

# CMS Karlsruhe @NAF resources

Volker Büge

for the Karlsruhe CMS group

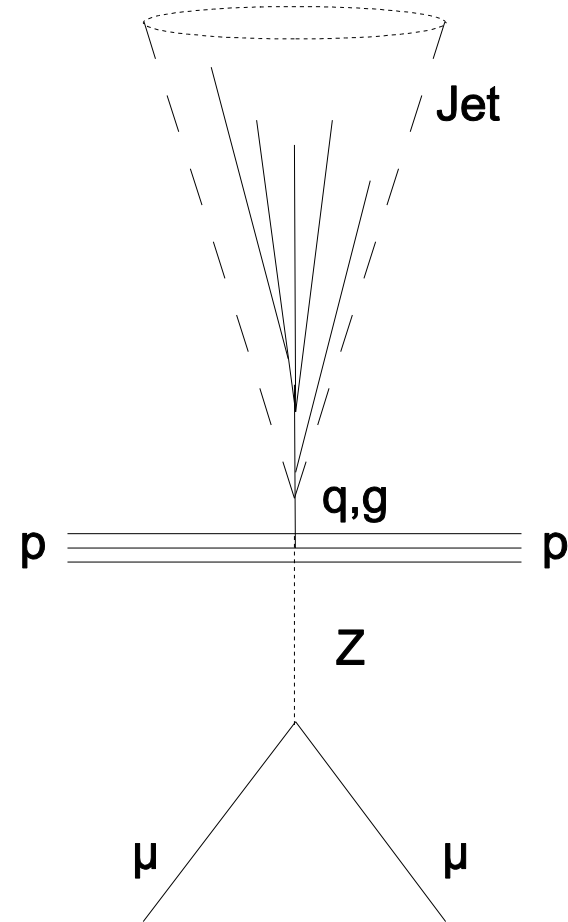
Institut für Experimentelle Kernphysik  
Karlsruhe Institute of Technology

## Z boson production at the LHC

- Large number of Z boson events available
- Can be used for:
  - Commissioning of the **tracker**
  - **Luminosity** measurement
  - Calibration of the **calorimeters** and **jets**
  - Measurement of the **jet resolution** and the **underlying event**

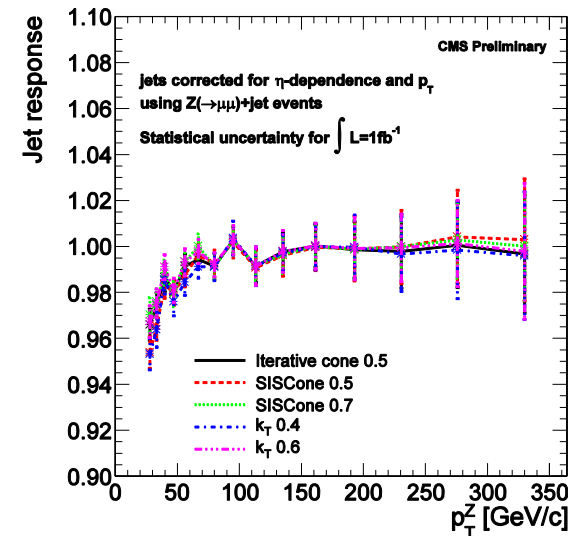
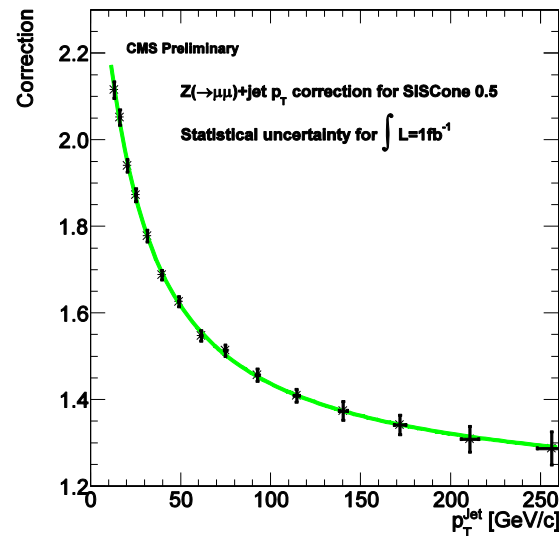
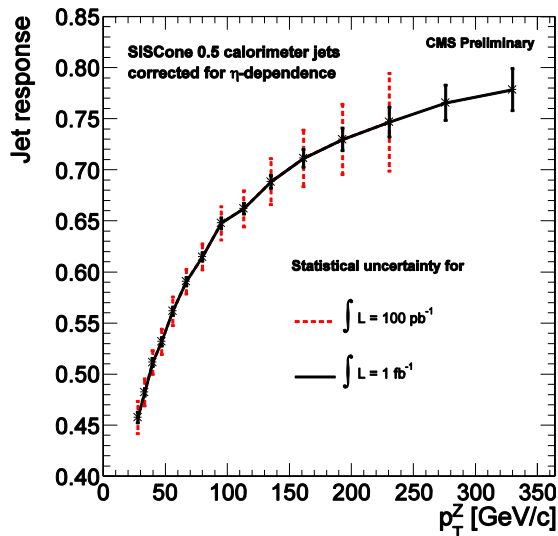
## Jet calibration with **Z+1 jet** events:

