

Distributed Analysis Using DaVinci In the Glite Framework



- DaVinci the Analysis Program for LHCb

- Rests on the Gaudi Framework
- OO-based, programmed in C++
- Steered through options files
- User specific libraries can be added in a simple way



- To run DaVinci for a generic channel one needs:-

Standard options files

User channel specific algorithms

User options files

- Output primarily comes in three forms:-

Histograms Ntuples DaVinci Standard Output

- Physics Overview

- LHCb will look into evidence of asymmetry in particle behaviour to explain why matter is more prevalent than antimatter

- This is made possible by searching for B mesons

- The asymmetry is known as CP Violation

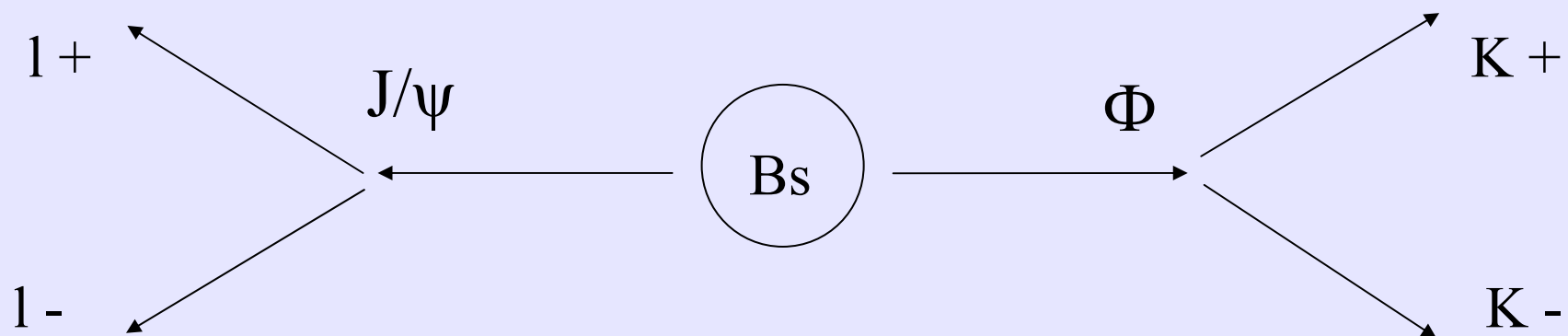
- Non-invariance of a process under charge conjugation and parity transformation

