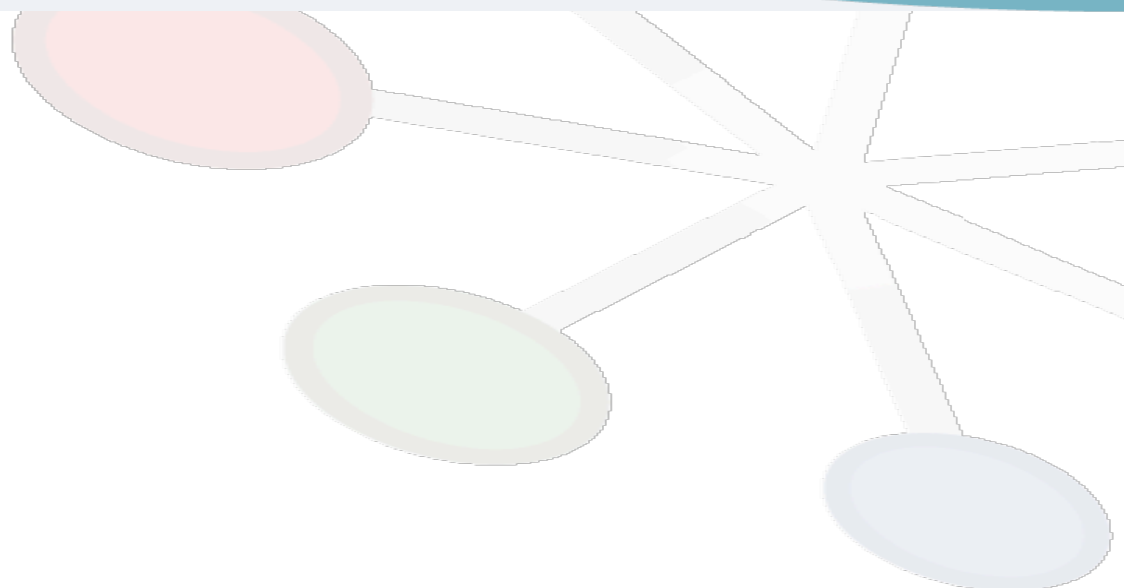




LHCb Computing Model Update





- The LHCb Computing Model was defined in 2005 in the Computing TDR
 - Define Grid access tools: DIRAC and ganga
 - Defines Grid sites role:
 - ☆ Tier0 (CERN): online data collection
 - ☆ Tier1s + CERN-CAF: reconstruction, stripping, analysis
 - ☆ Tier2s and others: simulation
- Some points remained open or not well defined:
 - Grid Usage policy: how should LHCb users access the Computing Grid?
 - Besides Tier1s, which sites can be eligible for running analysis?
- Already some presentations at LHCb week in September 2008!





- **Baseline model:**

- **MC simulation at non-Tier1 sites**

- ☆ **Use all opportunistic sites (Tier2 + others)**

- **Real data processing and analysis at Tier1s**

- ☆ **Limited number of sites**

- ☆ **Datasets available at several sites**

- * All for real data

- * 3 sites for MC

- **Analysis**

- ☆ **Batch processing on the Grid**

- * Submission to DIRAC from ganga

- * Output data uploaded to Tier1 storage

- MicroDST or Ntuples: small size for few events

- Dedicated USER area at all sites (permanently on disk)

