhD

US Lead of Daya Bay Offline Computing

NERSC Users' Group Meeting, Berkeley

February 13, 2013

1st major US/China experiment

US PI – Kam-Biu Luk (LBNL)
CN PI – Yifang Wang (IHEP)

North America (16)

LBNL, BNL, Caltech, Iowa State Univ., Illinois Inst. Tech., Princeton, RPI, Siena, UC-Berkeley, UCLA, Univ. of Cincinnati, Univ. of Houston, Univ. of Wisconsin-Madison, Univ. of Illinois-Urbana-Champaign, Virginia Tech., William & Mary

Europe (2)

JINR, Dubna, Russia
Charles University, Czech Republic

Asia (20)

IHEP, Beijing Normal Univ., Chengdu Univ. of Sci. and Tech., CGNPG, CIAE, Dongguan Univ.Tech., Nanjing Univ., Nankai Univ., NCEPU, Shandong Univ., Shanghai Jiao tong Univ., Shenzhen Univ., Tsinghua Univ., USTC, Zhongshan Univ., Univ. of Hong Kong, Chinese Univ. of Hong Kong, National Taiwan Univ., National Chiao Tung Univ., National United Univ.

~230 Collaborators
US/China ~50/50

Daya Bay Nuclear Power Complex

Why θ_{13}?

❖ Based on an assumption of three generations, a 3x3 neutrino mixing matrix was proposed – PMNS.

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13}e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- **The SM has no prediction power on the values of these mixing angles and the CPV phase. It relies on experimental input.**
- **θ_{13} was the last unobserved mixing angle.**
- **Provide knowledge of the basic assumptions:**
  - **— The unitarity of PMNS matrix**
  - **— Three generations of neutrinos**
- **A critical input for other researches, for example:**
  - **— Search for leptonic CP violation**
  - **— Determine the neutrino mass hierarchy**
  - **— Understand the 'effective' neutrino Majorana mass limit**

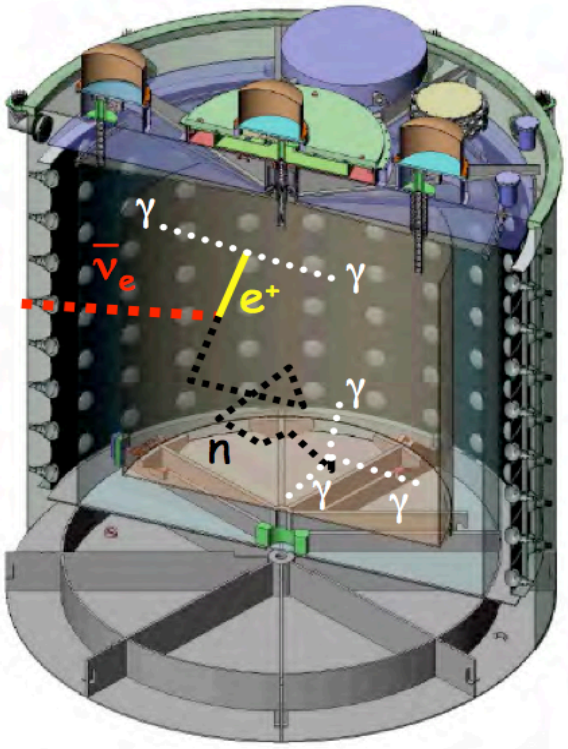
Detecting Reactor $\bar{\nu}_e$

- Use the inverse β-decay reaction:

$\bar{\nu}_e + p \rightarrow e^+ + n$ (prompt signal)

~180μs
$\rightarrow + p \rightarrow D + \gamma(2.2\ MeV)$ (delayed signal)

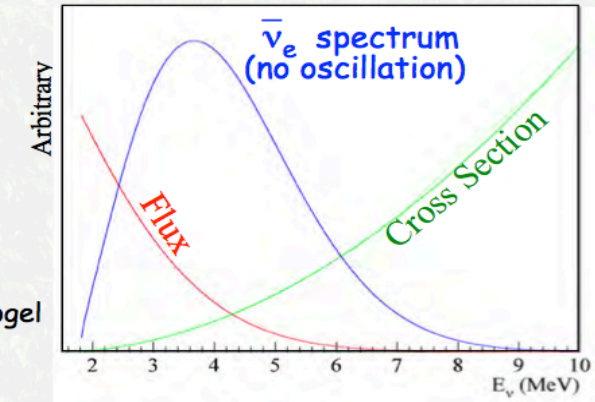
$\rightarrow + Gd \rightarrow Gd^*$
~30μs for 0.1% Gd
$\rightarrow Gd + \gamma$'s(8 MeV) (delayed signal)



- Time- and energy-tagged signal is a good tool to suppress background events.
- Energy of $\bar{\nu}_e$ is given by:

$E_{\bar{\nu}} \approx T_{e+} + T_n + (m_n - m_p) + m_{e+} \approx T_{e+} + 1.8\ MeV$

