



# Machine/Job Features Update

Stefan Roiser

GDB

9 Apr 2014



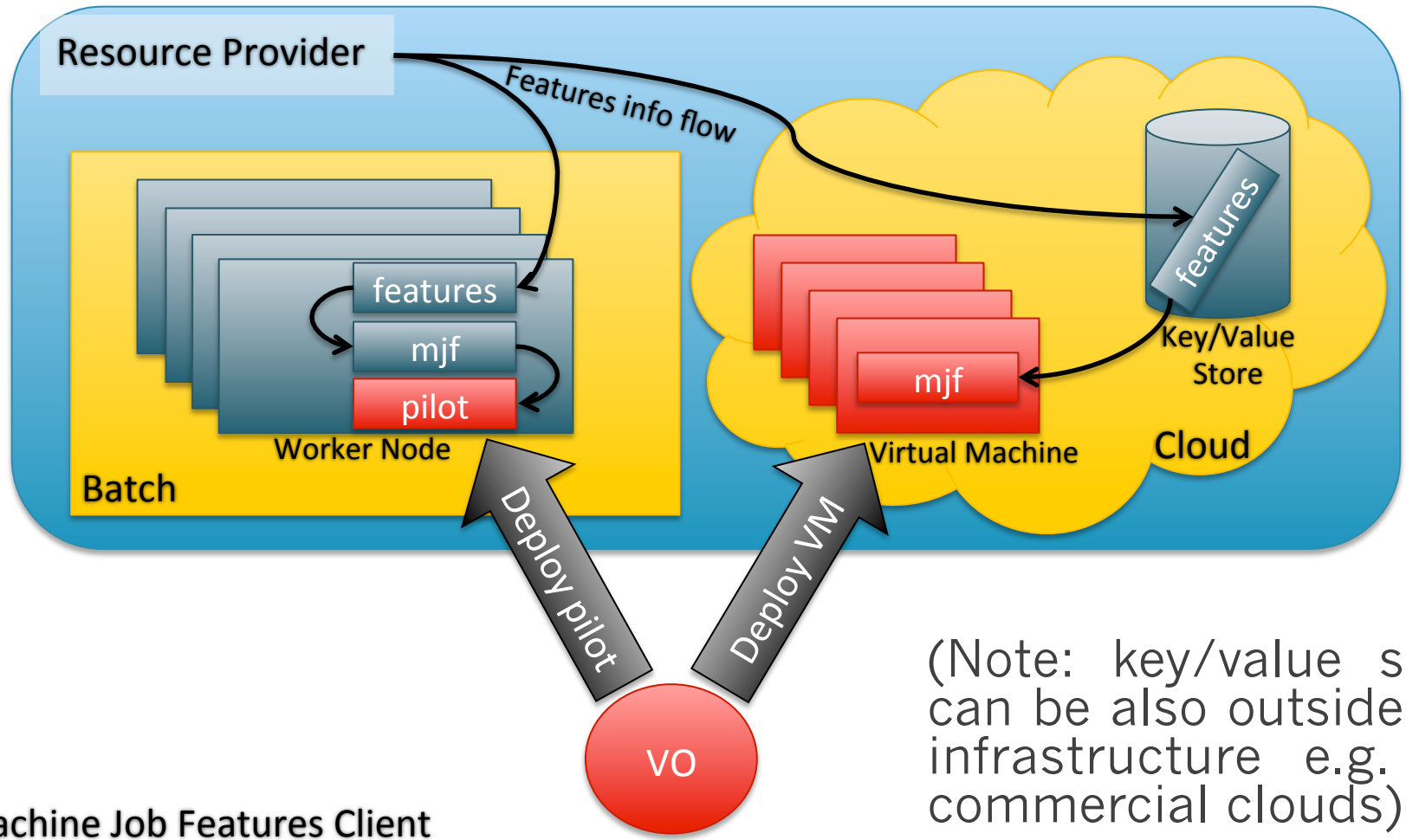
# Content

- Overview
- Status of main constituents
- What the TF does not cover

# Machine/Job Features Taskforce

- Mechanism
  - Possibility to communicate information from the resource provider to the VO about the compute infrastructure per worker node / VM
    - And vice-versa
  - Eg. # cores, shutdown time, wall/cpu limits, ...
- Major constituents of the task force
  - Handling of batch systems
  - Handling of cloud infrastructures
  - Development/deployment of a thin client to consume the information
  - Bi-directional communication (VO -> resource)

