

b-tagging a muy alto p_T
Producción privada de MC (CSC)

L. March, E. Ros, M. Vos
IFIC – Valencia

ATLAS Valencia Analysis Meeting - 24 sept. 2007

Tipo de análisis y motivación

- **Análisis previos en física exótica** (notas de física de ATLAS):
 - Search for Kaluza-Klein excitations of the gluon in models with extradimensions (ATL-PHYS-PUB-002)
 - Search for hadronic decays of Z_H and W_H in the Little Higgs model (ATL-PHYS-PUB-003)
 - Twin Higgs model (mirar charla de M. Vos)



Desintegraciones hadrónicas con b-jets a muy alto p_T en estado final



b-tagging a muy alto p_T

$Z_H \rightarrow b \bar{b}$ $M(Z_H) = 2 \text{ TeV}$ Little Higgs s-channel

Importante para nueva física, más allá del Modelo Estándar, y con desintegraciones de quarks b que producen b-jets a muy alto p_T

Producción privada de MC (CSC) [1]

- Producción privada CSC dentro del grupo de Exóticos para realizar el estudio de b-tagging a muy alto p_T
- La producción consiste en:
 - 3 datasets: $Z_H \rightarrow u \bar{u}, b \bar{b}, c \bar{c}$, con $M(Z_H) = 2 \text{ TeV}$ y
 - 20000 sucesos/dataset \rightarrow total: 60000 sucesos con Full Simulation
- Full Simulation (dividida en 3 pasos):
 - EvGen con release 12.0.5.1
 - Input data: no requiere (en este caso)
 - Output data: evgen.pool.root
 - Simul+Digit (Geant4) con release 12.0.31.8
 - Input data: evgen.pool.root
 - Output data: simul.HITS.pool.root y digit.RDO.pool.root
 - Recon con release 12.0.6.1
 - Input data: digit.RDO.pool.root
 - Output data: ESD.pool.root, AOD.pool.root y NTUP.root

Producción privada de MC (CSC) [2]

- Full Simulation (dividida en 3 pasos):
 - EvGen con release 12.0.5.1 (no requiere input data)
 - ATHENA jobOptions (Generador: Pythia), uno por cada proceso físico
 - Validado por Exotics MC production managers
 - Generados por el Sistema de Producción (ProdSys)
 - 5000 sucesos/fichero, ~ 193 MB/fichero (12 ficheros en total)
 - Simul+Digit (Geant4) con release 12.0.31.8 (input data: EvGen)
 - Simul + Digit en un mismo trabajo
 - 50 sucesos/(simul and digit job) → 1200 ficheros (simul y digit) en total
 - ~ 88 MB/simul file + ~ 138 MB/digit file → x 1200!
 - Recon con release 12.0.6.1 (input data: Digit)
 - 250 sucesos/recon job → 5 digit files como input (50 sucesos/digit file)
 - 80 recon jobs/dataset → x 3! (240 ficheros en total)
 - ~ 760 MB/ESD file, ~ 137 MB/AOD file, ~ 37 MB/NTUP file → x 240!

Problemas encontrados y sus implicaciones

- Competencia en la utilización de recursos con la producción oficial de CSC:
 - El coord. de física dijo que en los Tier-2 tiene preferencia la producción oficial
 - Se ha logrado hacer la producción completa en 1 mes (a partir de obtener los sucesos de generación), utilizando recursos Grid de ATLAS
- Baja eficiencia de reconstrucción de trazas en el Inner detector
 - Debido a que en el algoritmo de reconstrucción de trazas, newTracking, se observa una baja eficiencia a muy alto p_T (en comparación con otros algoritmos: iPatRec, xKalman) se está estudiando este efecto



Con cada nueva release, se hace una nueva pequeña producción para comprobar sus resultados: 12.0.6.1, 13.0.10.1 y en proceso 13.0.20.X

- Análisis de los datos
 - En lxplus (CERN), dado que todas la releases de ATHENA están en AFS. En lxplus, los datos deben copiarse al PC o estar en CASTOR-CERN
 - En el IFIC con la release 12.0.6 (instalada en AFS)

Estimación de recursos necesarios

- **Procesado de datos:**
 - **Simul + Digit (Intel Xeon CPU @ 2.00 GHz)**
 - Simul: ~ 866 seg./suceso
 - Digit: ~ 13 seg./suceso
 - **Recon (Intel Pentium 4 CPU @ 3.20 GHz)**
 - ESD + NTUP: ~ 105 seg./suceso
 - AOD: ~ 4.1 seg./suceso
 - **Tiempo total (60000 sucesos): ~ 16470 horas ~ 686 días (en un PC)**
 - **Estimación de recursos: ~ 20 CPUs potentes para pequeñas producciones (obteniendo ~ 60000 sucesos/semana) (las grandes producciones se deben hacer en los Tier-2, pero hace falta recursos locales para “pequeñas” producciones para usuarios IFIC)**
- **Almacenamiento: ~ 0.5 TB = 500 GB (para esta producción)**

Conclusiones

- Producción privada de MC (dentro de CSC)
 - b-tagging a muy alto p_T (importante para física exótica)
 - 60000 sucesos con Full Simulation en ~ 1 mes utilizando recursos Grid
 - Competencia de recursos: prioridad producción oficial CSC
- Análisis de datos MC
 - Normalmente en lxplus (CERN) porque tienen todas las releases de ATHENA, aunque se compite con otros usuarios conectados y los datos hay que copiarlos a lxplus o estar almacenados en CASTOR-CERN
 - En el IFIC, utilizando la release 12.0.6 instalada en AFS (comparando, sin duda mucho más cómodo y rápido que lxplus)
- Estimación de recursos locales (IFIC)
 - Procesado: ~ 20 CPUs potentes (pequeñas producciones de usuarios)
 - Almacenamiento: 0.5 TB (para esta producción de 60000 sucesos)