

# SAMGrid and D-Zero developments

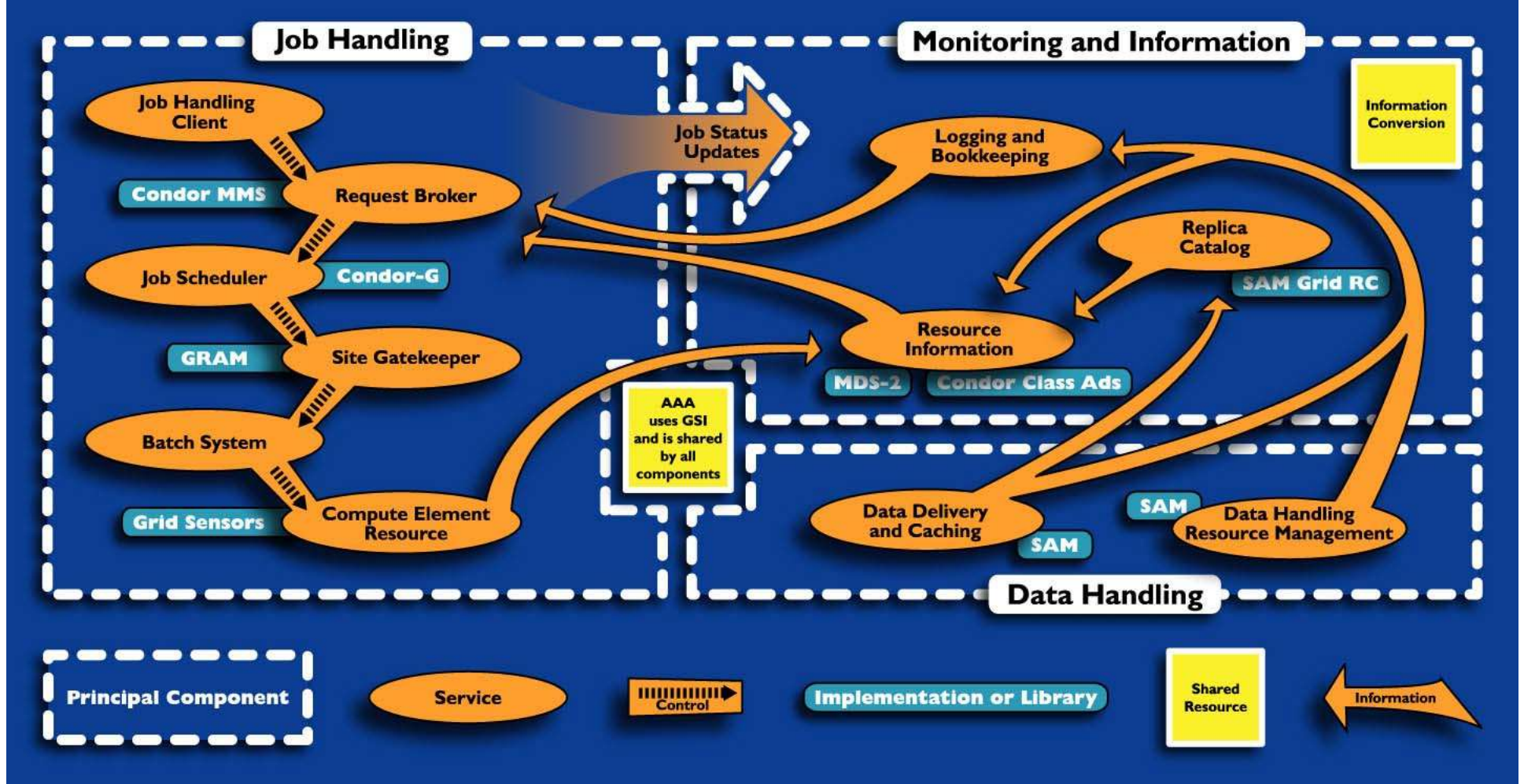
GridPP9 Edinburgh  
4-5 Feb 2004  
Peter Love  
Lancaster University



