Computing Framework and MC Production

Francesca Di Lodovico (QMUL)
6th Open Meeting
Kashiwa, IPMU, 28-31 Jan 2015

Overview

Ongoing effort focussing on:

Software release
- Need software for the collaboration
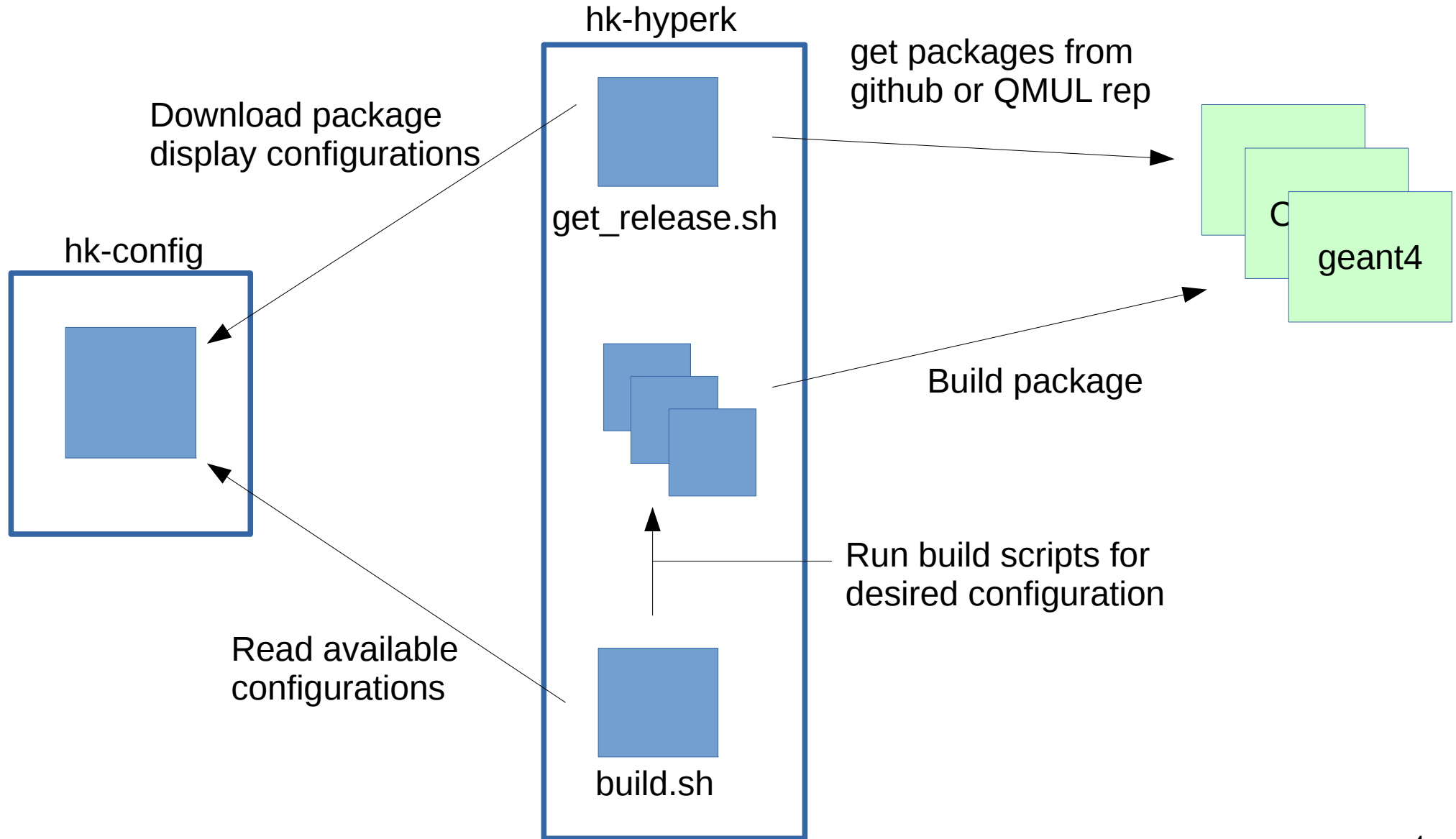- Easy to install
- Comprehensive

Production
- Automatic production
- Write scripts to run on the Grid
- Simulation production

Release Structure

- Instructions documented in: https://wiki.hyperk.org/Software/Release

- If you need the account on the wiki, please drop me an email <f.di.lodovico@qmul.ac.uk>.