10-40 keV



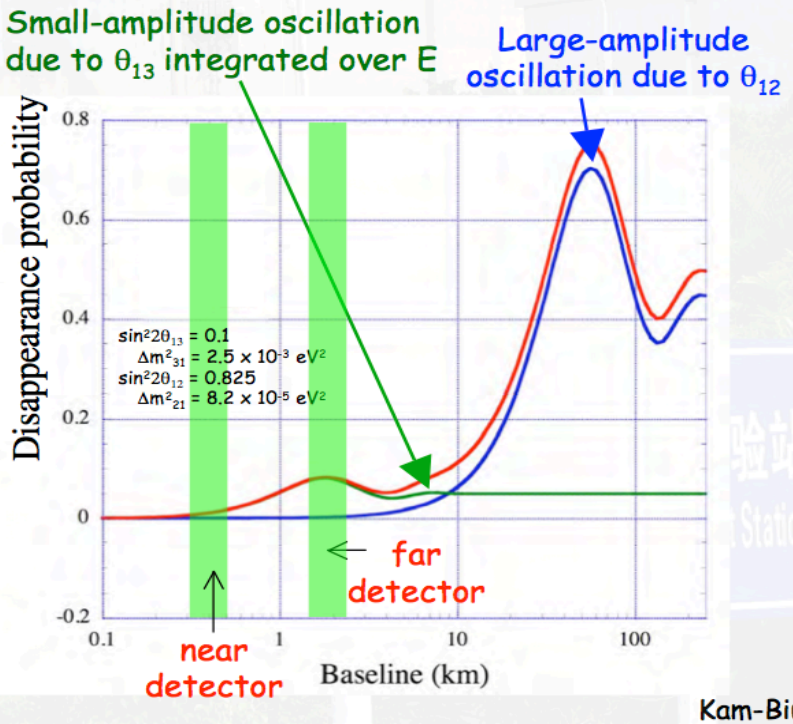
From Bemporad, Gratta and Vogel

Kam-Biu Luk LP2011 6

Reactor Measurement of θ_{13}

- High statistics: powerful nuclear reactors, big detectors, long run-time
- Optimize baselines:

$$P(\bar{\nu}_e \rightarrow x) \approx \sin^2 2\theta_{13} \sin^2\left(\frac{\Delta m^2_{31} L}{4E}\right) + \cos^4 \theta_{13} \sin^2 2\theta_{12} \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right)$$



- Reduce systematic uncertainties:
  - detector-related: 'identical' detectors, careful calibration
  - reactor-related: relative measurement with near and far detectors

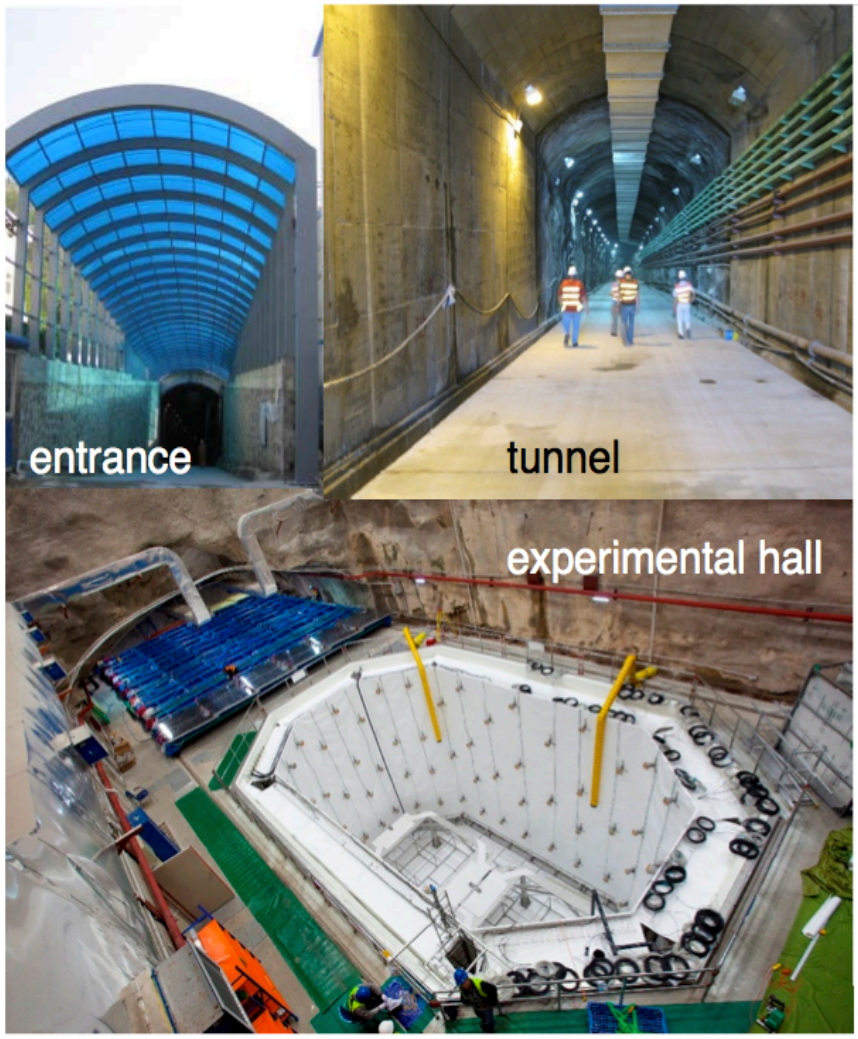
$$\frac{R_{Far}}{R_{Near}} = \left(\frac{L_{Near}}{L_{Far}}\right)^2 \left(\frac{N_{Far}}{N_{Near}}\right)\left(\frac{\varepsilon_{Far}}{\varepsilon_{Near}}\right)\left(\frac{P_{Far}(L_{Far})}{P_{Far}(L_{Far})}\right)$$