# SAMGrid

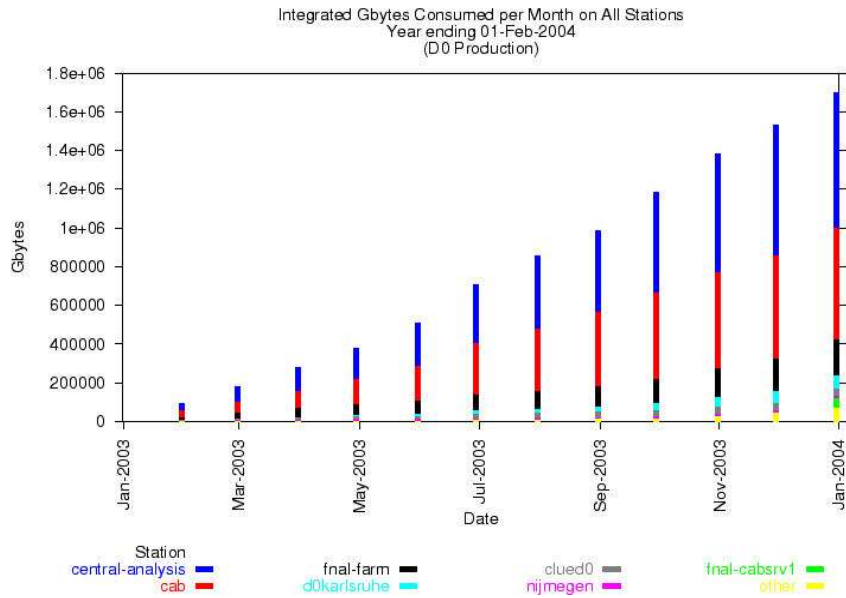
- SAM - Data Handling system used at D-Zero for last seven years
- JIM–Job and Information Management
- JIM packages submit and monitor
- SAM DataGrid + JIM -> SAMGrid computational grid
- [www-d0.fnal.gov/computing/grid/](http://www-d0.fnal.gov/computing/grid/)

# SAM-Grid Architecture



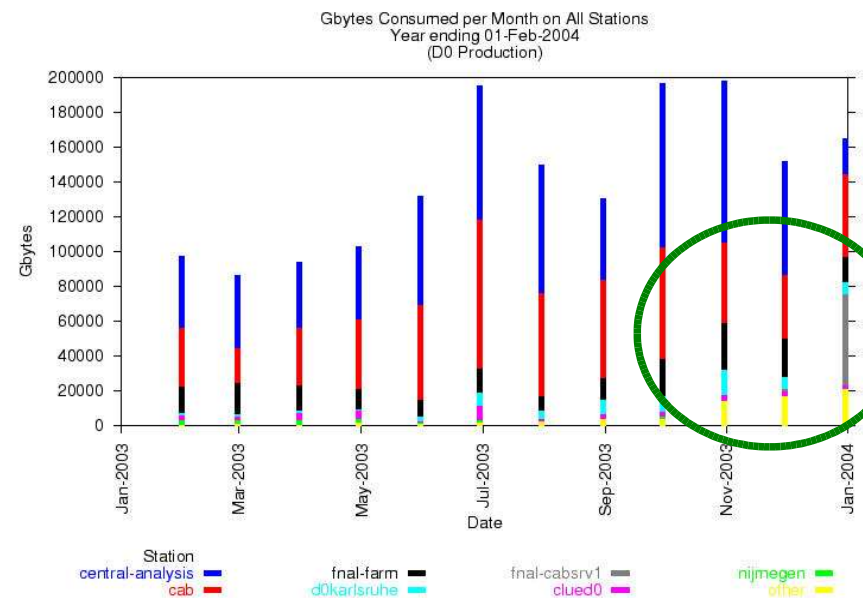
UPD install 1 afternoon (with experts on hand)  
XMLDB or FS configuration