• Bs to J/ψ Φ Channel

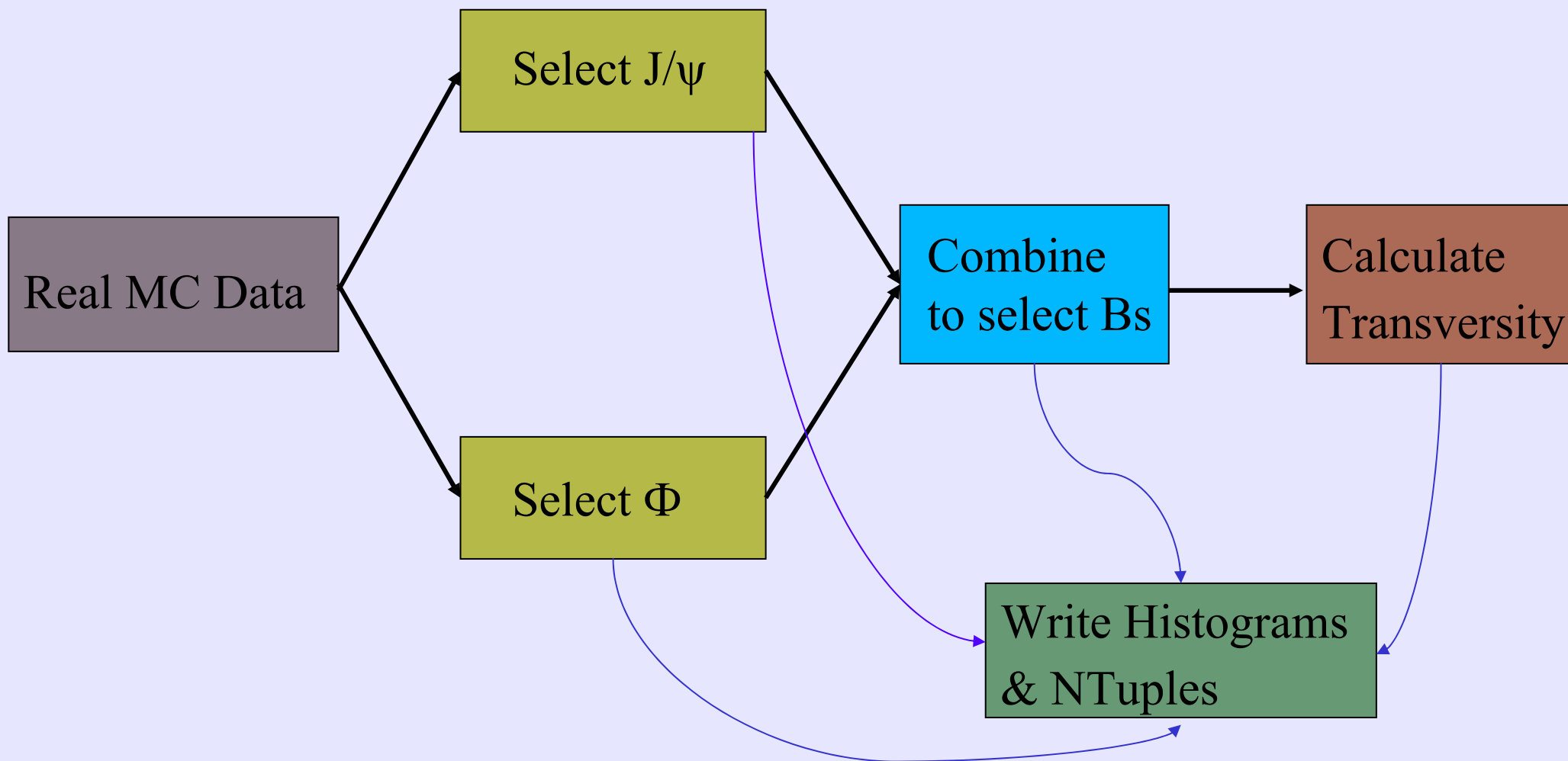
- Chosen as a ‘typical’ analysis
- Annual signal yield of ~ 100k events
- Possible to extract CKM property $\delta\gamma$



• If measured $\delta\gamma$ is large \longrightarrow new physics beyond SM



• DaVinci – Algorithm Overview



- GLite Prototype

- Based on the AliEn file catalogue

- To run a typical job one needs:-

JDL file script to run application user specific files

- Must decide to use one tarball of desired package directly, or the GLite package manager

