Introduction to Analysis Example tutorial

P. Skubic

Email: pskubic@ou.edu

Who we are

- We are DOSAR: Distributed Organization for Scientific and Academic Research http://www.dosar.org/
- You are welcome to join our bi-weekly video (EVO) meetings. Send request to be added to DOSAR email list to Prof. Greenwood: zdgreenwood@gmail.com
- If you want long-term grid access, you can request membership in the DOSAR VO

Typical data analysis steps in particle physics

- Step 1: create files containing simulated data
- Step 2: analyze simulated data
- Step 3: collect real data from detector
- Step 4: analyze real data
- Step 5: compare simulation with real data
  - If there is good agreement, limits can be set on existence of new physical states (i.e. particles)
  - If there is disagreement, further study is needed
    - Possible mistake?
    - Possible new discovery?
- We will illustrate steps 1 and 2 today

Notes

- In particle physics software tools such Madgraph or Isajet (event generators) and GEANT (to simulate our detector response) are used in Step 1
  - We will use a simple random generator of Gaussian distribution in Root
- Typically (almost) the same reconstruction software is used for Step 4 and Step 2
- Root is a powerful tool to read and analyze large amounts of data

Condor submission script

```
universe=grid
grid_resource=gt2 osgitb1.nhn.ou.edu/jobmanager-condor
executable=run-root.sh
transfer_input_files = run-root.C
transfer_executable=True
when_to_transfer_output = ON_EXIT
log=run-root.log
transfer_output_files = root.out,t00.root,t01.root
output=run-root.out.$(Cluster).$(Process)
error=run-root.err.$(Cluster).$(Process)
notification=Never
queue
```

Step 1: Create simulated data by running Root on the Grid

- Contents of execution script: **run-root.sh**

```
#!/bin/bash
/usr/local/bin/root -b < run-root.C > root.out
```

This command executes Root in batch mode using macro **run-root.C** and routes output to file **root.out**

Step 1: Create simulated data by running Root with macro **run-root.C**

- Create TFile 0 for “run 0” (t00.root)
- Create TTree object (“t0”) to store data in Root
  - Generate 100 “events” each with Gaussian distributed “Energy”
  - Fill TTree branches for each event
- Write TFile 0
- Close TFile 0
- Repeat above steps to create TFile 1 for “run 1” (t01.root)

Step 2: make TSelector

```
TFile f("t00.root");  //open file
t0.MakeSelector("s0");  //create TSelector “s0”
f.Close();  //close file
```

This creates two files with code: s0.C and s0.h

We will modify these files to add a histogram of the Energy variable and use them to process the simulated data on the Grid

Conclusion

- After completing Steps 1 and 2 you are in principle ready to scale up and make TTree's with hundreds of variables and create and analyze thousands of files
- If time permits you can try adding your own features to the existing example by adding variables and histograms, etc.
- Good luck and have fun!!