- ☆ **End user analysis**

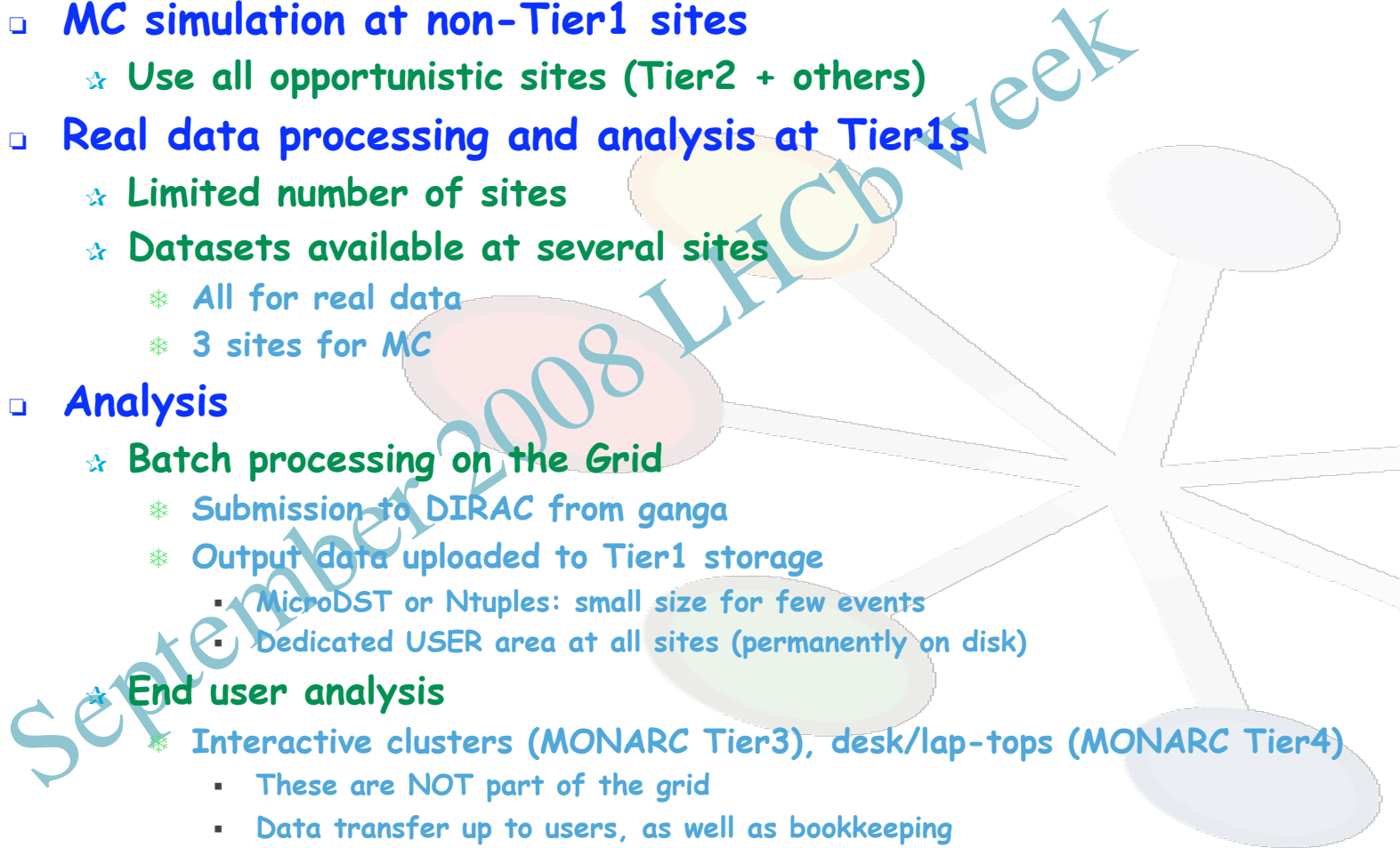
- * Interactive clusters (MONARC Tier3), desk/lap-tops (MONARC Tier4)

- These are NOT part of the grid

- Data transfer up to users, as well as bookkeeping

- * Amount of data usually small

- Typically < 1 TB, can fit on a cheap local disk





- What we (Computing project) can and will support:
 - Software installation / environment tools
 - ☆ Install_project, extCMT, SetupProject
 - Data management tools
 - ☆ Copy data from the Grid to local storage
- What we cannot support:
 - Handle group/private datasets
 - Sysadmin your clusters
 - Install and certify software on your installation
 - Requests for National Grid computing
 - ☆ Grid resources reserved for nationals
 - ☆ Two orthogonal approaches
 - ☆ Grid == Shared resources by definition
- What can be discussed:
 - Local Grid shared resources (storage + CPUs)
 - ☆ If these are **additional** resources (we are short in MC resources at Tier2s)
 - ☆ If support is provided by sites for data management and data access problems
 - ☆ If access is granted to the whole LHCb collaboration