- The procedure for running a job is:-



- GLite Package Manager

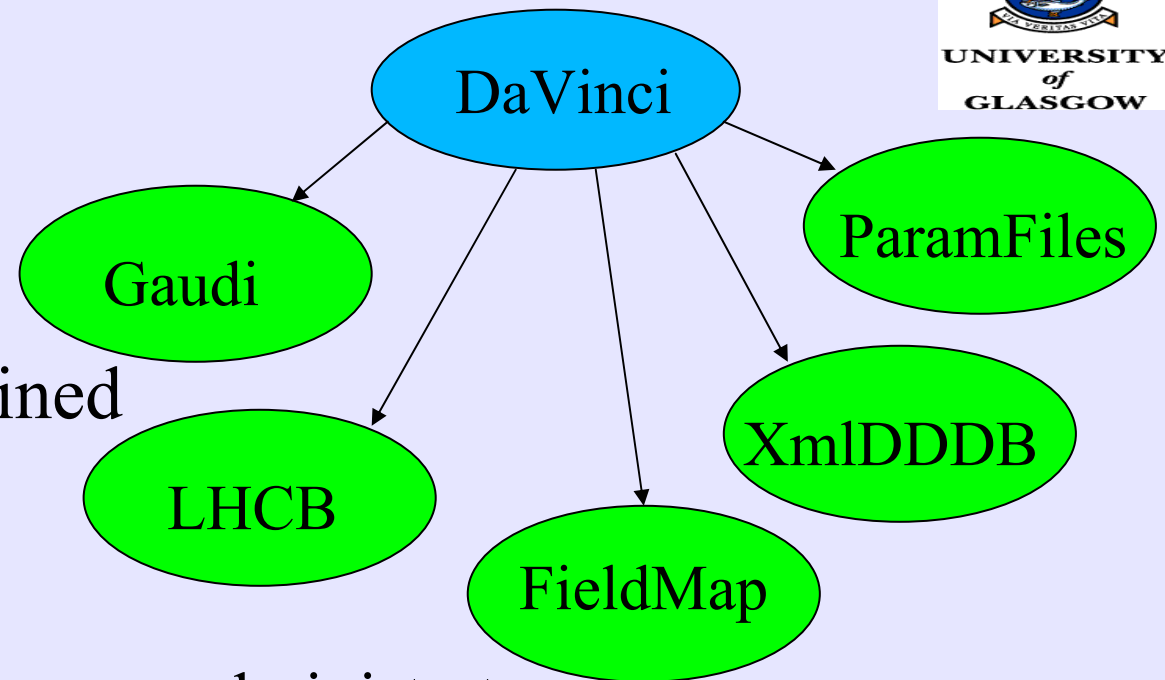
- Works! Is now very streamlined

- LHCb Use Case

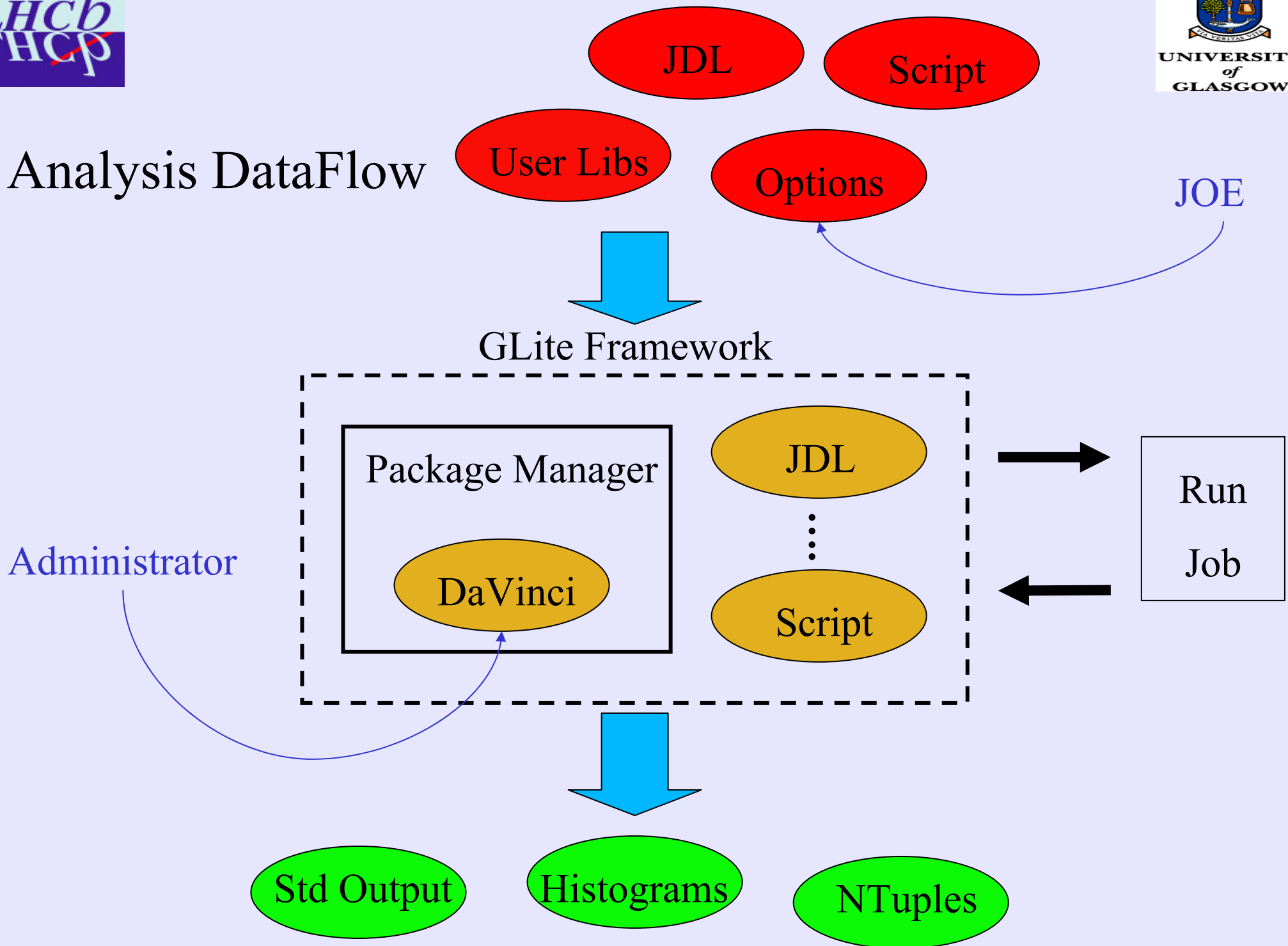
- Would urge each VO to name an administrator
- Create Generic 'lhcb' account
- All versions of DaVinci could be inserted once
- Any LHCb user could then make use of a particular package via