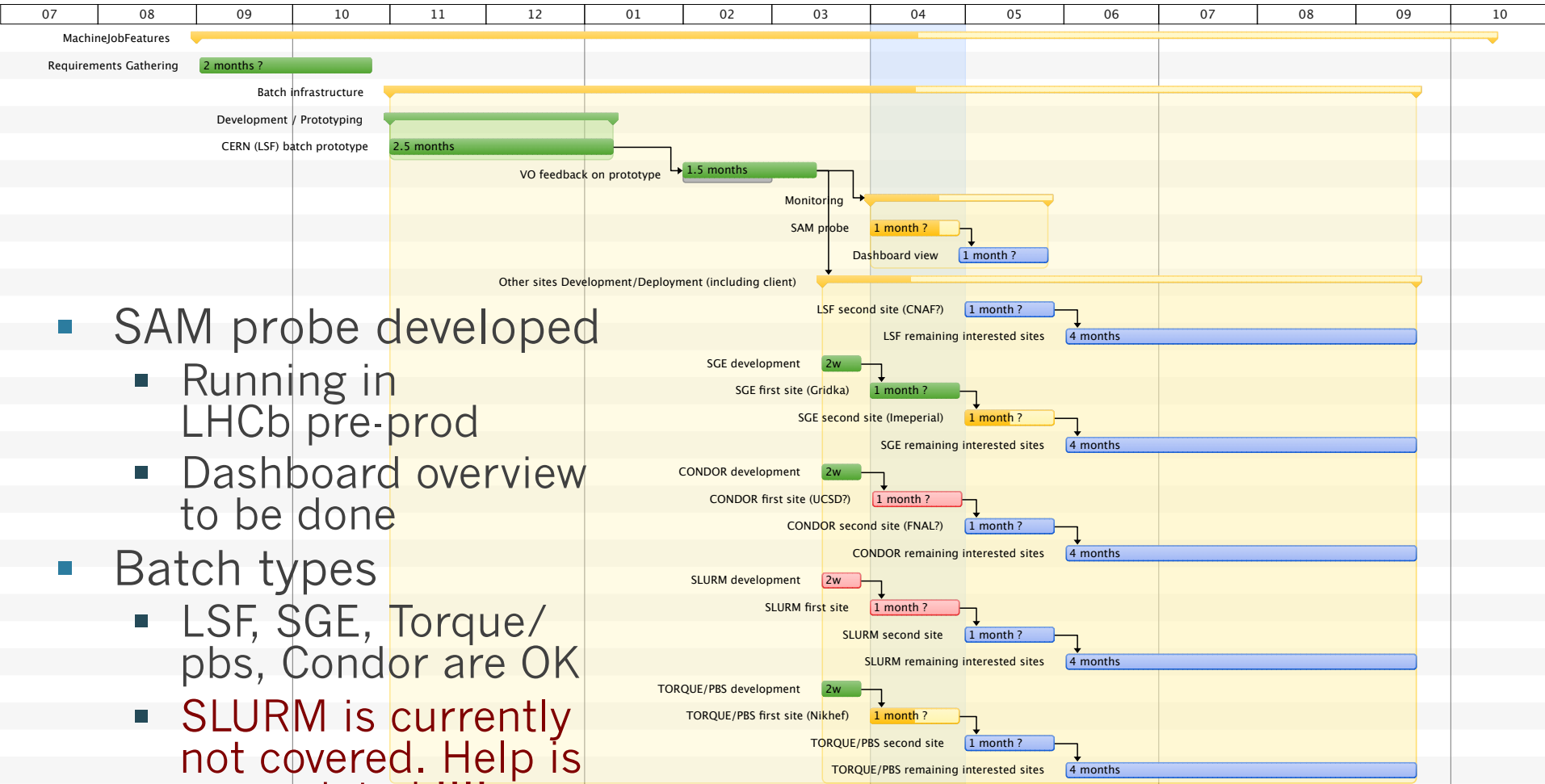
# Machine/Job Features



(Note: key/value store can be also outside the infrastructure e.g. for commercial clouds)