<http://d0db.fnal.gov/sam/plots-and-stats.html>



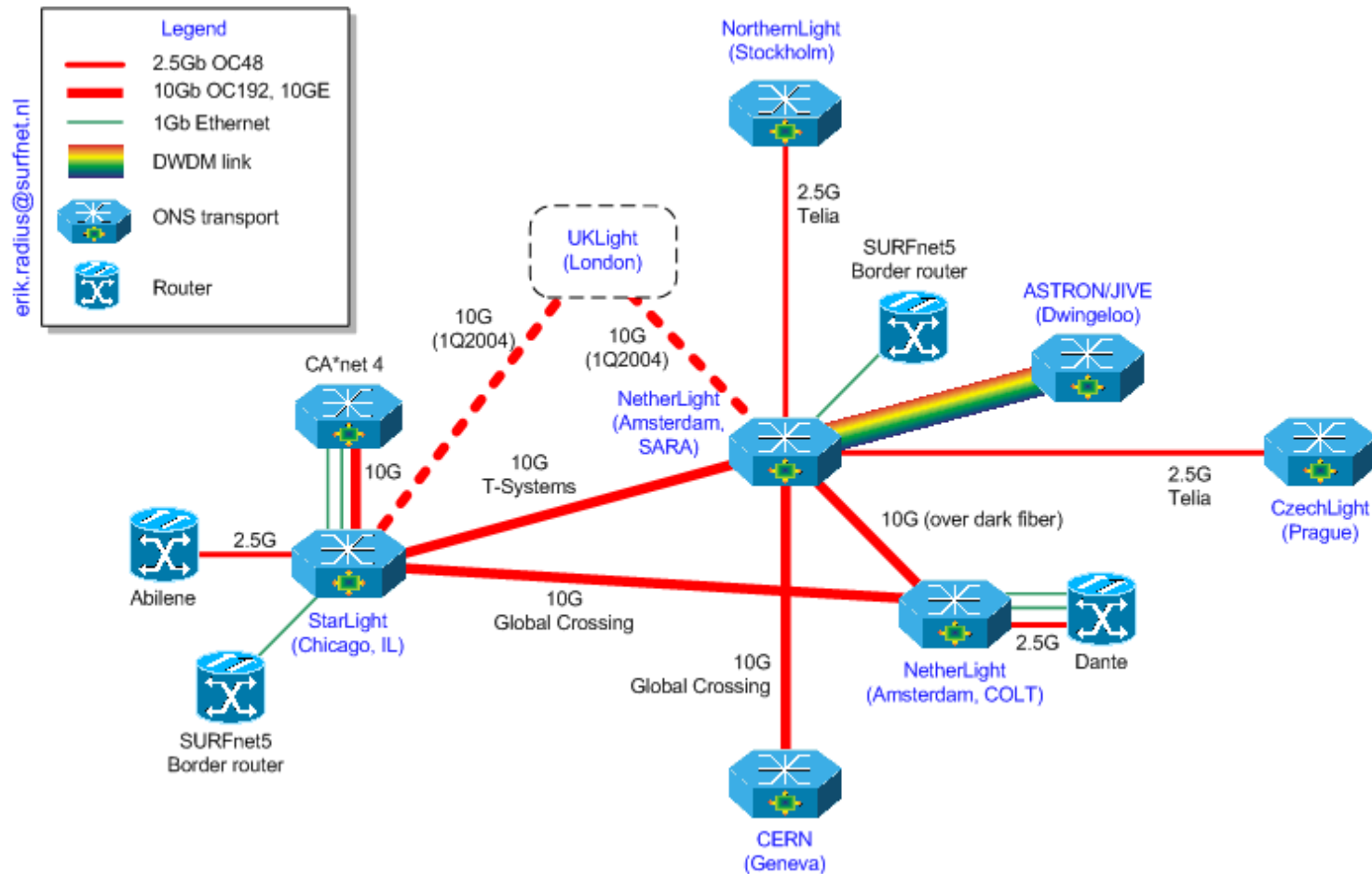
1.8 Pb for 2003

200 Tb peak months



Reprocessing

# UKLight - 10Gbit/s



~ 500 Tb / week?