`Packages= 'lhcb@DaVinci::v12r3'`

- Avoids each user inserting the same packages unnecessarily

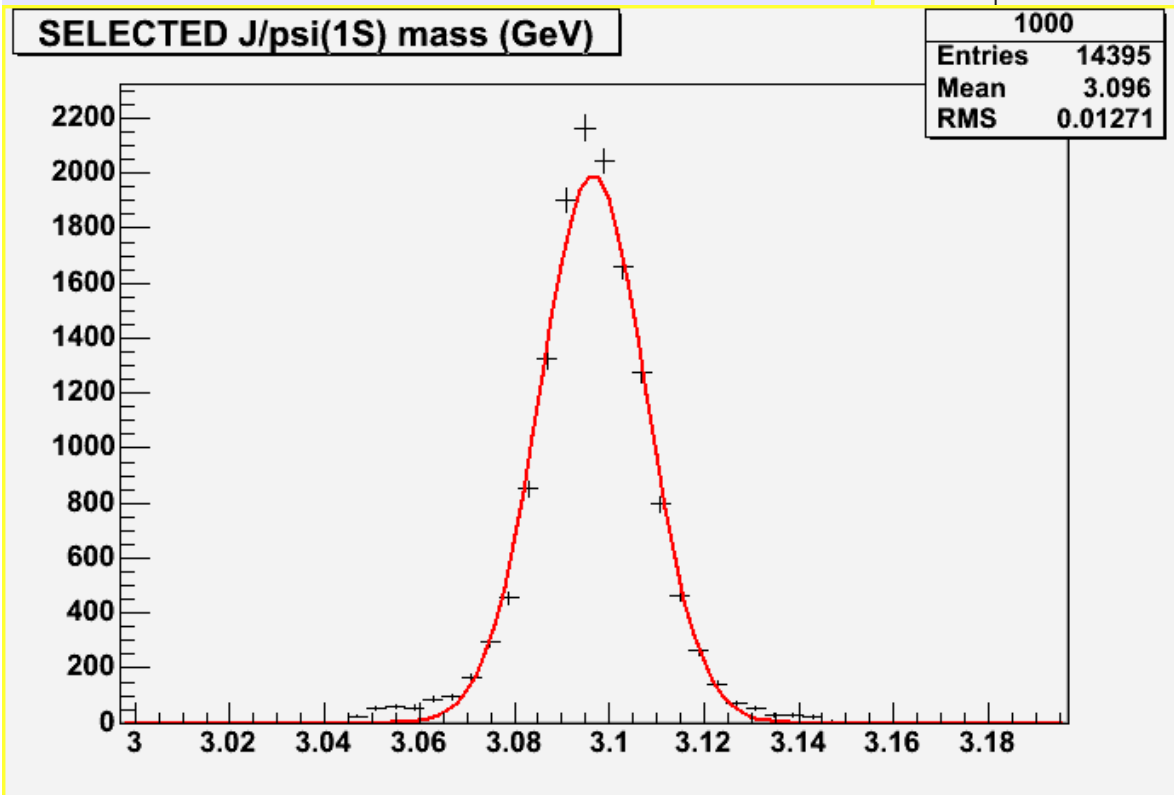
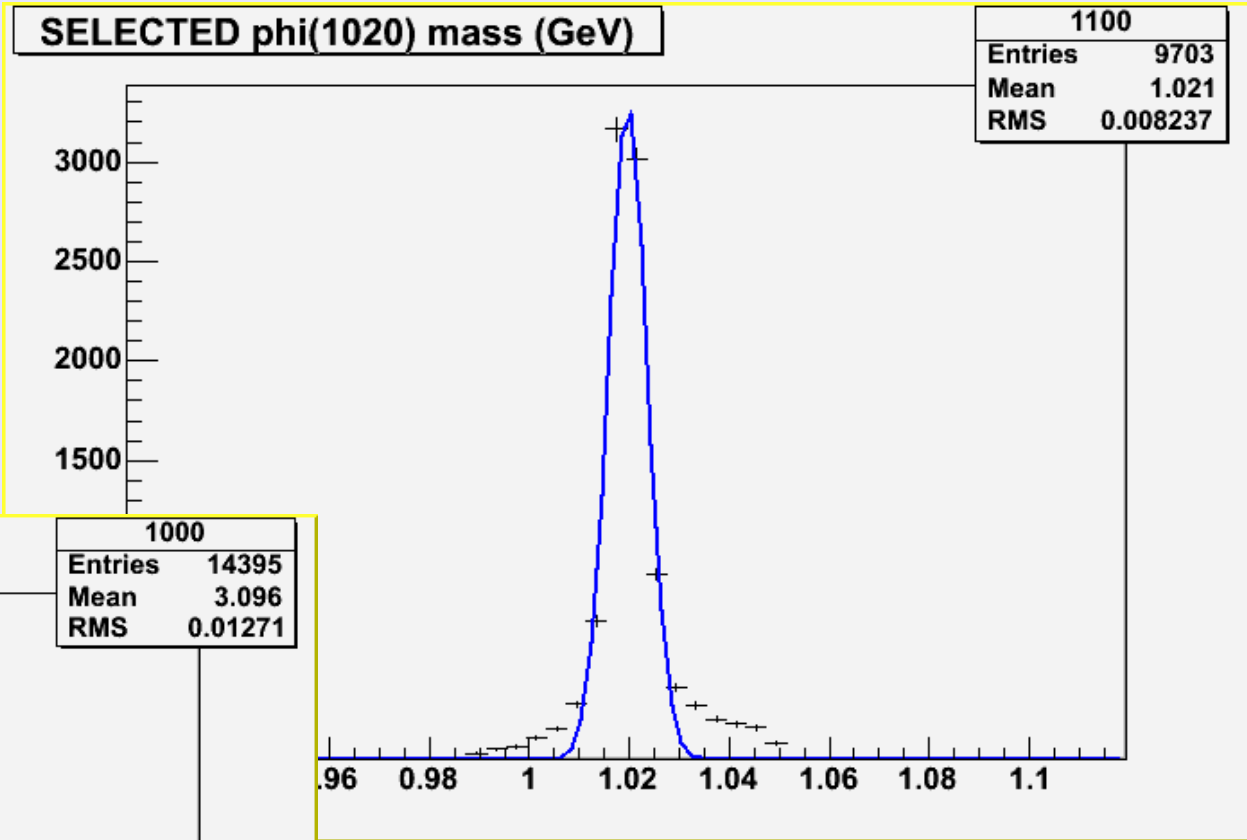


