Migration to new Information System and decommissioning the BDII

Alastair Dewhurst, Alessandra Forti

Motivation

- Since December 2012 there have been discussions about replacing the BDII.

  - It's too complex and doesn't provide VOs with the information they need.

- In June 2015 OSG announced plans to stop using the BDII.

  - OSG BDII has been decommissioned since 31 / 3 / 2017.

- Since September 2015, "No effort will be put in [to the development of the] BDII as the idea is to reduce its dependencies".

  - We are not recommending any new communities that GridPP may support rely on the BDII for information.

- Even if a BDII is simple to run, it is still an extra piece of grid middleware.

  - How many know the ldap command to extract the wallclock time of a queue off the top of their head?

  - Higher fraction of effort required at small sites.

UK BDII instances

Site	Site Bdii	Top Bdii
ECDF	info4.glite.ecdf.ed.ac.uk	
BHAM	epgr09.ph.bham.ac.uk	
Liverpool	hepgrid4.ph.liv.ac.uk	
Manchester	site-bdii.tier2.hep.manchester.ac.uk	top-bdii.tier2.hep.manchester.ac.uk
Sheffield	lcg.shef.ac.uk	
Durham	site-bdii.dur.scotgrid.ac.uk	
Brunel	dc2-grid-68.brunel.ac.uk	
RHUL	sbdii2.ppgrid1.rhul.ac.uk	
Cambridge	vserv02.hep.phy.cam.ac.uk	
Bristol	lcgbdii02.phy.bris.ac.uk	
Sussex	grid-bdii-02.hpc.susx.ac.uk	
Lancaster	py-fjalar.hec.lancs.ac.uk	
QMUL	bdii02.esc.qmul.ac.uk	
Oxford	t2bdii06.physics.ox.ac.uk	
Imperial	bdii.grid.hep.ph.ic.ac.uk (2)	topbdii.grid.hep.ph.ic.ac.uk (2)
Glasgow	svr030.gla.scotgrid.ac.uk	
RALPP	site-bdii.pp.rl.ac.uk	
RAL Tier-1	site-bdii.gridpp.rl.ac.uk (2)	lcgbdii.gridpp.rl.ac.uk (2)
TOTAL hosts	20	5

Alastair Dewhurst, 3rd August 2018

UK BDII instances

Site	Site Bdii	Top Bdii
ECDF	info4.glite.ecdf.ed.ac.uk	
BHAM	epgr09.ph.bham.ac.uk	
Liverpool	hepgrid4.ph.liv.ac.uk	
Manchester	site-bdii.tier2.hep.manchester.ac.uk	top-bdii.tier2.hep.manchester.ac.uk
Sheffield	lcg.shef.ac.uk	
Durham	site-bdii.dur.scotgrid.ac.uk	
Brunel		
RHUL		
Cambridge		
Bristol		
Sussex		
Lancaster		
QMUL	bdii02.esc.qmul.ac.uk	
Oxford	t2bdii06.physics.ox.ac.uk	
Imperial	bdii.grid.hep.ph.ic.ac.uk (2)	topbdii.grid.hep.ph.ic.ac.uk (2)
Glasgow	svr030.gla.scotgrid.ac.uk	
RALPP	site-bdii.pp.rl.ac.uk	
RAL Tier-1	site-bdii.gridpp.rl.ac.uk (2)	lcgbdii.gridpp.rl.ac.uk (2)
TOTAL hosts	**20**	**5**

~0.5 FTEs of effort spread across GridPP to keep this infrastructure running.

InfoSys Evolution TF

- An WLCG Information System Evolution Task Force was formed in July 2015:

  - https://twiki.cern.ch/twiki/bin/view/EGEE/WLCGISEvolution

  - Contributions from Alessandra Forti, Andy McNab, David Meredith (former GOCDB developer).

- OSG decommissioned the BDII and put static information in to Github:

  - e.g. BNL's Configuration

New Information System

New Information System

OSG

GitHub

HT-Condor CE

CREAM CE

ARC CE

EGI

GOCDB

{;} JSON

For OSG sites the JSON is stored in Github. It is committed via pull requests.

HT-Condor CE can automatically generate json. ARC CE can automatically generate json in next major release.
No ETA for when CREAM CE can generate json.

For EGI sites a link to a webserver hosting the json is stored in the GOCDB.