- Production activities
 - ▣ Simulation, reconstruction, stripping, WG analysis (μ DST)
 - ▣ Uses DIRAC and the LHCb Production System
- User analysis
 - ▣ Data analysis
 - ☆ For testing, use local resources (including local batch system)
 - ☆ For large datasets, use Grid Computing
 - ▣ Toy MC
 - ☆ Use Grid Computing for large samples
 - ▣ Non-Grid user analysis
 - ☆ On local clusters (Tier3), desk/lap-top (Tier4)
 - ▣ Policy for Grid Analysis
 - ☆ Supported access: exclusively via ganga + DIRAC
 - ❄ No direct Grid job submission supported
 - ☆ Data Analysis performed on LHCb Analysis Centres (LACs)
- Grid sites
 - ▣ Grid sites provide pledged resources to the whole Collaboration without restrictions



LHCb Analysis Centres

- All Tier1s and CERN-CAF are LHCb Analysis Centres
- A Tier2 may request to become a LAC if:
 - It provides sufficient storage in a Grid-SE, supported by LHCb Core Software (Castor, StoRM, dCache, DPM...)
 - It provides resources that are in addition to resources provided for simulation
 - It provides enough local LHCb manpower for managing the datasets, in coordination with the LHCb Data Manager.
 - It is open to the whole collaboration as specified before
 - Conclusion of an agreement between LHCb and the site
 - ☆ Full commitment by the site (MoU-like)



Non-Grid Analysis support

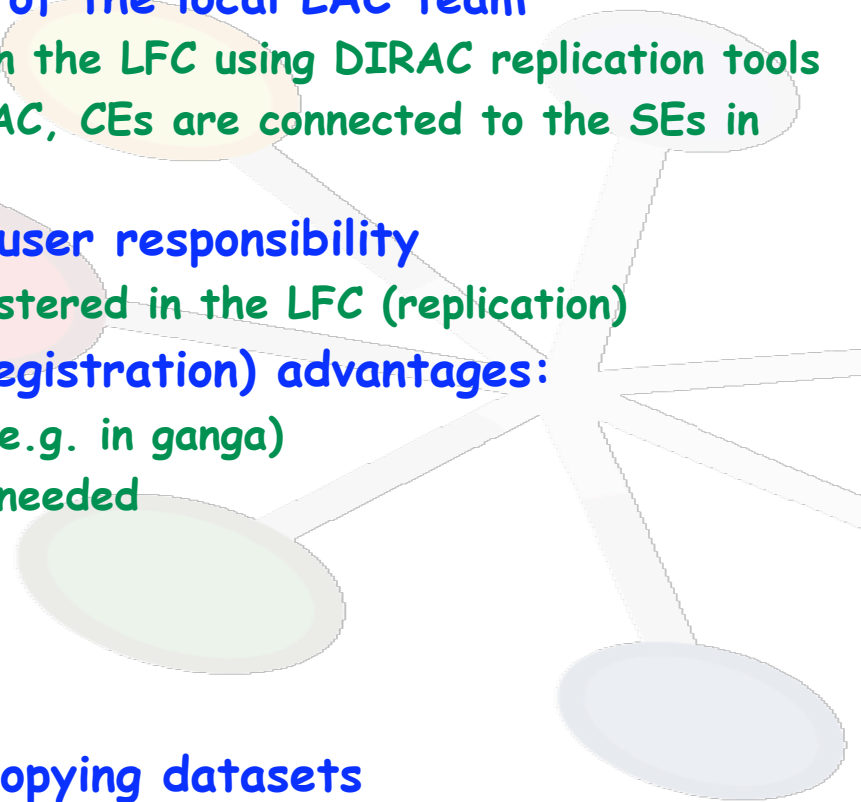
- Tier3/4 may or may not have direct access to a Grid-SE
 - None of these resources are pledged to not accounted for the LHCb Collaboration in the WLCG
- The Computing projects provides support for:
 - Data Management: with DIRAC tools (see next slide)
 - Software distribution and installation
 - ☆ For applications, Gaudi, ganga and DIRAC clients
 - ☆ install_project