• Analysis DataFlow



• Progress Made

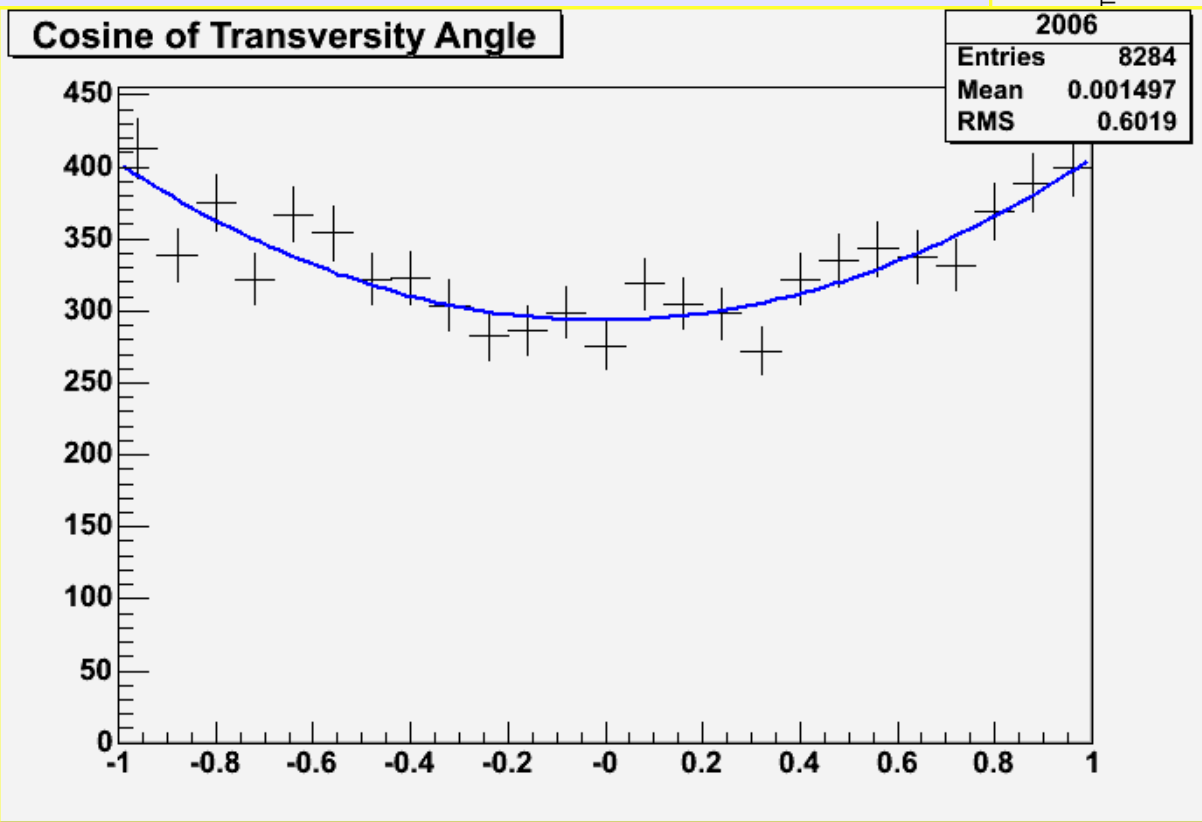
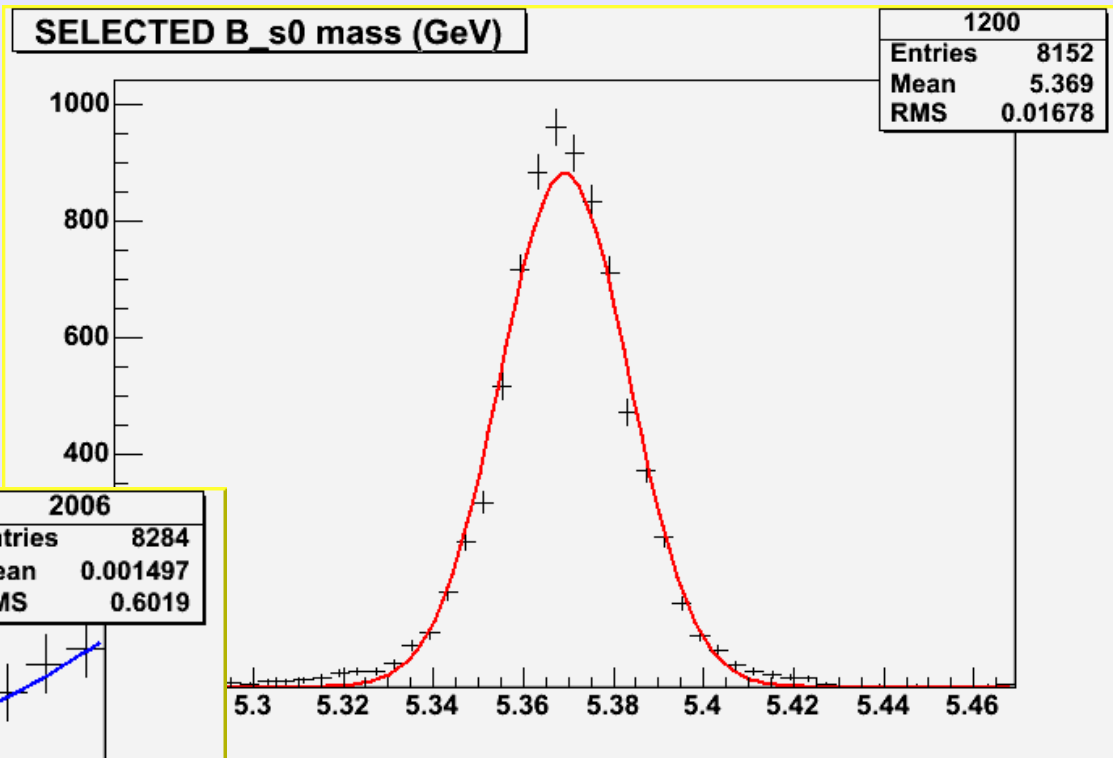
- Ran over 100,000 Events
- J/ψ Efficiency $\sim 14.4\%$
- Φ Efficiency $\sim 9.7\%$