- hk-hyperk package downloads release packages from GitHub (public packages) and QMUL git (private packages).

- The above structure is totally transparent to the users.

- Structure of hk-hyperk scripts to download and build s/w recently simplified

  - One build script per package.

  - Configuration package contains mapping between packages and config (eg PROD = clhep, geant4, root, wcsim, fitqun, irods).

  - Versions of configuration package for each production release. Makes it easier for people to be in synch with production.

  - Can have more than one configuration package for different types of detectors (e.g. at the moment TITUS) or operations each with own release cycle.

Release Structure

hk-hyperk

get packages from
github or QMUL rep

get_release.sh

Download package
display configurations

hk-config

geant4

Build package

Run build scripts for
desired configuration

Read available
configurations

build.sh

Hyper-K Beam Event MC Production

- Used beam flux from Beam WG (M.Hartz)

- Vertex generator code from Okajima => STANDARD horizonthal tank

- NEUT 5.3.1 (latest release used in T2K)

- Used latest available packages for MC production:

  - CLHEP 2.1.0.1

  - ROOT v5-34-23

  - GEANT4 4.9.4p04

  - WCSim v1.2.0

  - FiTQun v4r2

- New Version of WCSim on 22/Jan

- Deployed and produced 100K ν-mode and 100k $\bar{\nu}$-mode events by 25/Jan

- Overall fairly quick, regarless of a few hiccups with the Grid.

MC Production

- Production made use of UK Grid nodes

  - Imperial, QMUL, RAL (more nodes coming soon).

  - Plan to include other Countries very soon.

- As before 1 simu job = 10K events

  - Now 1 reco job = 400 events (25 reco jobs/simu)

- Nomenclature as before:

  - Simu: neutvect-<run>-<part>.root (always neutvect-<n>-<n>.root)
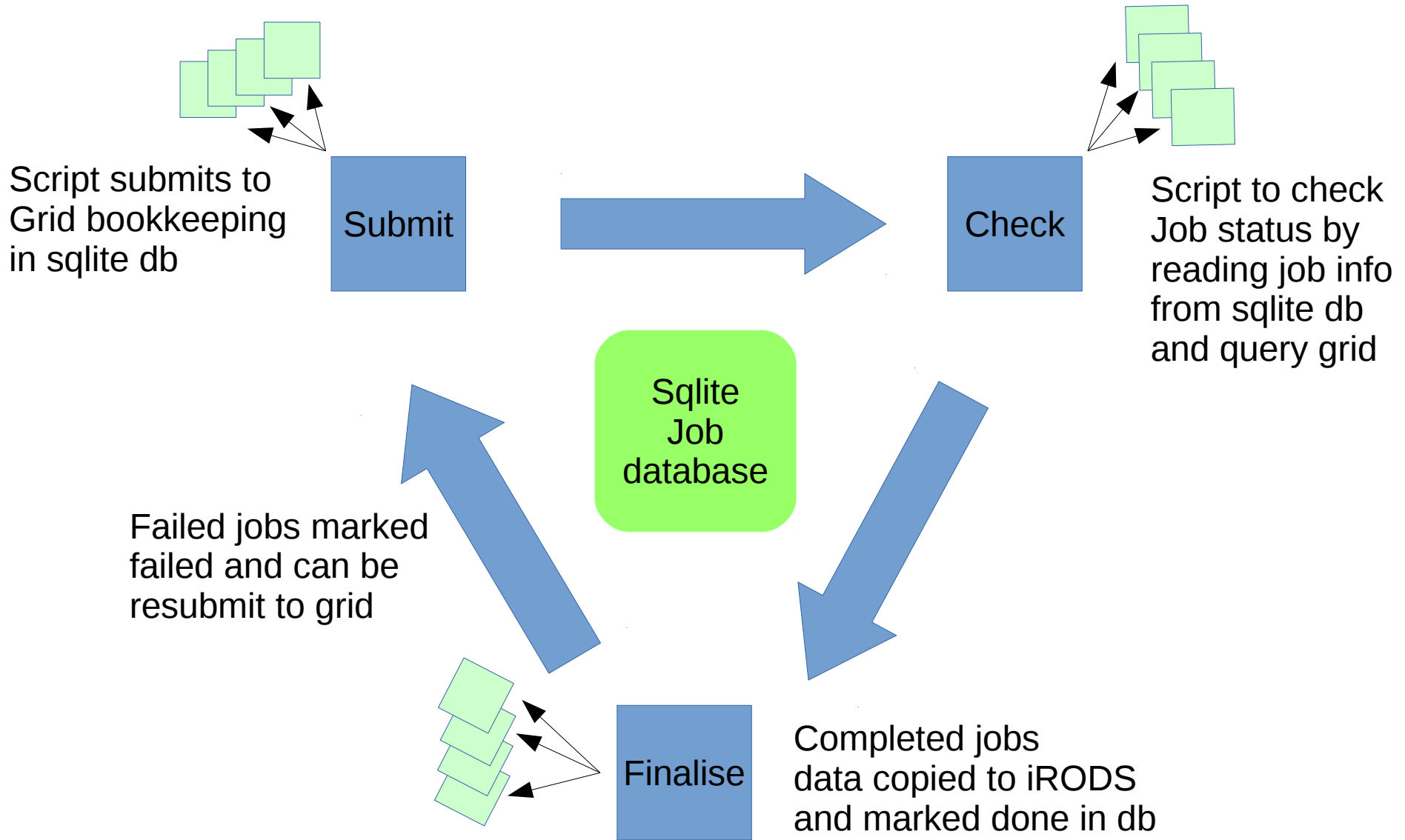
  - Reco: fitqun-<run>-<part>.root

Time for reco jobs
from when they are
submitted

MC Production (reconstruction)

6
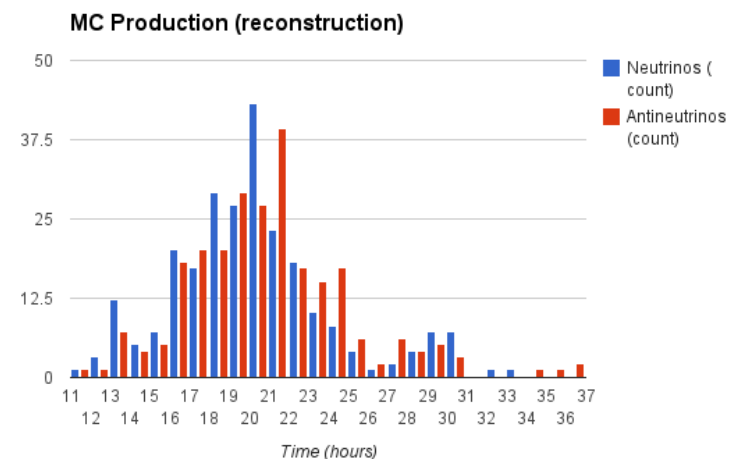
MC Production

- Files stored on the iRODS data management system.

- No Grid certificated needed. Just email <f.di.lodovico@qmul.ac.uk> to get the account. Information in: https://wiki.hyperk.org/Software/iRODSHK

- Simulation files in iRODS:
  - /QMULZone2/home/hyperk/production/numode/simu/P1a

  - /QMULZone2/home/hyperk/production/antinumode/simu/P1a

- Reconstruction files in iRODS:

  - /QMULZone2/home/hyperk/production/numode/reco/P1a

  - /QMULZone2/home/hyperk/production/antinumode/reco/P1a

- It is your responsibility to check these files before increasing the production.