New Information System

- Both ATLAS and CMS are consuming information from OSG sites (via github) for over a year using new Information system.
  - CMS need it less than ATLAS.
- ALICE only ever used Dynamic information from the BDII.
  - This is now provided by CEs directly.
  - No actual need for new system.
- LHCb DIRAC is able to consume information from json format using their GOCDB2CSAgent.
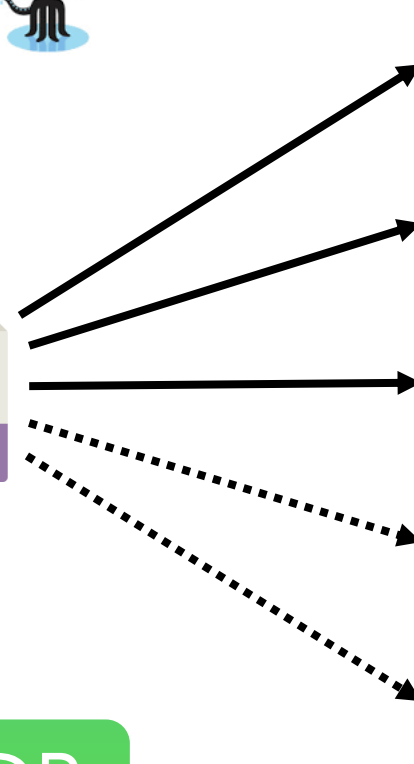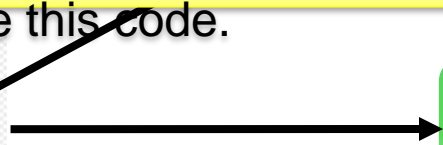 - GridPP DIRAC could hopefully re-use this code.

ATLAS

CMS

LHCb

ALICE

Other

GOCDB

Json

"computingservices" [
    {
   "ce_id": 1234,
   "ce_name": "arc-ce01.gridpp.rl.ac.uk",
   "cs_endpointurl": "https://arc-ce01.gridpp.rl.ac.uk:2811/",
   "cs_flavour": "ARC-CE",
   "cs_version": "5.64",
   "cs_jobmanager": "condor",
   "cs_jobmanager_version": "2",
   "cs_status": "production",
   "cs_state": "production",
   "cs_queue_name": "grid3000M",
   "cs_queue_maxcputime": 0,
   "cs_queue_maxwalltime": 345600,
   "cs_queue_maxmainmemory": "8048",
   "cs_queue_maxrunningjobs": "25000",
   "cs_assigned_vo": ["ATLAS", "CMS", "LHCb", "ALICE",],
   "cs_message" : "some free form text" }
…
 ],

Manual vs Automated

- Time taken to manually edit json file ~5 minutes.

- The information is static, so only needs updating when site make a significant change.

  - Doing more than one update a month seems unlikely…

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE? (ACROSS FIVE YEARS)
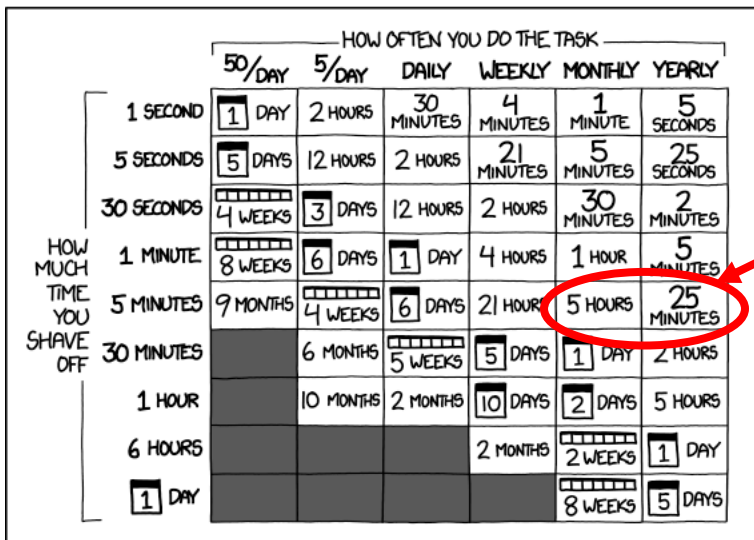
Not automating the json update will cost between 30 minutes and half a day of work for each site admin over the next 5 years.

Keeping the automated script running may use more time.

Running a BDII is significantly more effort than either.