• Typical 1000 event job takes ~ 9 minutes

• Progress Made contd.

- Bs Efficiency $\sim 8.2\%$
- Transversity distribution promising



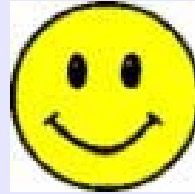
- Contributions from CP odd ($l=1$) & CP even ($l=0,2$) states
- Need to investigate helicity states in generation (EvtGen)
- Event selection is quite clean



• Problems Encountered

- Issues with reliability for individual jobs and commands
- File manipulation is cumbersome
 - Would be better if this could be automated for the user
- Some problems with datasets
 - Originally these were picked up from Castor but not now
 - Further investigation into this is required
- Worker nodes had 20Gb disks, needed to split large jobs
- System prone to hanging, often needs to be rebooted
- Often not clear whether user is at fault or the prototype itself

• Overall Impressions



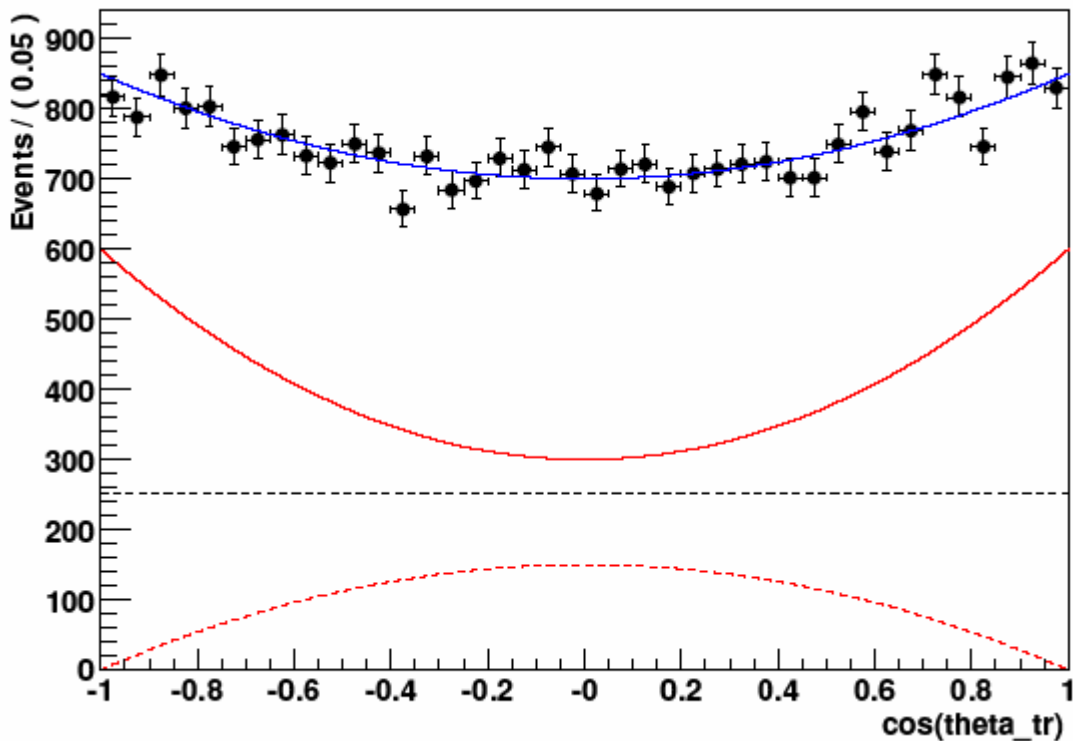
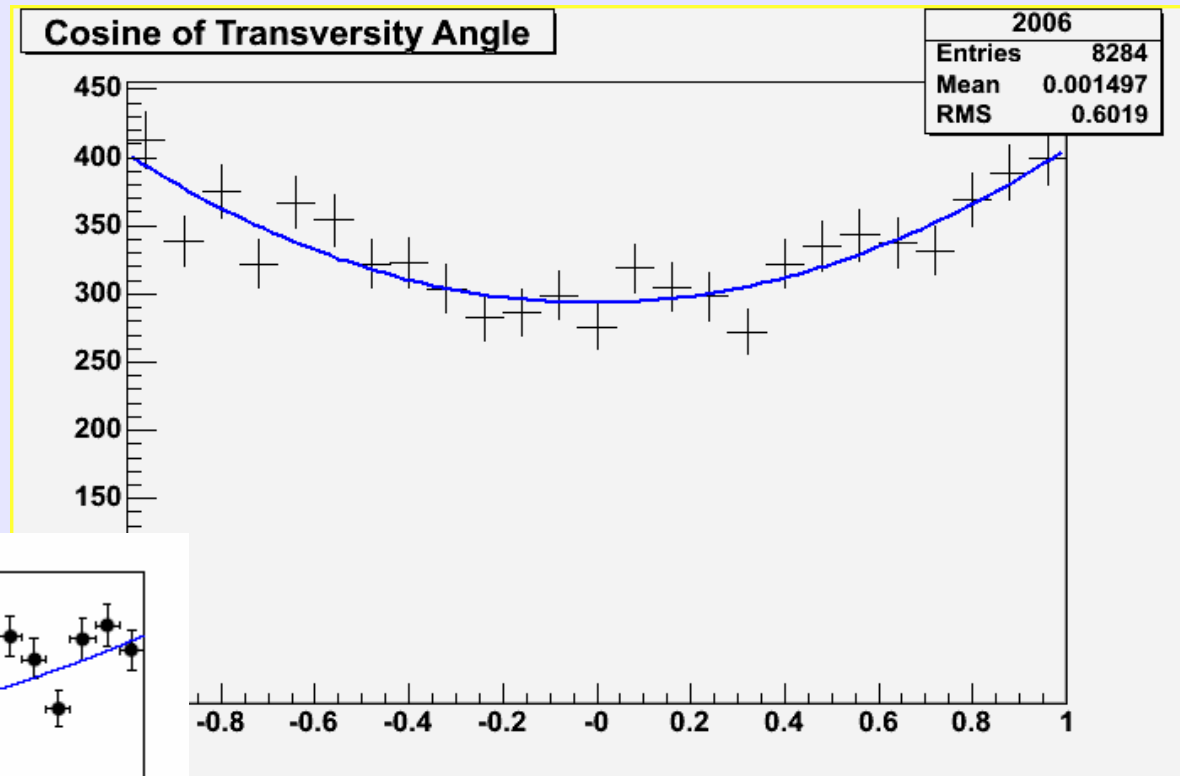
- Analysis is possible
- When it works, system is relatively painless to use
- Most errors were due to the prototype being very fragile
- Lots of scope for improvement in more stable releases

• Conclusions

- DaVinci was successfully introduced to GLite
- Job splitting was possible
- Use of the GLite package manager resulted in tarballs from release area being put straight to work
- Large jobs over 100000 events have led to the exploitation of Grid resources
- The mechanism in place is scalable and could become autonomous for other versions of DaVinci
- As it stands, LHCb users could utilise the GLite Prototype

•Additional Material

•Below plot taken from LHCb 2003-119 by Gerhard Raven, used a toy MC



•Above plot is on real MC data, shows very good correlation