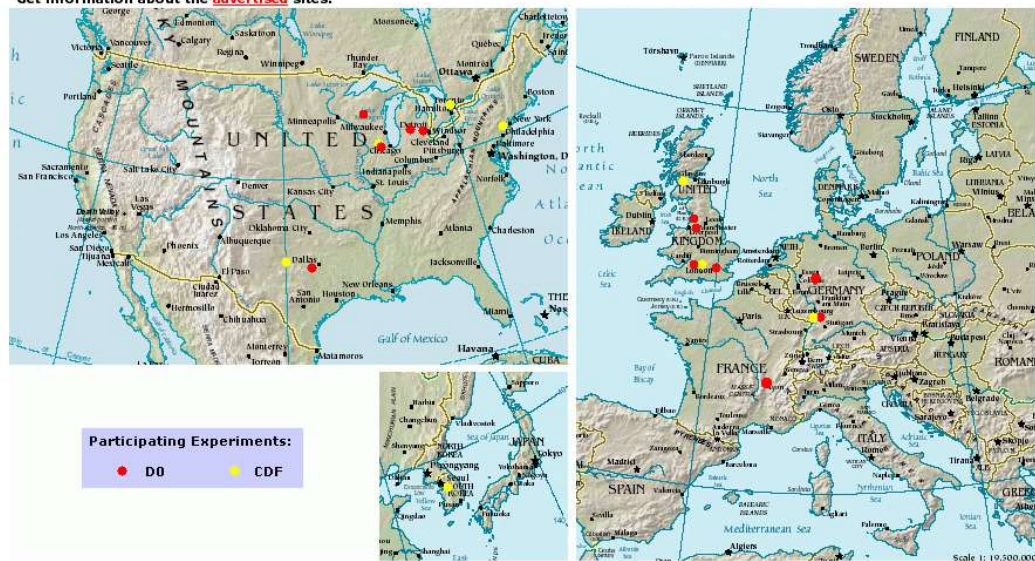
# Monitoring

- Job status via JIM
  - pending, idle, running etc
- Runtime tools
  - task info published to XMLDB
  - specific executable info, num events processed etc.
- Output retrieval via web (small files) and SAM (event data)

<http://samgrid.fnal.gov:8080/>

Launching the Monitoring System:

Please click at the map to monitor the execution sites.  
 Get information about the **submission** sites  
 Get information about the **advertised** sites.



```
File Edit View Terminal Go Help
executable = /home/farm/projects/mystuff/Req2Mac.macros
job_type = mc_runjob
runjob_requestid = 7658
runjob_numevts = 1
sam_experiment = d0
sam_universe = prd
log = mctest.log
group = grid
instances = 1
```

**SAM Grid List of Resources**

Station Name	Type	Site	Universe
d0karlsruhe	SamAnalysis	GridKa	d0
d0rutherford	Monte Carlo	RALd0	d0
d0rutherford	SamAnalysis	RALd0	d0
lancs	Monte Carlo	LANCS	d0
lancs	SamAnalysis	LANCS	d0
imperial-test	Monte Carlo	imperial	d0
imperial-test	SamAnalysis	imperial	d0
d0-wisconsin	SamAnalysis	Wisconsin	d0
cdf-scotgrid-2	SamAnalysis	scotgrid-glw	d0
manhep	Monte Carlo	manchester	d0
manhep	SamAnalysis	manchester	d0
wuppertal	SamAnalysis	wuppertal	d0

Please try clearing your browser cache and **reloading** the page if the information does not seem current.

