LHCb Experience with Parallel Processing And Virtualization

Some LHCb background

- Event sizes
  - Raw data: 35kB 60k events / file
  - Reconstructed data (DST): 70kB (300kB with MC)
  - MicroDST: 5kB
- Event samples
  - Raw data / year: $2\ 10^{10}$
  - Events / physics analysis stream: 0.1 to 10 million
- Processing time
  - Full reconstruction and stripping: ~1s / event
  - After stripping, no more reconstruction needed, user provided with B decay candidates, “only” tighter selection required to suppress background, ~10 ms / event or less
  - MC production, full detector simulation: ~100s / event
  - Toy MC studies: ?

Analysis Pattern

- For developing a selection / analysis algorithm, fast turnaround on reasonable amount of data is important
  - Standard batch queues, e.g. lxplus, queuing issue, requires additional step to merge output from different jobs
  - Grid jobs, e.g. using Ganga, merging ok, however large overhead to get results, 80% success rate, additional bookeeping of failed jobs !
  - Quick interactive running most effective
- For final running on full data set, waiting time less important, however, the smaller the number of individual jobs, the smaller the number of potential problems, less bookkeeping issues.
  - Waiting half of day for final result, not a problem
  - Getting feedback after 5 minutes if something is wrong, forgotten, etc., makes all the difference.

Running on Multicore Machines

- LHCb: Job configuration and starting done with python
- Many people also use python scripts for handling the event loop
  - simple processing, steering of C++ algorithms and tools from python using dictionaries
- Some people in LHCb started about 1 year ago with the help from Pere Mato to use the processing python module to run one GaudiPython script on multicore machines in parallel.
  http://pyprocessing.berlios.de/
- Very simple to use, almost no change to normal scripts:

```
class MyTask(Task):
   def process(self, file) :
task = MyTask()
wmgr = WorkManager(ncpus = 8)
wmgr.process(task,files)
myHistograms  = task.output
```

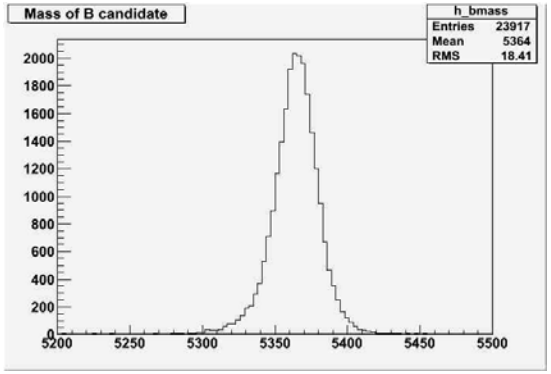
contains event loop, reco/analysis steps