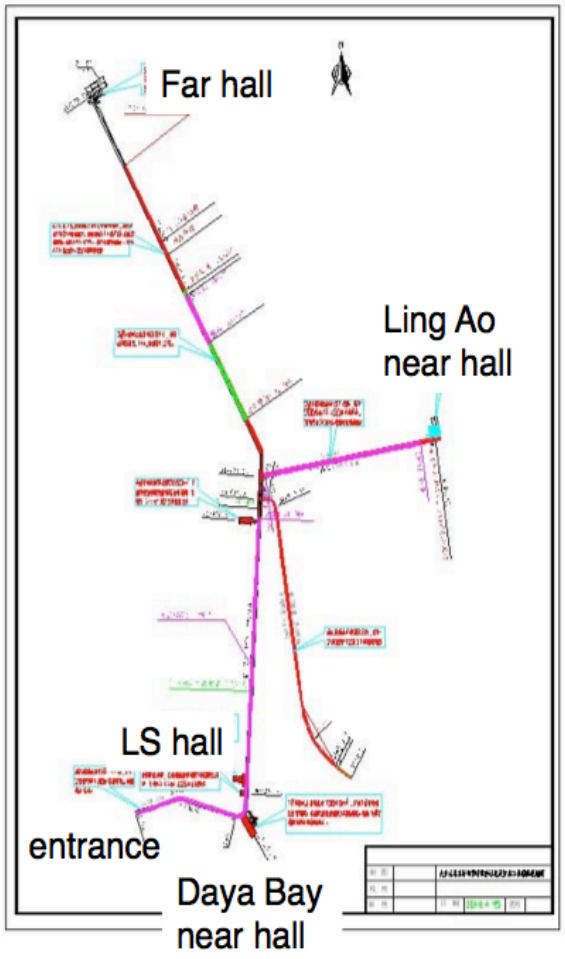
$\bar{\nu}_e$ flux | $1/r^2$ | number of protons | detection efficiency | yield $\sin^2 2\theta_{13}$

- Reduce background: shield and veto

Kam-Biu Luk LP2011 7

Daya Bay Underground Laboratory



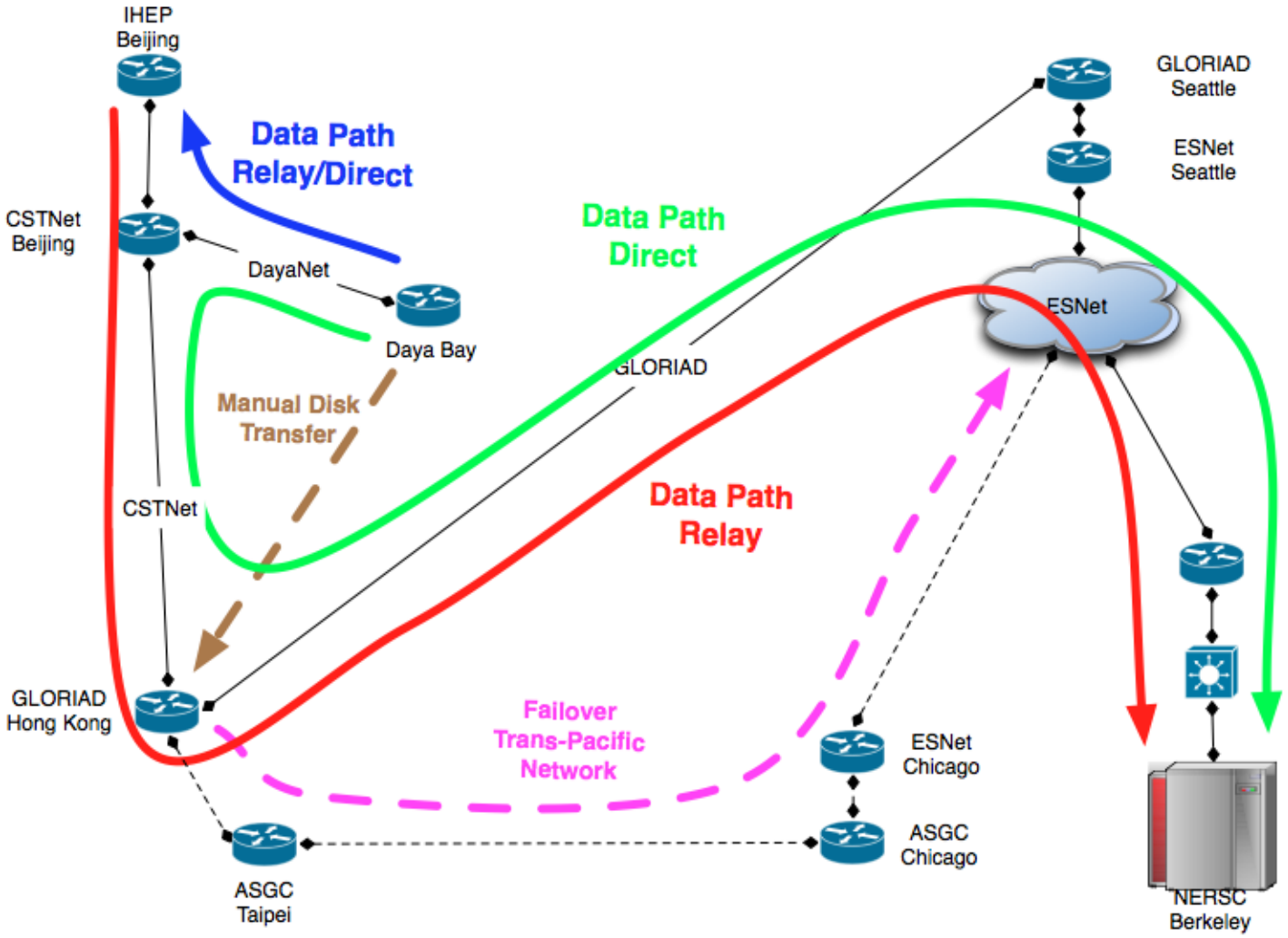


NERSC is US Tier 1

- 10 Onsite Offline Servers: Spade, disk cache, Offline DB, PQM, Web, …
- Network: Mix of national, international, and institutional
  - DayaNet (OC3) stable at 150 Mbps (Dec 2011)
  - CSTNet – Chinese national network
  - GLORIAD – Trans-Pacific network
  - ESNet – US-DOE national network
- PDSF (NERSC): 350 cores, 624 TB disk (604 TB used), >600 TB on HPSS
- IHEP (Beijing): 248 cores (64 bit), 161 TB disk (95 used), Castor Tape
- Offsite Servers: DBs, SVN, Web, Spade, 6 NuWa build slaves, eLog, …