- Info on the Wiki: https://wiki.hyperk.org/Software/Production

Production Lifecycle

Script submits to
Grid bookkeeping
in sqlite db

Submit

Check

Script to check
Job status by
reading job info
from sqlite db
and query grid

**Sqlite
Job
database**

Failed jobs marked
failed and can be
resubmit to grid

Finalise

Completed jobs
data copied to iRODS
and marked done in db

8

Production Lifecycle

- Failed jobs for many reasons:

  - Proxy certificate expired

  - Problems with compute node

  - Problems with job (not seen these)

  - Failure to access release filesystem

- Recognise many things can go wrong in complex distributed system so made resubmission very easy.

Production Lifecycle

- Scripts and db logic forbids reco to be submit unless simulation successful.

- Scripts to generate jobs use template files and write details to sqlite database.

- Continue to try to reduce as much as possible management overhead
  - People have more important things to do

Summary of How to Get the Files

- Register on the Wiki (https://wiki.hyperk.org/) to gather the most up-to-date information.
- Current production on the Wiki: https://wiki.hyperk.org/Software/Production
- To get the files you need an iRODS account. Instructions in: https://wiki.hyperk.org/Software/iRODSHK
- You can access vector, simulation and reconstruction files.

Neutrino beam (replace numode with antinumode for antinu beam):
- The Vector files are in the iRODS collection:
/QMULZone2/home/hyperk/production/vector/numode/P1a
- The Simulation files are in the iRODS collection:
/QMULZone2/home/hyperk/production/numode/simu/P1a
- The Reconstruction files are in the iRODS collection:
/QMULZone2/home/hyperk/production/numode/reco/P1a
- Each reconstruction job contains 400 events and there are 25 reconstruction files for each simulation file.

Summary of How to Get the Files

Names of the files:

Vector File	Simulation (10k/file)	Reconstruction (400/file) <n>=0-24
t2hk_320a_fluka2011_vectors_1.dat	neutvect-0-0.root	fitqun-0-<n>.root
t2hk_320a_fluka2011_vectors_2.dat	neutvect-1-1.root	fitqun-1-<n>.root
t2hk_320a_fluka2011_vectors_3.dat	neutvect-2-2.root	fitqun-2-<n>.root
t2hk_320a_fluka2011_vectors_4.dat	neutvect-3-3.root	fitqun-3-<n>.root
t2hk_320a_fluka2011_vectors_5.dat	neutvect-4-4.root	fitqun-4-<n>.root
t2hk_320a_fluka2011_vectors_6.dat	neutvect-5-5.root	fitqun-5-<n>.root
t2hk_320a_fluka2011_vectors_7.dat	neutvect-6-6.root	fitqun-6-<n>.root
t2hk_320a_fluka2011_vectors_8.dat	neutvect-7-7.root	fitqun-7-<n>.root
t2hk_320a_fluka2011_vectors_9.dat	neutvect-8-8.root	fitqun-8-<n>.root
t2hk_320a_fluka2011_vectors_10.dat	neutvect-9-9.root	fitqun-9-<n>.root

Grid Jobs Future

- WMS (workload management system) being phased out.

- Replacement will be Dirac (distributed computing Grid system, currently used by LHCb). Will look at interfacing the scripts to this new service once test system in place.

Future Production

- Future production will follow directions from the tank/physics (others?) groups regarding the needed types of events and vector files.

- Simulation and reconstruction will be provided by the WCSim/FitQun and BONSAI groups.

- Will organize a small group for validation within the software group.

- After validation, the files will be released to the physics groups for analysis.

Conclusions

- Streamlined tools to decrease the overhead.

- Ongoing progress, they will continue to improve.

- Possible to download/compile HK release automatically.

- Produced neutrino and antineutrino files (100k+100k) with latest release. For everyone to check!

- All documentation is on the Wiki.

- Getting ready for future production.

Backup slides

Numbers

Simulation CPU time:~1.7 sec/ev
Reconstruction CPU time:~95 sec/ev

1 vector size: 3.2MB
1 simu file size: 235MB
1 reco file size: 800kB