mjf: Machine Job Features Client

# Status and Plans for Batch



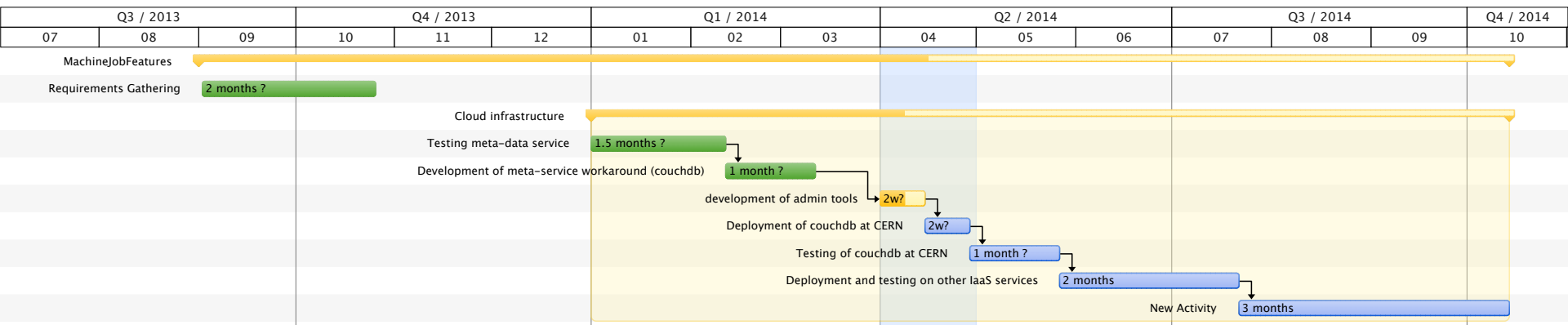
- SAM probe developed
  - Running in LHCb pre-prod
  - Dashboard overview to be done
- Batch types
  - LSF, SGE, Torque/pbs, Condor are OK
  - **SLURM is currently not covered. Help is appreciated !!!!**

## Batch (ctd.)

- SAM probe tests for existence of features / mjf client
  - + course sanity check of values (e.g.  $1 \leq \#cores \leq 128$ )
  - Will help in deployment phase
- Many thanks to all sites involved (so far)
  - Cern, Gridka, Nikhef, Imperial College,
  - More “early adopters” welcome to join in

# Status and Plans for Cloud

- First idea was to use “meta-service” for feature deployment
  - Concerns about bringing down the service with high load
- Development of an alternative key/value store to hold features
  - Store was tested, deployment for CERN/Openstack pending
  - Then move on to different cloud infrastructures / providers

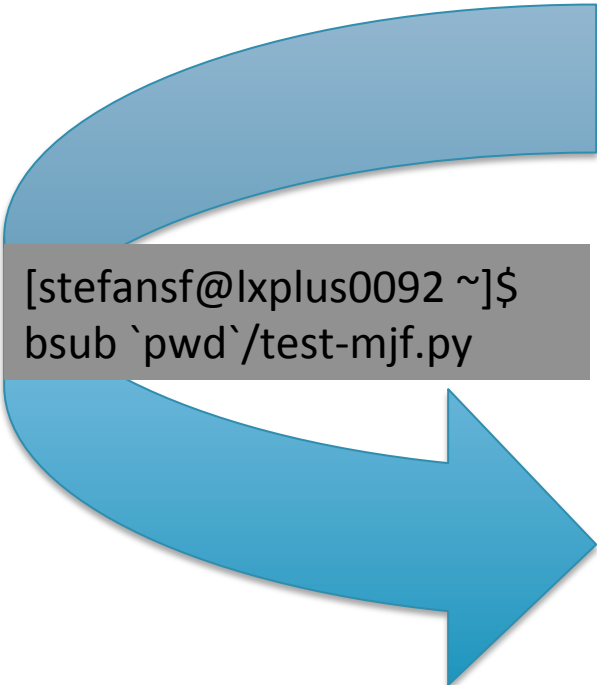


# Status and Plans for Client (mjf)

- Thin python module to abstract from the different ways of feature provisioning
  - Import as python module or run standalone
  - Returns json data structure with features info
- Client developed and deployed via necessary channels
  - WLCG repository (package python-mjf)  
<http://linuxsoft.cern.ch/wlcg>
  - LCG/AA area:  
[/afs/cern.ch/sw/lcg/external/Grid/mjf/](http://afs.cern.ch/sw/lcg/external/Grid/mjf/)
- Current version 0.0.2 ;-)
  - Updates provided through the channels above



# mjf.py example



```
[stefansf@lxplus0092 ~]$  
bsub `pwd`/test-mjf.py
```

```
#!/usr/bin/env python
```

```
test-mjf.py
```

```
from mjf import mjf,MJFException
```

```
m = mjf()
```

```
try :
```

```
    m.collect()
```

```
    print m.features()
```

```
except MJFException, e: print e
```

```
[...]
```

```
LSF/STDOUT
```

```
{'machinefeatures': {'hs06':  
76.1050000000000004, 'jobslots': 6,  
'log_cores': 4, 'phys_cores': 4},  
'jobfeatures': {'disk_limit_GB': 0,  
'wall_limit_secs': 885.51343538532296,  
'cpufactor_lrms': 2.4392628205128202,  
'cpu_limit_secs': 196.78076341896099,  
'cpu_limit_secs_lrms': 480,  
'allocated_CPU': 1, 'mem_limit_MB':  
4000000, 'wall_limit_secs_lrms': 2160,  
'jobstart_secs': 1390213188}}
```

```
[...]
```