DIRAC Data Management tools

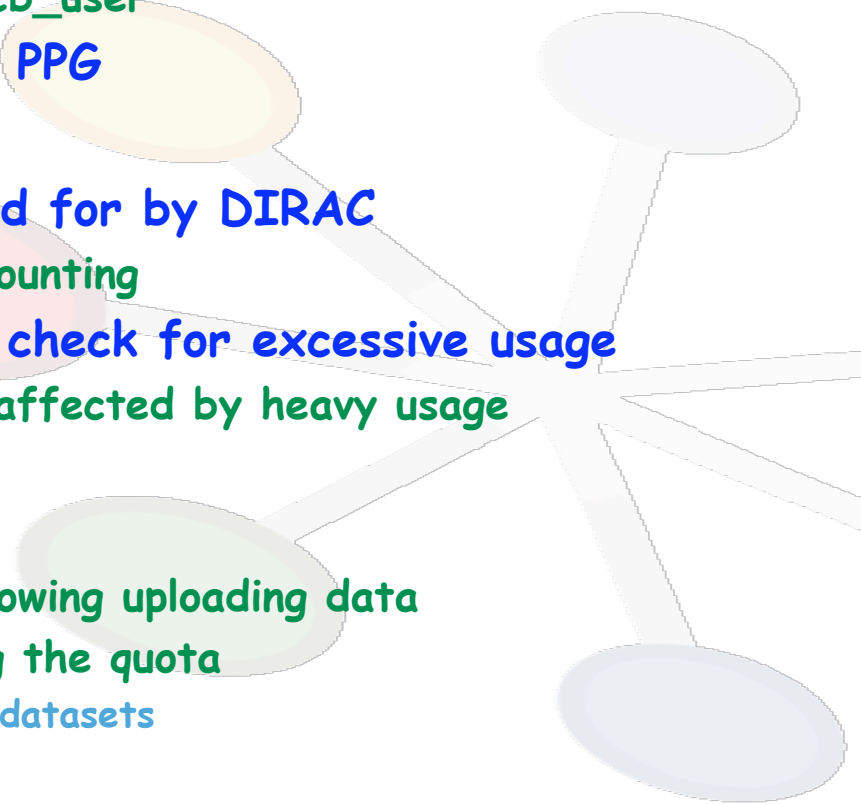
- **Grid-SE**
 - Tier1: data management is under control of the LHCb Data Manager
 - Other LACs: under control of the local LAC team
 - ☆ Datasets fully registered in the LFC using DIRAC replication tools
 - ☆ SEs are registered in DIRAC, CEs are connected to the SEs in DIRAC
 - Tier3/4: under local team/user responsibility
 - ☆ Datasets may be fully registered in the LFC (replication)
 - Full replication (with LFC registration) advantages:
 - ☆ Allows using LFNs in jobs (e.g. in ganga)
 - ☆ No handmade bookkeeping needed
- **Non-Grid SE**
 - Only for Tier3/4
 - Local disk array usually
 - DIRAC provides tools for copying datasets
 - ☆ No LFC registration, purely manual bookkeeping