~125 TB raw data per year

Daya Bay Networking

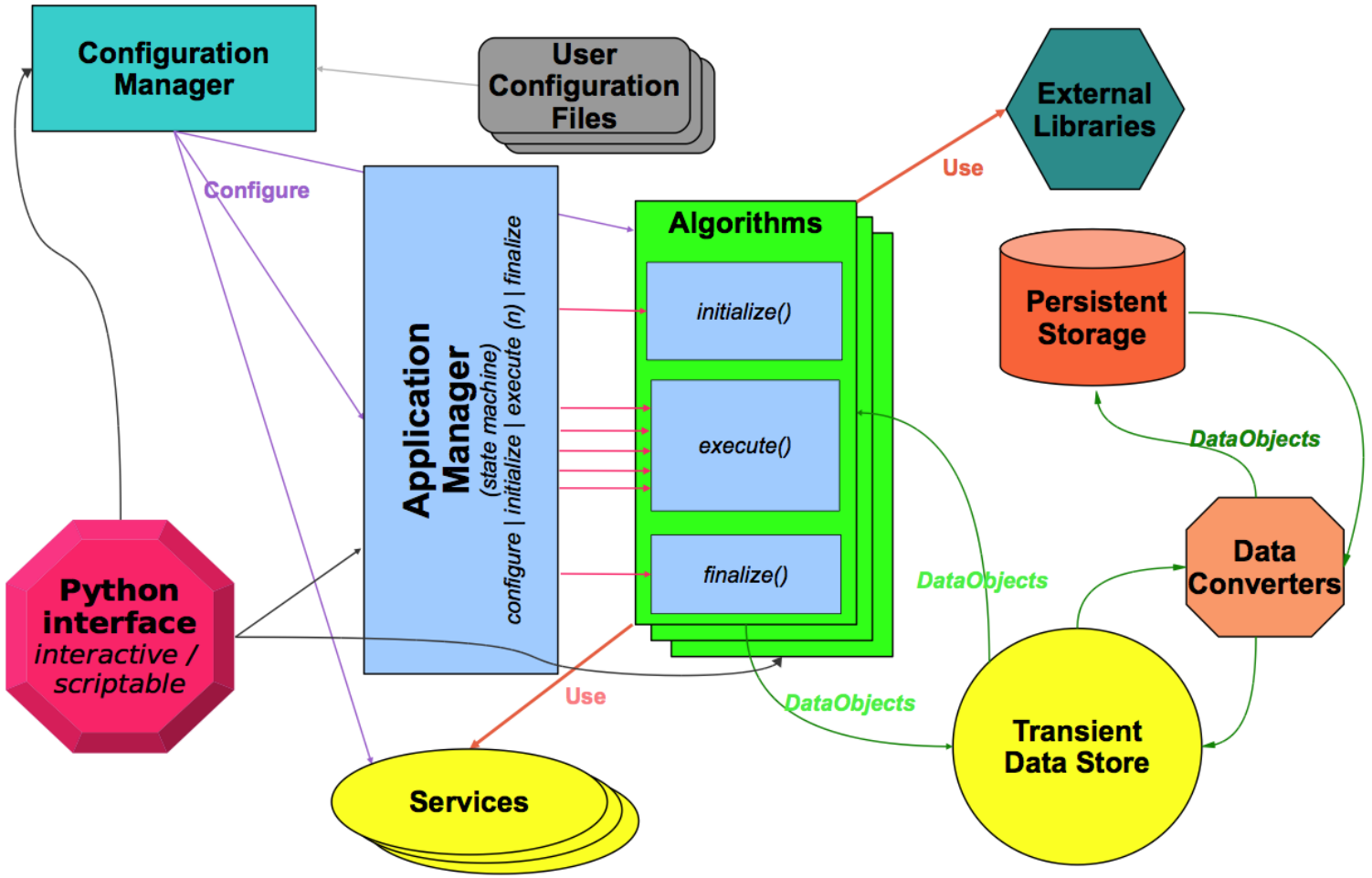


- Relay Path requires daisy-chaining 2 data transfers (default)
- Direct Path requires 2 data transfers out of site.
- DayaNet: Dedicated 150 Mbps optical link
- CSTNet: Chinese national network
- GLORIAD: Trans-Pacific scientific network (NSF)
- ASGC: Trans-Pacific eScience network (fallback for GLORIAD)
- ESNet: US national Energy Science Network
- Hot-swappable disk transport between Daya Bay and Hong Kong in case of long-term network failure of either DayaNet or CSTNet

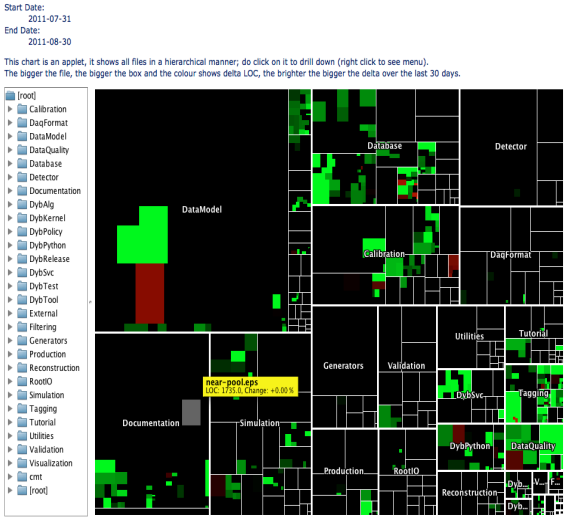
Daya Bay applications are reusable by design