https://xkcd.com/1205/

Alastair Dewhurst, 3rd August 2018
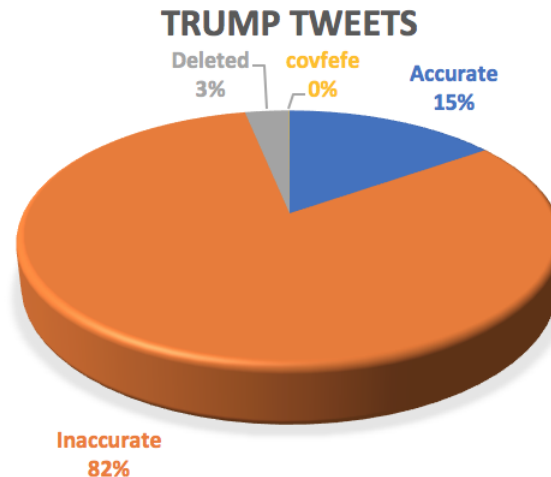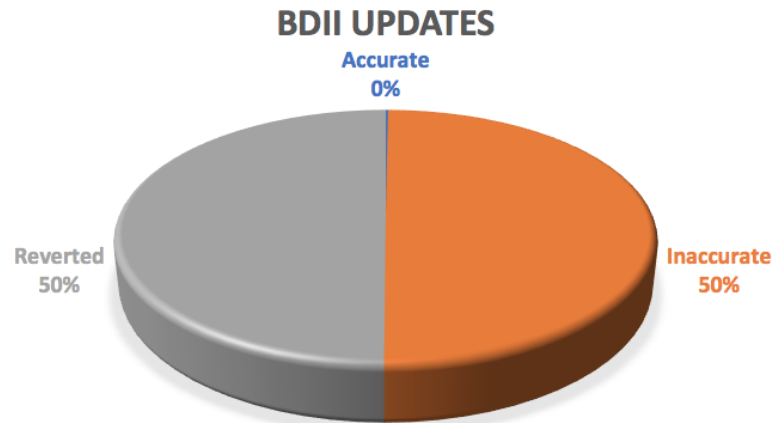
Info System updates

- ATLAS Grid Information System polls BDII every 2 hours.

  - All updates are logged.

- Comparison between new OSG system and BdII has been performed over last 1.8 years of updates.

- New OSG system updates each Panda Queue on average once every 5 years.

- The vast majority of BDII changes are flip-flopping between two values.

  - We assume a correct one and a "default" one when the automated updates breaks.

	EGI (BDII)	OSG (New InfoSys)
Number of Panda Queues	722	349
Maxwallclock time updates	32643	75
Status updates	92925	45
Maxcputime updates	31761	1
Total updates per year	**120 / PQ**	**0.2 / PQ**

Alastair Dewhurst, 3rd August 2018

Trump vs BDII

- Even if some sites ensure their information is accurate in the BDII there is too much noise for it to be effectively used by VOs.

  - It has proved too complex to debug.

- The accuracy of BDII updates compares unfavourably with Donald Trump's twitter feed!!

BDII UPDATES

Accurate
0%

Inaccurate
50%

Reverted
50%

TRUMP TWEETS

Deleted
3%

covfefe
0%

Accurate
15%

Inaccurate
82%

https://www.washingtonpost.com/graphics/politics/trump-claims-database/?noredirect=on&utm_term=.d98396d65527
https://www.bustle.com/p/how-many-times-has-trump-tweeted-as-president-twitter-is-his-best-friend-8011368
https://factba.se/topic/deleted-tweets
Alastair Dewhurst, 3rd August 2018

Plan

- 30th August (GridPP41) - Finalize plan for the UK to migrate to the new Information System and decommission BDII.

- 12th September - Announce plan at GDB meeting.

- By 1st October. - All UK sites to have created and published static json file.

  - They can upgrade to automated systems if/when they choose.

- 1st October - EGI broadcast announcing decommissioning of UK BDII service scheduled for February 1st 2019.

  - 4 months* notice + After data taking has ended.

- By end of 2018 - VOs to run a validation that new information they are consuming is accurate.

  - UK has expertise to deal with technical issues as they arise.

- Start of January 2019 - All UK sites announce long downtimes (starting on February 1st) for BDII in GOCDB.

- 1st February 2019 - Decommission BDII at all UK sites.

* 6 months notice would coincide with a possible hard Brexit day so possibly not the best time to make changes!

Alastair Dewhurst, 3rd August 2018

Residual issues

- Some work would be needed for the DIRAC instance at Imperial.

  - Should be able to reuse LHCb DIRAC code.

- Need to migrate LFC users to DIRAC File Catalogue (ongoing).

- Inevitably some problems will be uncovered during the migration.

  - Forgotten about VO still uses BDII?

    - Good reason for them to get in touch with us!

  - Ambiguity in json format?

    - Alessandra Forti wrote current proposal, best placed to make corrections.

  - GOCDB problem?

    - Developer based at RAL.

GridPP Impact

- Migration to new Information System and decommissioning the BdII should be seen as a success (assuming we actually implement it)

  - Leadings roles in Task Force from GridPP.

  - UK well placed for future evolution (e.g. CRIC).

  - Usable system for new communities (e.g. IRIS).

- If not now when?

  - The vast majority of the work has been done, the remainder is progressing extremely slowly.

    - The remaining work will only get done by us actually trying to implement it.

  - Tier-1 has to reduce manpower by April 2019.

  - GridPP6 bid is starting to be written.

Alastair Dewhurst, 3rd August 2018