# Status and Plans for Bi-directional Communication

- Currently discussing last details of communication
  - Summer: plan to equip mjf client with possibility to provide VO info to resource (invert data flow arrows from overview picture)

# Reminder: What this TF does NOT cover

- Further content of the machine/job features
  - The TF starts with the proposed values of the twiki
    - other changes need to be discussed in WLCG Ops / GDB
- Evaluation of the correctness of the values
  - SAM probes will provide a “course” validation (within bounds) but a priori not try to validate the e.g. power up to a certain percentage level
- Conventions / agreements
  - The TF can make a proposal about e.g. the grace period for shutdown but this needs to be discussed on WLCG Ops / GDB level

# Conclusion

- Client (mjf) is ready for use
- Batch in general in good shape
  - ... help for SLURM much appreciated !!!
- Development for Cloud and Bi-directional communication ongoing

# Batch systems as reported by the BDII (GLUE2)

=== condor (7) === BelGrid-UCL INFN-MILANO-ATLASC RAL-LCG2 UKI-LT2-IC-HEP UKI-SOUTHGRID-BRIS-HEP UKI-SOUTHGRID-OX-HEP === KVM (4) === 100IT INFN-CATANIA-NEBULA INFN-CATANIA-STACK PRISMA-INFN-BARI === loadleveler (1) === NDGF-T1 === lsf (19) === CERN-PROD CRS4 GRISU-COMETA-INFN-CT INFN-BOLOGNA-T3 INFN-CAGLIARI INFN-CATANIA INFN-CNAF-LHCB INFN-GENOVA INFN-LECCE INFN-LNL-2 INFN-PADOVA INFN-PISA INFN-ROMA1 INFN-ROMA1-CMS INFN-T1 INFN-TRIESTE JP-KEK-CRC-02 mainzgrid TRIGRID-INFN-CATANIA === pbs (75) === AEGIS01-IPB-SCL AEGIS02-RCUB AEGIS03-ELEF-LEDA AEGIS04-KG AEGIS09-FTN-KM AEGIS11-MISANU ARAGRID-CIENCIAS Australia-ATLAS BA-01-ETFBL BCBR BG01-IPP BG03-NGCC BG08-MADARA BIFI BMRZ-FRANKFURT BRGM-ORLEANS CIEMAT-TIC CIRMMIP CNR-ILC-PISA EFDA-JET FBF-Brescia-IT GARR-01-DIR GE-01-GRENA GILDA-INFN-CATANIA GILDA-SIRIUS GR-06-IASA GRISU-SPACI-NAPOLI HG-02-IASA HG-04-CTI-CEID HG-08-Okeanos ICEAGE-CATANIA IEPSAS-Kosice IGI-BOLOGNA IN2P3-SUBATECH INFN-Bari INFN-BOLOGNA INFN-CNAF INFN-COSENZA INFN-FERRARA INFN-NAPOLI-ATLAS INFN-PAVIA INFN-PERUGIA INFN-ROMA2 INFN-ROMA3 IR-IPM-HEP JP-HIROSHIMA-WLCG LSG-AMC LSG-AMS LSG-EMC LSG-KUN LSG-LUMC LSG-NKI LSG-RUG LSG-TUD LSG-VU LSG-WUR MA-01-CNRST MREN-01-CIS NIKHEF-ELPROD prague\_cesnet lcg2 PSNC RU-Protvino-IHEP SARA-MATRIX SNS-PISA TW-EMI-PPS UKI-LT2-RHUL UKI-NORTHGRID-LIV-HEP UKI-SCOTGRID-GLASGOW UKI-SOUTHGRID-BRIS-HEP UKI-SOUTHGRID-OX-HEP UPV-GRYCAP USC-LCG2 WCSS64 WCSS-PPS ZA-WITS-CORE === sge (15) === BMEGrid BUDAPEST CESGA CETA-GRID IFCA-LCG2 IFISC-GRID IN2P3-CC IN2P3-CC-T2 LIP-Coimbra LIP-Lisbon