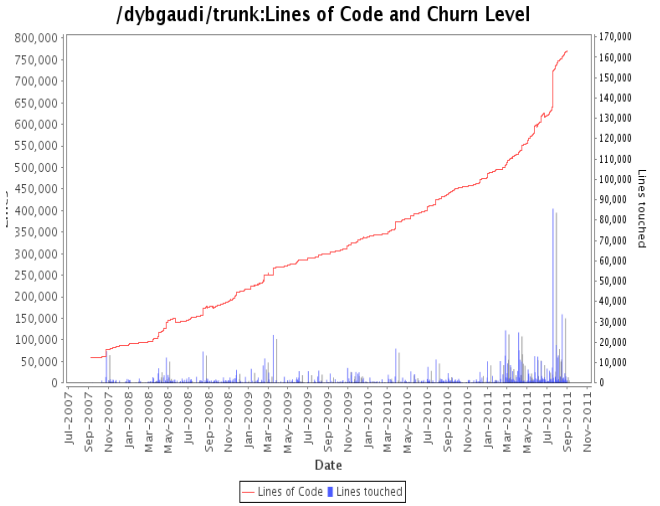
- We operate in a reusable software ecosystem.
- NuWa (Gaudi) – Extended LHCb / ATLAS Architecture
  - Component simulation and analysis framework
  - Repurposed from 'event' to 'time-window' model.
- DBI (DataBase Interface) – Extended from MINOS
  - Interval-of-Validity DB with roll-back/provenance
- Spade – Extended (major rewrite) from IceCube
  - Manages data workflow from detector to warehouse.
- P2 (psquared) – Custom built
  - Manages execution of file processing.
- ODM – Custom built with standard components
  - Near-time (2 hr.) feedback for scientists
- Geant4 (NuWa) – HEP / NP standard
  - Simulation engine for detector/physics studies
- ROOT (NuWa) – HEP / NP standard
  - Interactive data exploration & libraries

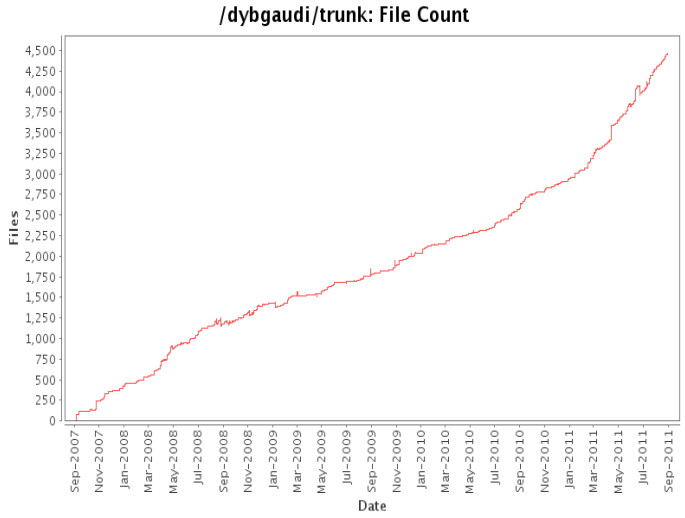
NuWa - Gaudi-based Framework

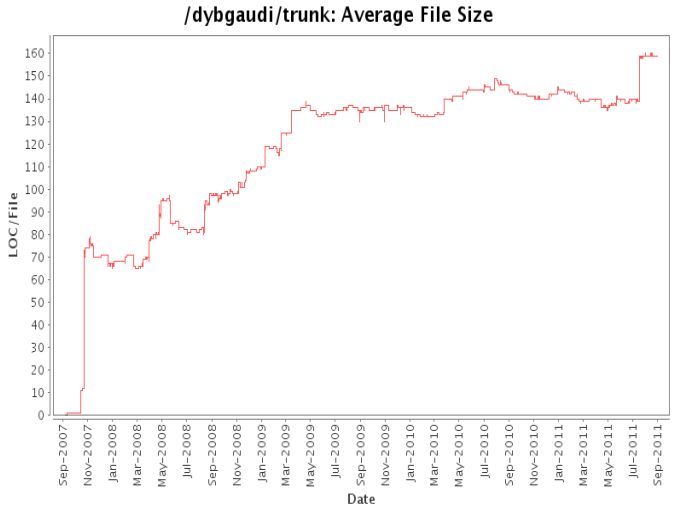


NuWa SVN Statistics (Jan 2012)