Resource accounting and priorities

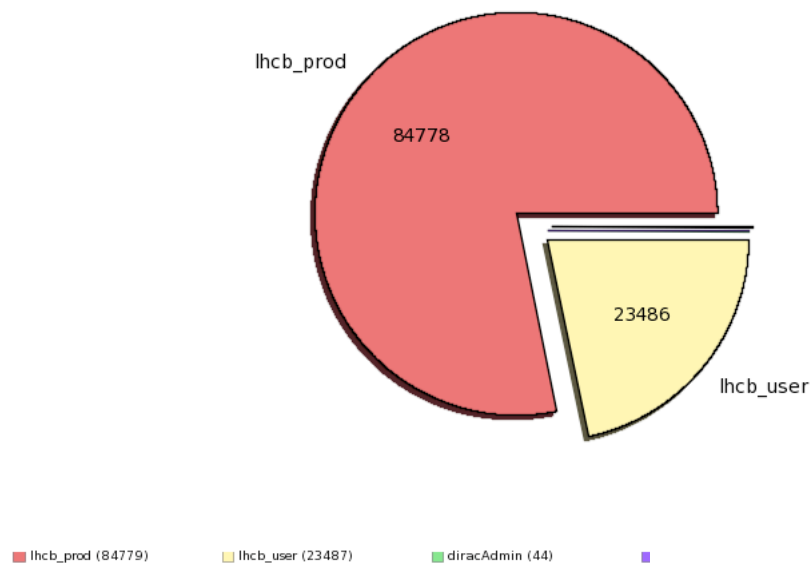
- **Priorities**
 - **Assigned forgroups of users**
 - ☆ Groups to be defined by PPG
 - ☆ Currently a single group lhcb_user
 - **Group priorities defined by PPG**
- **Accounting**
 - **Grid CPU usage is accounted for by DIRAC**
 - ☆ Group accounting, user accounting
 - **No CPU quota, but regular check for excessive usage**
 - ☆ Individual priority may be affected by heavy usage
 - **Storage accounting**
 - ☆ System of quota in place
 - ☆ No "quota exceeded" disallowing uploading data
 - ☆ Mail sent when approaching the quota
 - * Asking for clean up of old datasets



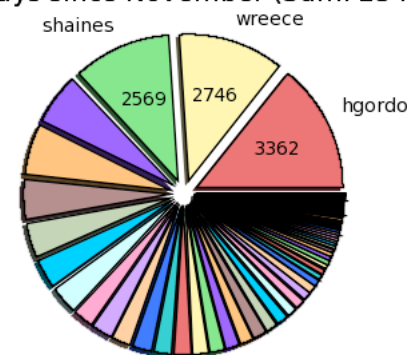


Accounting examples

CPU days since November (Sum: 108351)



CPU days since November (Sum: 23487)



hgordo (3363)	wreece (2746)	shaines (2570)	apuignav (1345)	mcoombes (1326)
bkhanji (1048)	xieyu (908)	conti (773)	jhe (737)	coftzpa (525)
rvazquez (523)	dvaneijk (507)	adinolfi (506)	skottowe (472)	serranoj (455)
pgandini (416)	frodriku (375)	phunt (363)	amartens (344)	riambert (334)
rharji (323)	barsuk (288)	awebber (279)	bonivent (267)	crodrigu (216)
mcharles (192)	emaurice (166)	dsavrina (148)	larrabit (145)	rogers (120)
powell (112)	asarti (100)	fmacheffe (98)	spradlin (81)	jdickens (74)
hvc (74)	pkoppenb (70)	seaso (69)	calvom (68)	ukerzel (57)
fsomro (56)	kimym (50)	nmangiaf (49)	karl (49)	cadrover (46)
snos (46)	mlernv (41)	thears (38)	acarlone (36)	plus 81 more



- Grid usage policy
 - LHCb Grid sites are democratically available to the whole collaboration
 - Production activity uses LHCb Production System + DIRAC
 - User analysis must use ganga + DIRAC
 - ☆ No support for direct submission
- LHCb Analysis Centres
 - Tier1s + CAF
 - Tier2s
 - ☆ Enough storage resource
 - ☆ Enough local LHCb manpower
 - ☆ Additional resources for analysis
 - ☆ Subject to evaluation and conclusion of an agreement
 - * Commitment of the LHCb institute, not individuals
 - * Binding like an MoU
- Two documents submitted to TB for approval, then CB
 - EDMS 1059800 and 1059810

