University of Zurich
CMS Group

Thomas Speer

CHIPP LHC Computing workshop

CSCS Manno

25-26 August 2005

University of Zurich CMS group

- Group is based at CERN
- Present involvement in CMS (software/analysis related)
  - 2 postdocs (Vincenzo Chiochia, TS)
  - 3 students (shared with PSI and CERN, Kirill Prokofiev, Enver Alagoz, Dimitrios Tsirigkas)
- Future involvement will stay of that order/ grow by 1 postdoc/senior scientist, 1 student

Current activities

- Pixel detector hardware
- Track/Vertex reconstruction:
  - software development
  - Physics TDR: performance studies in different topologies

  ($B_s \rightarrow J/\psi\,\varphi$, b-jets, light jets, H $\rightarrow \gamma\gamma$, tt, ttH, etc…)
- Pixel detector software: monitoring, reconstruction
- Analysis:
  - $B_s \rightarrow J/\psi\,\varphi$: Benchmark channel for Physics-TDR
  - Rare B decay ?
  - One additional channel ? (Higgs?)

Computing: current usage

- Facilities at CERN:
  - Code development, evaluation
    - reconstruction software
    - pixel software
  - Analysis: B_s samples on CASTOR

✓ *lxplus* interactive services

✓ *lxbatch* batch services

✘ nearing saturation

✘ not all data samples are at CERN anymore

✓ Desktops (expand to small "local cluster" for analysis?)

Computing: current usage

- LCG:
  - $B_s \rightarrow J/\psi\,\varphi$: sample at CERN or Bologna.
  - Track/vertex reconstruction : samples scattered around different centres

    H $\rightarrow \gamma\gamma$(Lyon), DY (FNAL), *ttH* (Perugia), light jets (CNAF), etc

  - ✓ CRAB: (CMS Remote Analysis Builder) job creation submission tool for analysis of published data with ORCA (CMS reconstruction software), based on requested dataset
    - ✘ not all centers run with the same OS, version of ORCA (being improved, SLC3 now nearly everywhere).
    - ✘ random (in)stability (but, when everything works, it works very well!)
    - ✘ Little/no information in case of problems (e.g. aborted jobs…)

Computing: current usage

- All data samples used until now were produced centrally (e.g. in DC04)
- Only very limited amounts were produced privately for specific studies
- Analysis:
  - ORCA reconstruction job produces Root Trees
  - Analyse these Trees on your desktop

The $B_s \rightarrow J/\psi\,\varphi$ decay

- Current MC samples, located at CERN and CNAF (Bologna):
  - signal - $B_s \rightarrow J/\psi\,\varphi$: 200k events
  - background - $B_s \rightarrow J/\psi X$: 200k events

    - $bb \rightarrow \mu\mu\, X$: 100k events

  All samples with low-lumi PU ($2\cdot10^{33}\,\mathrm{cm^{-2}s^{-1}}$): ~1MB per event
- New samples will be generated in the near future
  - signal - $B_s \rightarrow J/\psi\,\varphi$: 2M events
  - background - $B_s \rightarrow J/\psi X$: 2 M events

    - $bb \rightarrow \mu\mu\, X$: O(100k) events

    - prompt J/ψ: O(100k) events

  $52\cdot10^6$ bb already available: 52 GB

The $B_s \rightarrow J/\psi\,\varphi$ decay

- Data per 10 fb^{-1}: ~ 420k events (preliminary estimates)
  - signal: ~ 170k events
  - background: ~ 250k events
- CMS computing model:
  - RAW event size at startup ~ 1.5 MB/event at L = $2 \cdot 10^{33}\ cm^{-2}s^{-1}$
  - RECO ("reconstructed event format") ~ 250 kB/event
  - AOD (Analysis Object Data) ~ 50kB/event

- Other samples are also needed for the study of this decay ($B^0 \rightarrow J/\psi K_s$)

Computing: short-term

- Start using Manno Tier 2 - first tests successful
- Use of Manno Tier 2 for B_s analysis of new MC production (also for the production itself?)
- Small local cluster for analysis?

Computing: long-term

- Approx. 4/5 users actively doing analysis
- Resources per user (CMS computing model):
  - CPU: 15 kSi2k
  - Storage: 3.5 TB
- Hope that most of the data will be at a Tier 2 (3?), only limited amount on our “local cluster”
- Same for MC production
- Since we are based at CERN, we would also like to use the CERN facilities… to what extent is this possible? Practical?