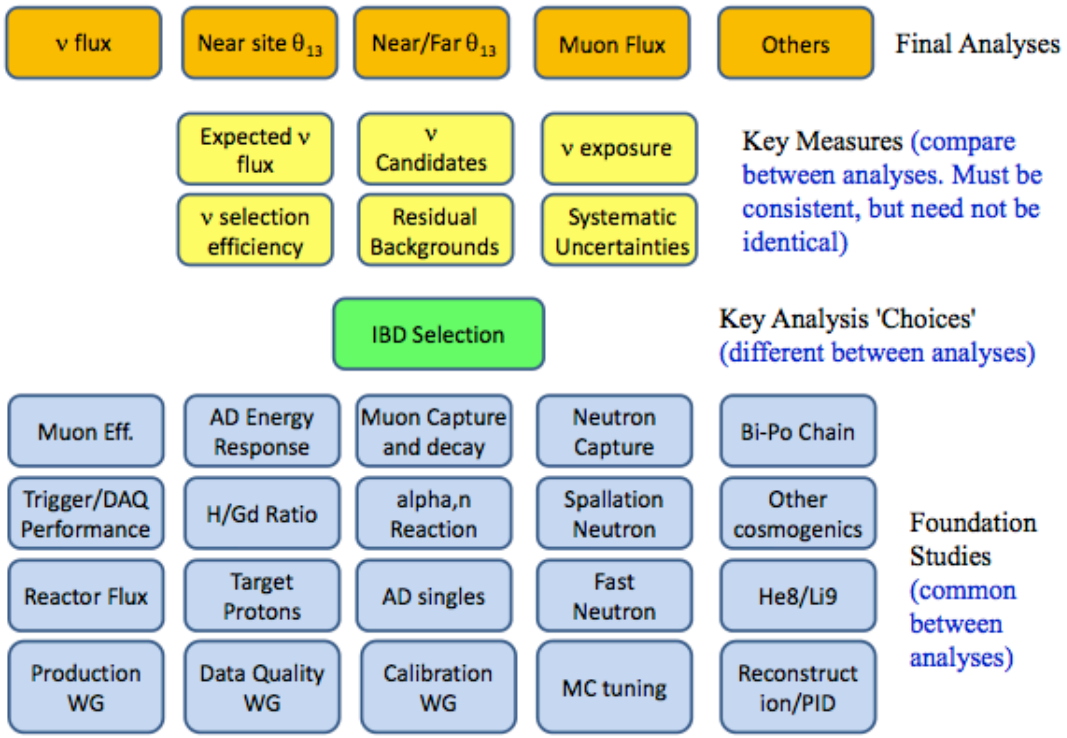




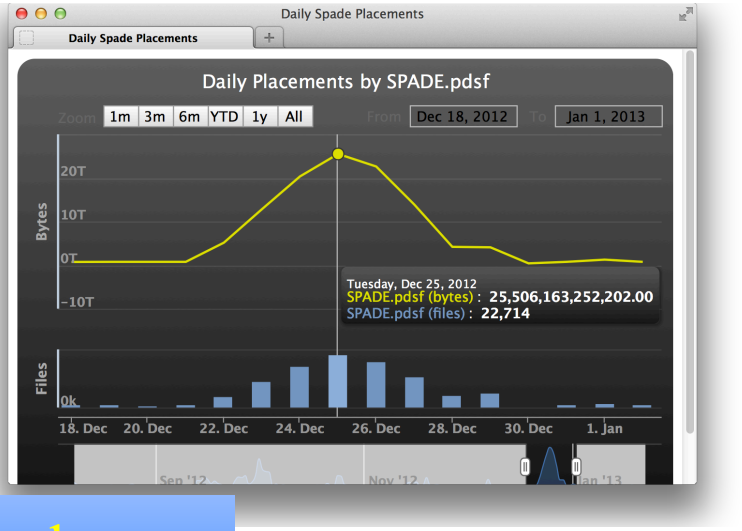


NuWa allows multiple independent analyses

- March result rate-only analysis (blind analysis) was the result of 5 competitive internal analysis campaigns.
- Share foundation, but differ in:
  - Energy calibration & reconstruction
  - Candidate selection/ efficiency
  - Background studies
- Shared services, comparibility, internal consistency, and reproducibility led to the best-of-breed analysis and physics results in record time.

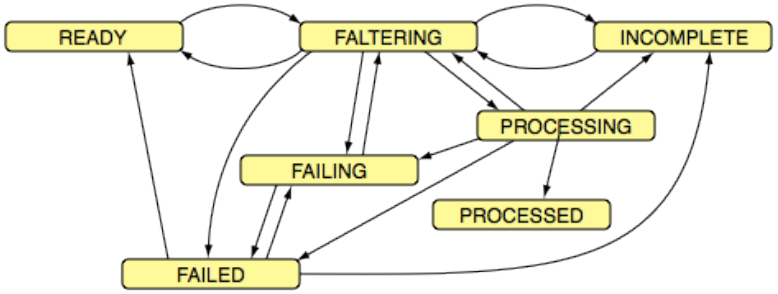


Interoperating data-driven workflow components

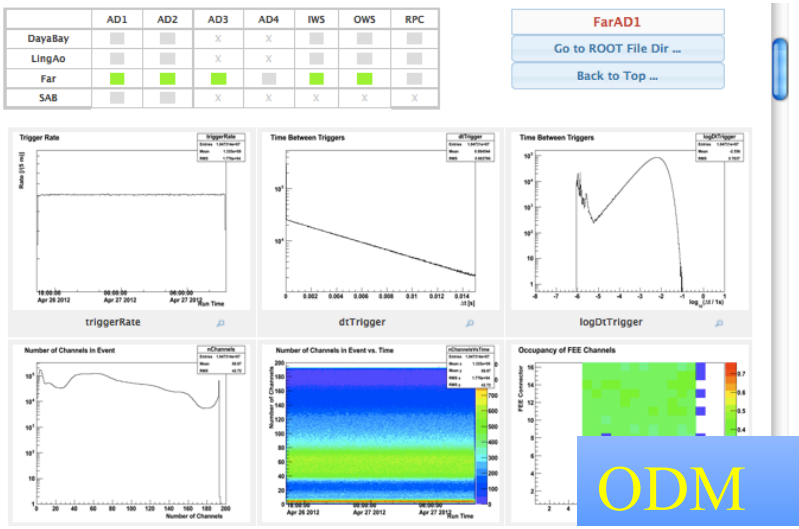


Spade

- Components work together, but operate independently.
- Enhances robustness & recoverability from outages.
- Allows evolution of components.