**SAM Grid Projects at a Submission Site**

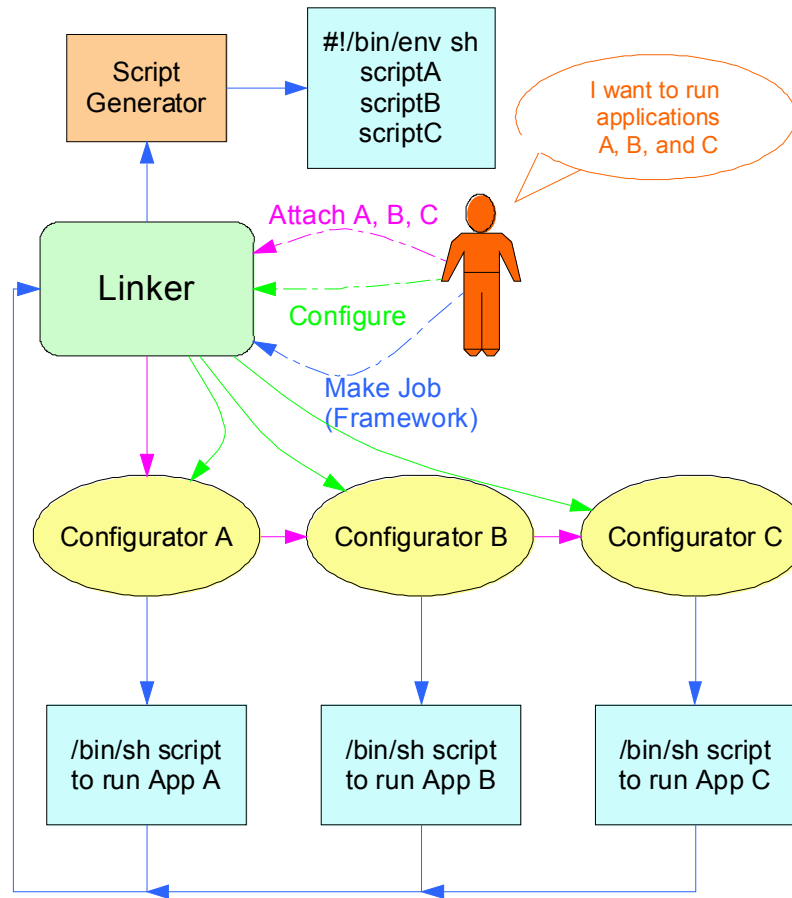
Projects submitted from [fal000phys03.lancs.ac.uk](http://fal000phys03.lancs.ac.uk)

For projects that have been matched with a resource, information becomes available about the execution site, the station and the project's process/consumer details.

Global Job ID	Owner	Status	Type	Execution Site	Station	Universe
farm_fal000phys03.lancs.ac.uk_111545_23540_0	farm	Removed	mc_runjob			
farm_fal000phys03.lancs.ac.uk_124433_20707_0 (BS)	farm	Held	mc_runjob			
farm_fal000phys03.lancs.ac.uk_120624_13491_0 (BS)	Peter_Love	Idle	mc_runjob			
farm_fal000phys03.lancs.ac.uk_121009_13764_0 (BS)	Peter_Love	Idle	mc_runjob			
farm_fal000phys03.lancs.ac.uk_121908_14772_0 (BS)	Peter_Love	Idle	mc_runjob	RALd0	d0rutherford	prd
farm_fal000phys03.lancs.ac.uk_192058_18467_0 (BS)	Peter_Love	Idle	mc_runjob			

Please try clearing your browser cache and **reloading** the page if the information does not seem current.

# mc\_runjob



Workflow and metadata

# mc\_runjob developments v06

- New runtime tools
  - threaded execution and monitoring
  - task health monitoring
  - runtime metadata completion
- New configuration system
  - rc-style system and user settings
- Continued support of user needs

# Current MC Production

- Via global request system
  - Farmers chose and run a request
  - Request defines mc\_runjob macro
  - Results stored into SAM, request marked finished
- One farmer per farm, inefficient

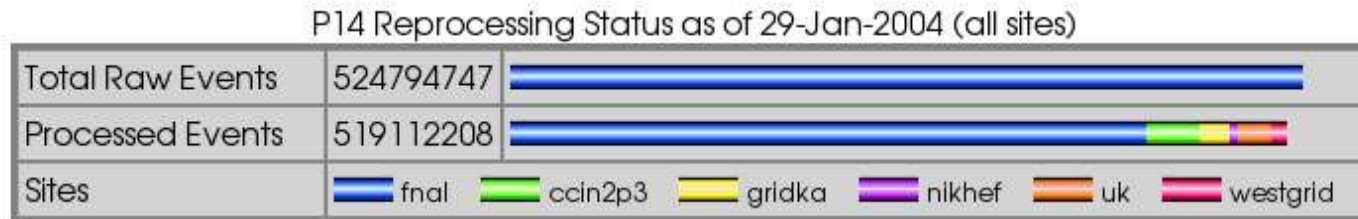
# SAMGrid MC Production hurdles

- Tested at Wisconsin, Manchester, Lyon
- Request system and jobmanager-runjob, specify ReqID, NumEvents
- MinBias and Dzero software installation, now retrieved via SAM
- Need mc\_runjob fixes to handle all MC production requests – fault tolerance
- Temp storage in SAM, merging