NCG-INGRID-PT UKI-LT2-IC-HEP UKI-LT2-QMUL UKI-NORTHGRID-LANCS-HEP UKI-SCOTGRID-ECDF === slurm (13) === ARNES CSCS-LCG2 ICM IISAS-Bratislava SiGNET csc.fi CSCS-LCG2 EENet LRZ-LMU NDGF-T1 NO-NORGRID-T2 UA\_ICYB\_ARC === sungridengine (6) === LRZ-LMU UA\_ICMP\_ARC UKI-LT2-IC-HEP UNIBE-LHEP wuppertalprod === torque (155) === AM-04-YERPHI ATLAND AUVERGRID BG05-SUGrid BG06-GPHI BIOCAMP CA-ALBERTA-WESTGRID-T2 CAFPE-GRANADA CA-MCGILL-CLUMEQ-T2 CAMK CA-SCINET-T2 CA-VICTORIA-WESTGRID-T2 CBPF CIEMAT-LCG2 CREATIS-INSALYON CY-01-KIMON CYFRONET-LCG2 DESY-HH DESY-ZN EELA-UTFSM FMPHI-UNIBA GoeGrid GR-01-AUTH GR-04-FORTH-ICS GR-07-UOI-HEPLAB GR-09-UoA GR-10-UOI GR-11-UPATRAS GRISU-UNINA HEPHY-UIBK Hephy-Vienna HG-03-AUTH HG-05-FORTH HK-HKU-CC-01 ICM ICN-UNAM IEETA IEPSAS-Kosice ifae IFIC-LCG2 IFJ-PAN-BG IISAS-Bratislava IL\_COMP IL-TAU-HEP IN2P3-LPC INAF-TS IN-DAE-VECC-02 INDIACMS-TIFR INFN-NAPOLI-ARGO INFN-NAPOLI-ATLAS INFN-NAPOLI-CMS INFN-NAPOLI-PAMELA INFN-TORINO INSU01-PARIS ITP ITWM JINR-LCG2 JINR-T1 Kharkov-KIPT-LCG2 KR-KISTI-GCRT-01 KR-KISTI-GSDC-01 KR-KNU-T3 KR-UOS-SSCC LCG KNU MD-02-IMI MD-03-USM MD-04-RENAM MK-01-UKIM\_IL MK-02-ETF MK-03-FINKI MY-UPM-BIRUNI-01 MY-USM-GCL MY-UTM-GRID NCBJ-CIS NCP-LCG2 NIHAM pic PK-CIIT prague\_lcg2 PSNC RECAS-NAPOLI RO-02-NIPNE RO-07-NIPNE RO-09-UTCN RO-11-NIPNE RO-13-ISS RO-14-ITIM RO-15-NIPNE RO-16-UAIC RRC-KI RRC-KI-T1 RUG-CIT ru-Moscow-FIAN-LCG2 ru-Moscow-SINP-LCG2 ru-PNPI RU-SPbSU Ru-Troitsk-INR-LCG2 SAMPA SCAI SFU-LCG2 SZTAKI T2-TH-CUNSTDA T3\_HU Debrecen Taiwan-LCG2 TASK TECHNION-HEP TH-HAI TH-NECTEC-LSR TOKYO-LCG2 TR-01-ULAKBIM TR-03-METU TR-10-ULAKBIM TRIUMF-LCG2 TUDresden-ZIH TU-Kosice TW-eScience TW-FTT TW-NCUHEP TW-NTU-HEP UA-BITP UA-ISMA UA-KNU UB-LCG2 UKI-LT2-Brunel UKI-LT2-UCL-HEP UKI-NORTHGRID-LANCS-HEP UKI-NORTHGRID-MAN-HEP UKI-NORTHGRID-SHEF-HEP UKI-SCOTGRID-DURHAM UKI-SOUTHGRID-BHAM-HEP UKI-SOUTHGRID-CAM-HEP UMB-BB UNIANDES UNICAN UNI-DORTMUND UNI-FREIBURG UNINA-EGEE UNINA-EMI-HPC UNI-PERUGIA UNI-SIEGEN-HEP UPorto NDGF-T1 UA\_BITP\_ARC UA\_ILTPE\_ARC UA-IMBG UA-IRE UA-KNU UA-MHI UA-NSCMBR UA-PIMEE UPJS-Kosice WEIZMANN-LCG2 WUT ZA-CHPC ZA-UJ === unicore/x (1) === FZJ

## Links / Further Info

- Batch system implementations by: **Ulrich Schwickerath** (LSF), **Manfred Alef** (SGE), **Igor Sfiligoi** (Condor), **Jan Just Keijser** (Torque/pbs)
- Twiki: <https://twiki.cern.ch/twiki/bin/view/LCG/MachineJobFeatures>
- TF Meetings: <http://cern.ch/go/gTk8>
- RPM repository: <http://linuxsoft.cern.ch/wlcg>
- AFS: </afs/cern.ch/sw/lcg/external/Grid/mjf/>
- Source repository: <http://cern.ch/go/M6xw>