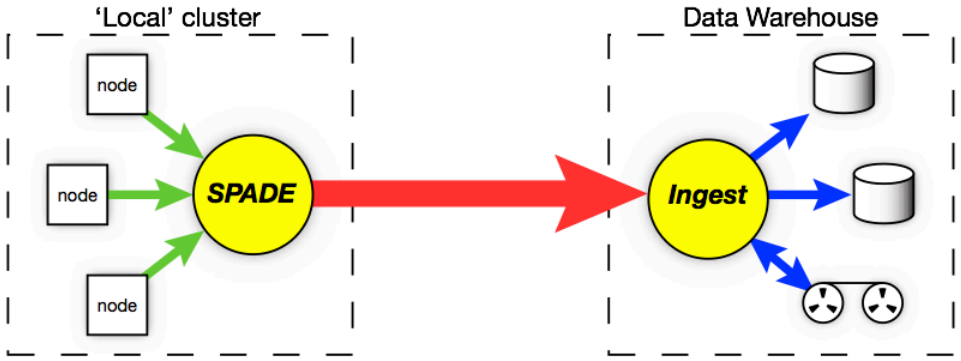


PSquared



ODM

Data Transfer & Management

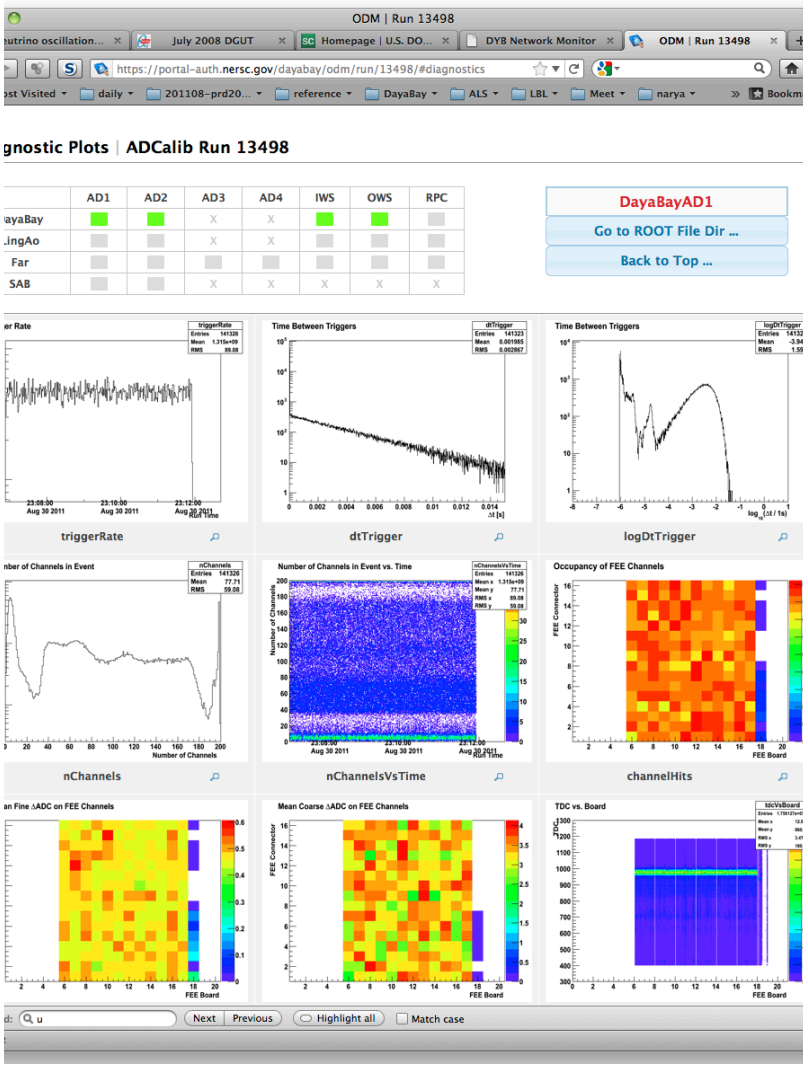


daq.Neutrino.0021508.Physics.EH1-Merged.SFO-1._0001

Tracking of daq.Neutrino.0021508.Physics.EH1-Merged.SFO-1._0001	
Closed by DAQ	Sat, Jan 7 2012, 16:34
IHEP warehouse	Sat, Jan 7 2012, 16:45
PDSF warehouse	Sat, Jan 7 2012, 16:51
HPSS at PDSF	Sat, Jan 7 2012, 16:51
KUP at PDSF	Sat, Jan 7 2012, 18:28

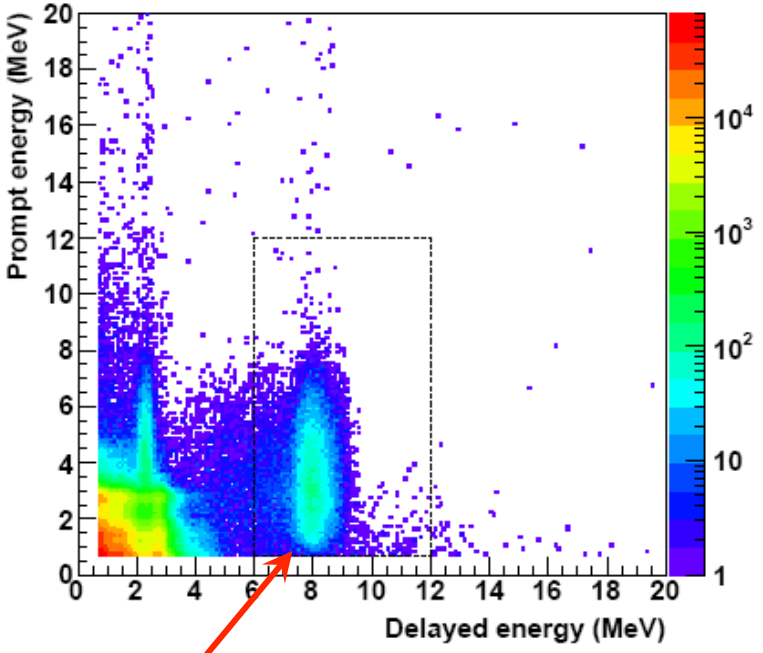
- Spade/Ingest suite is designed to reliably transfer data from an experiment to it's data warehouse.
- We have 2 production data warehouses (IHEP & LBNL) containing all raw data and all production processed data.
  - Web query interface and accessible by all collaborators
- Nominal data transfer latencies from Daya Bay are ~10-15 minutes (IHEP) and ~15-20 minutes (LBNL) after file is closed by DAQ.
- Data is buffered onsite and at IHEP to accommodate network interruptions (>30 days buffer at Daya Bay).
- Keep Up Production (KUP), data archiving and conversion to ROOT are triggered automatically by Ingest.