Attribute	value
MyType	"Machine"
TargetType	"Job"
Requirements	TRUE
gatekeeper_url_	"fal000phys03.lancs.ac.uk:2119/jobmanager-runjob"
sam_nameservice_	"IOR:000000000000002a49444c3a6f6f632e636f6d2f436f734e616d696e672f4f424e616d696e67436f6e74657874"
Name	"lancs.d0.prd.jobmanager-runjob"
DbURL	"http://fal000phys03.lancs.ac.uk:7080/Xindice"
d0software_cardfiles_v00_02_07	"Installed"
d0software_cardfiles_v00_03_15	"Installed"
d0software_cardfiles_v00_04_06	"Installed"
d0software_cardfiles_v00_04_38	"Installed"
d0software_cardfiles_v00_04_29	"Installed"
d0software_cardfiles_v00_04_53	"Installed"
d0software_cardfiles_p14_02_00	"Installed"
d0software_cardfiles_v00_01_02	"Installed"
d0software_cardfiles_p09_02_00	"Installed"
station_name_	"lancs"
station_experiment_	"d0"
station_universe_	"prd"
jobmanager_name_	"jobmanager-runjob"
gatekeeper_location_	"fal000phys03.lancs.ac.uk:2119"
cluster_architecture_	"Linux+2.4"
cluster_name_	"SamGrid"
schema_version_	"1_0"
site_name	"LANCS"

advert not needed now  
- dynamic install

# First remote reprocessing



<http://www-d0.fnal.gov/computing/reprocessing/>

20% reprocessed off-site, of that 1/4 in UK

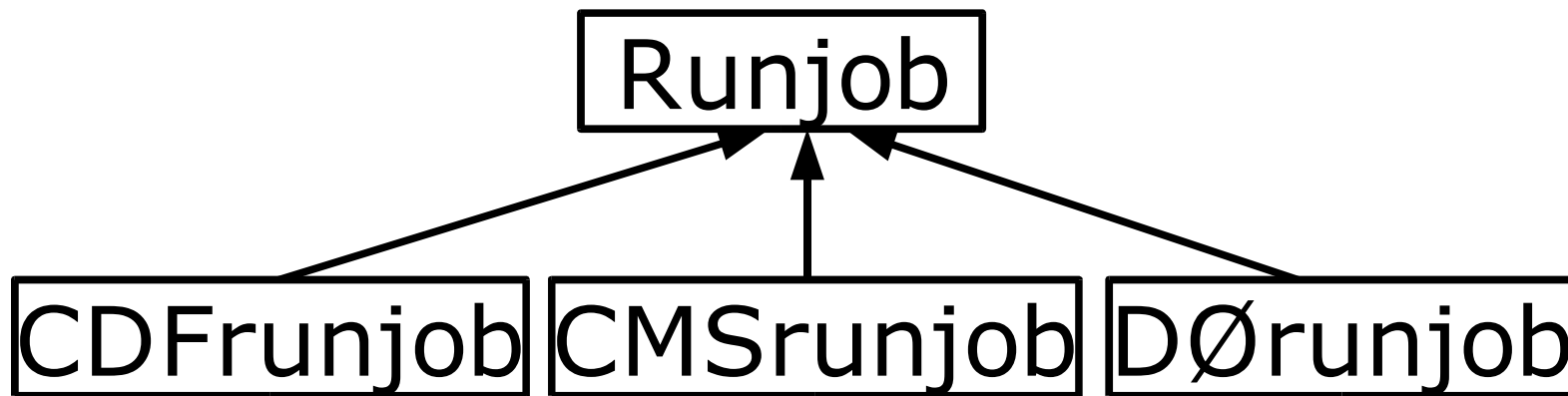
Nikhef *test-run* using LCG at Karlsruhe, RAL, etc.