- Z boson momentum is **balanced by one jet of the hard sub-process**
- Allows determination of the momentum of the balanced **parton** through comparison with the momentum of the Z boson
- Depends only on the **calibration of the tracker** ( $Z \rightarrow \mu\mu$ ) or **additionally on the calibration of ECAL**
- **Cross section smaller** compared to photon+jet but kinematics of the Z boson can be **reconstructed very precisely**



## Data samples

- Analysis performed on **official CMS JetMET datasets**
- DESY is **associated T2** for JetMET
- All required **datasets** for this analysis **available at DESY**  $\rightarrow$  **interactive access** perfect for debugging and optimisation.
- Recently investigated datasets and event numbers:
  - Summer08 Simulation: 1.5 Mio events
  - Winter09 FastSimulation: 2 Mio events
  - Summer09@10TeV: 1.5 Mio Events
  - Summer09@7TeV: 2 Mio events



## Workflow:

1. **Data access** currently via **CE (grid submission)**. However, the analysis is setup to **also support local submission** via SGE (in case of large latencies of the grid environment). Currently, grid based access is preferred (with proxy [cms:/cms/dcms](https://cms.cern.ch/cms/dcms))
2. Processing of the datasets and **creation of small and adapted skims**
3. The output is stored via the **DESY dCache SE** for later offline processing
4. Current data format allows an efficient **analysis** of all skimmed files **on a laptop**. Therefore, these files are copied from dCache to local resources. This **may change in the future** with growing data samples.

## Comments

- **Interactive** access to our datasets!
- **Local support** in case of problems
- **Reliable service and support** of the DESY/ NAF infrastructure, e.g. for the dCache Storage Element and the WMS!
- For parts of our analysis, we have to store and process **large amount of small files** via the SE  $\rightarrow$  dCache not optimal suited
- Would be a great improvement if **direct grid access to the Lustre file space** is possible.

## Summary of Karlsruhe analysis activities on NAF resources (list not complete):

- Manuel Zeise et. al. : “Modeling of background from Z boson decays for the Higgs boson search in the channel  $H \rightarrow \tau\tau$ ”
- Oliver Oberst et. al. : „Measurement of the underlying event with Z boson events at the LHC“
- Jasmin Gruschke et. al. : „Early top-quark production cross-section measurement in the semi-leptonic channel “
- Jeannine Wagner-Kuhr et. al. : „Top physics at the LHC“
- Fedor Ratnikov et. al. : „SUSY at the LHC“
- Hauke Held et. al. : „Energy Corrections for Semileptonically Decaying b-Jets“
- ...

## Typical workflow of the analyses:

- If data samples are located at T2 DESY: [Skimming via grid job submission](#) or local SGE submission. If not, skims are performed and the output is [transferred to DESY via the dCache SE](#). If necessary and possible, PHEDEX transfer of an example dataset block to DESY for grid job debugging.
- Depending on the analysis, the skimmed files are [further processed via SGE](#)
- In most cases, the [processed files are copied to local resources for the final analysis](#). Only few workflows are finalised on NAF resources. This may change in near future.

## Future plans:

- Some groups plan to move [CPU intensive jobs](#) to NAF resources
- Migration of more [end user analyses](#) to the NAF

## Summary of comments and wishes:

- [Access to the lustre](#) file system from the [grid](#)
- [Larger “/scratch”](#) partition, eventually with user or analysis group [quota](#)
- [Guaranteed disk space for datasets fragments](#) (1 block minimum) and analysis area (some TB per analysis group)
- [Guaranteed disk space per user](#)
- Currently, [mailing list](#) used [to report problems](#). As mailing lists do not allow to inform about the status of the request  
→ [Ticketing system like GGUS](#) would be more transparent to the user.

Users are very happy with the performance, availability and support at the NAF

Thank you!