Offline Data Monitor (ODM)



- Realtime analysis of data at NERSC
  - Uses Science Data Gateway
- Data transferred over DayaNet => CSTNet => GLORIAD => ESNet => GPFS => HPSS in less than 30 minutes by Spade.
- Automated KUP (Keep Up Production) jobs launched by Spade and managed by PSquared.
- NuWa production job using DBI and Offline DB execute in ~1.5 Hours.
  - Runs on PDSF (or Carver)
- Offline, DCS, and DAQ DBs scrapped, aggregated with onsite ELog entries and output of NuWa.
- Available globally to collaborators.
- Reference plots and IHEP plots used to validate consistency.
- Data and Root files can be downloaded from GPFS/HPSS.
- Best tool for immediate feedback on physics quality and used by onsite shifters and QA (offsite) shifters.

Robust infrastructure enabled science discovery

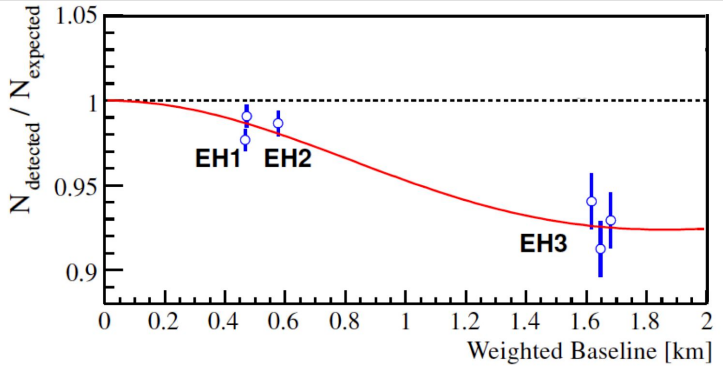


antineutrinos

- 24Dec2011: 6-AD running
- First θ_{13} result with 54.8 days of data (16 TB).
- US scientists at full parity with Chinese.
- Observed anti-neutrinos within 24 hours.
- Saw anti-neutrino deficit within 7 days.
- PRL submitted within 75 days (20 days after last file).
- 08Mar2012: Announce θ_{13}

Robust infrastructure enabled science discovery

"Physics-ready on day one."



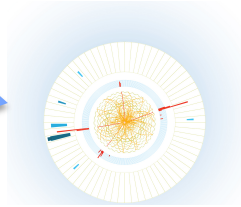
The Daya Bay experiment counts antineutrinos at detectors in three experimental halls (EH) near the Daya Bay nuclear reactor and calculates how many would reach the detectors if there were no oscillation. The plot shows measured disappearance of antineutrinos at the halls as a function of distance from the reactor. The 6.0% rate deficit at EH3 provides clear evidence of the new transformation.

- **24Dec2011: 6-AD running**
- **First θ_{13} result with 54.8 days of data (16 TB).**
- **US scientists at full parity with Chinese.**
- **Observed anti-neutrinos within 24 hours.**
- **Saw anti-neutrino deficit within 7 days.**
- **PRL submitted within 75 days (20 days after last file).**
- **08Mar2012: Announce θ_{13}**

Science AAAS One of *Science* Magazine's Top 10 Breakthroughs of 2012

Data Breakthroughs in 2012

- **AAAS Breakthrough of the Year 2012**
  - **—Higgs Boson**
- **Runners Up**
  - **—Denisovan DNA**
  - **—Genome Engineering**
  - **—Last Neutrino Mixing Angle**
  - **—ENCODE**
  - **—Curiosity Landing**
  - **—X-ray Laser Protein Structure**
  - **—Brain-Machine Interfaces**
  - **—Majorana Fermions**
  - **—Eggs from Stem Cells**



Data's role in high impact science is deep and getting deeper. Now is the time to bring the power of DOE computing at scale to transform data analysis

DEMO

- **Onsite Network:**
  - http://nms.dyb.ihep.ac.cn/nms/index-dyb.html
- **Onsite Spade Transfers:**
  - http://dayabay.lbl.gov/dybspade/daily-transfers/index.html?ihep
- **PDSF Spade Placements:**
  - http://dayabay.lbl.gov/dybspade/daily-placements/index.html?pdsf
- **Spade Summary:**
  - http://dayabay.lbl.gov/dybspade/daqdm/summary/
- **ODM (Offline Data Monitor):**
  - https://portal-auth.nersc.gov/dayabay/odm/
  - http://dybdq.ihep.ac.cn//odm/
- **PDSF Throughput and Performance:**
  - http://www.nersc.gov/users/computational-systems/pdsf/
  - http://portal.nersc.gov/project/pdsf/ganglia/

THANK YOU
(XIE-XIE 谢谢)