Data Handling with SAMGrid,

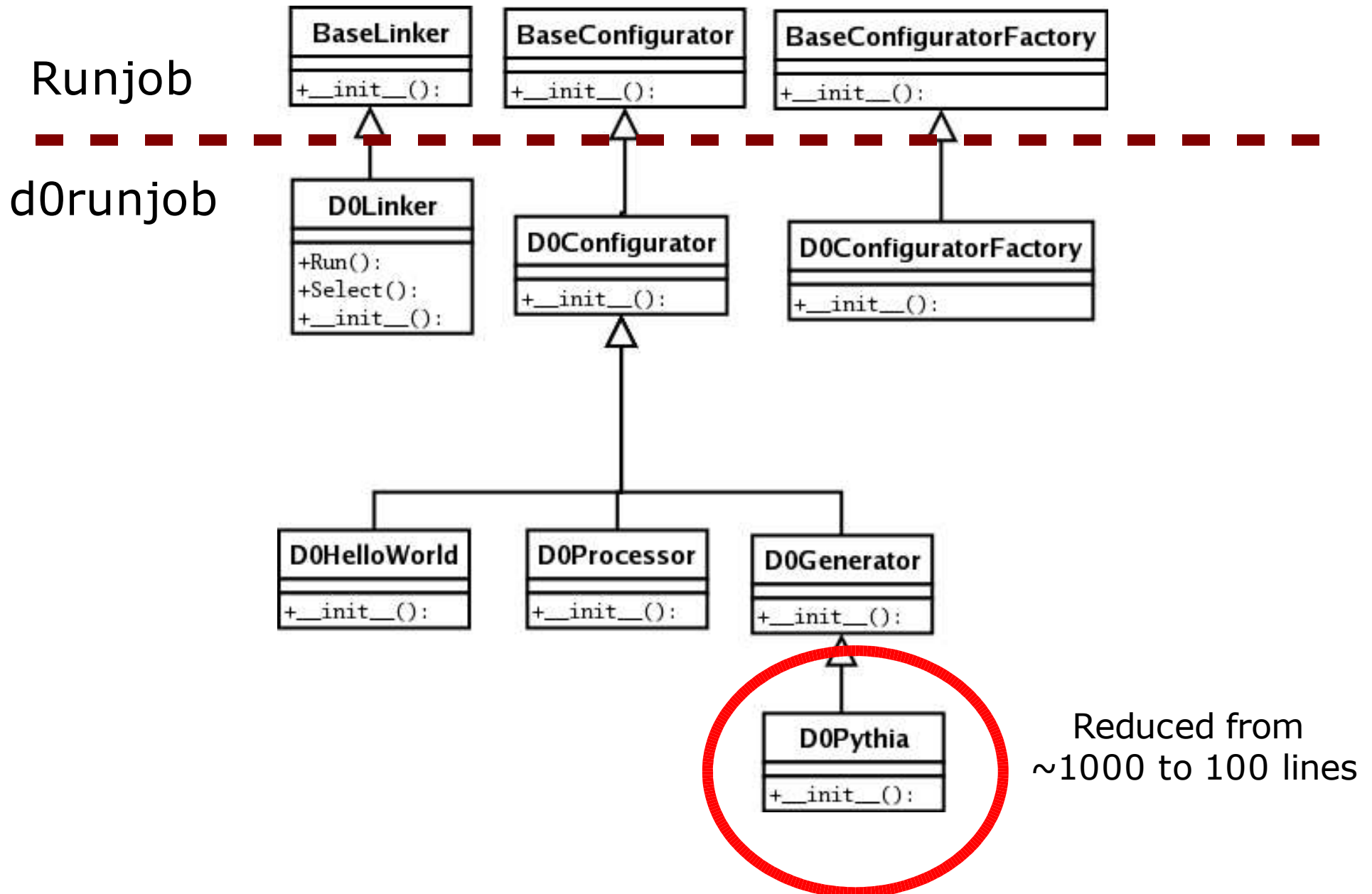
JIM submit/monitoring not used, priority to get remote reprocessing functional

# dØrunjob v07 – the re-write

- Joint FNAL-CD/CDF/CMS/DØ project
- Base classes from Runjob package



# Example generator inheritance



# D-Zero Status

- Remaining deliverables
  - 3.6.9 SAMGrid operating for expert users...done
  - March 1<sup>st</sup> deadline global production
  - 3.6.11 Robust system for all users Q4,2004
- Expert users will be running:
  - MC Production (in shifts)
  - Reprocessing
- General user wait for d0runjob v07