splitting per file

automatically merged

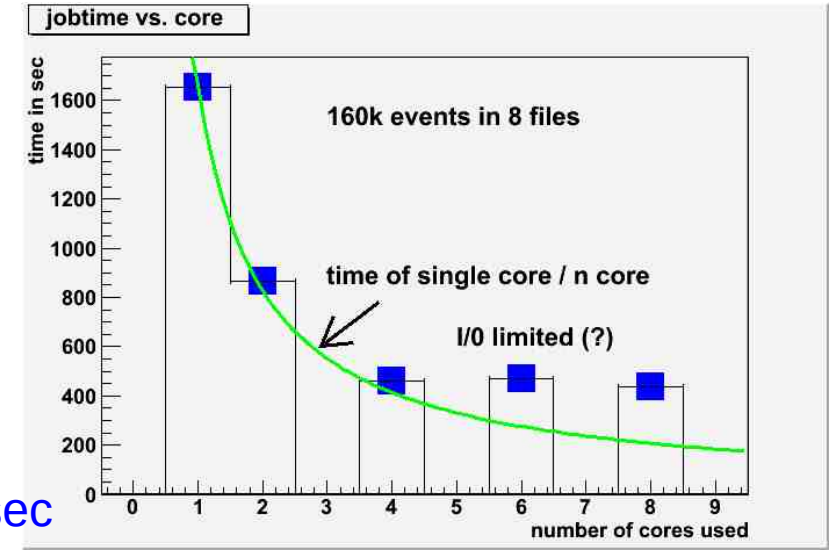
Examples

- run Physics selection,
  fill histogram with B mass



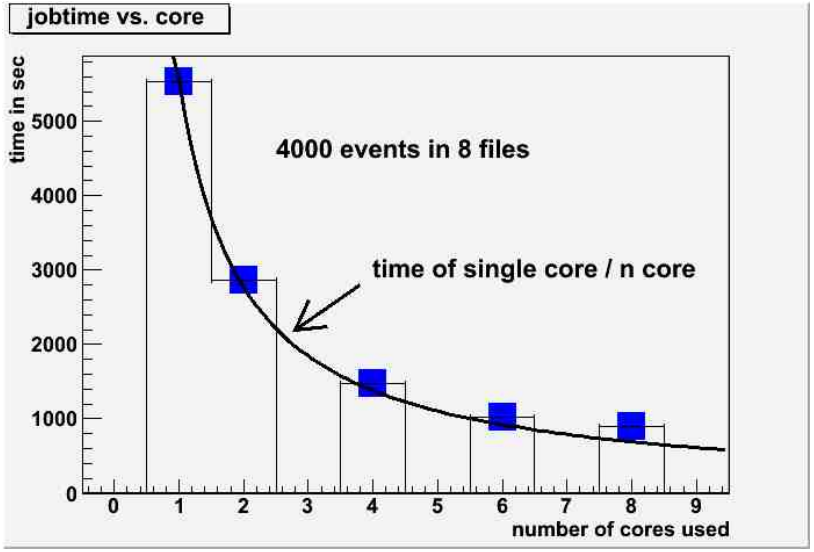
320 events/sec
~15 MB / sec

µDST :
~3000 events/sec



- run full reconstruction

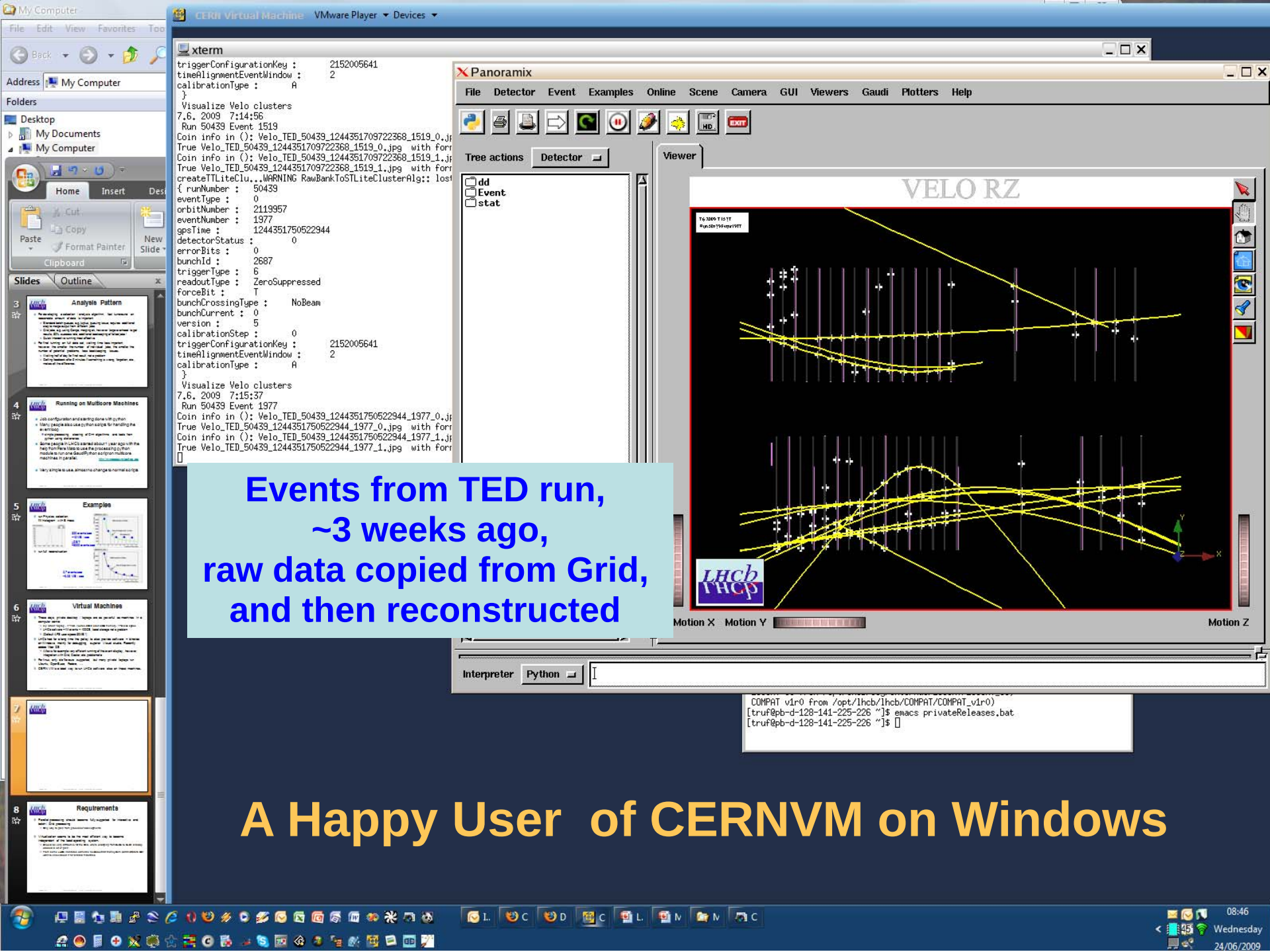
0.7 events/sec
~0.25 MB / sec



Virtual Machines

- These days, private desktop / laptops are as powerful as machines in a computer centre
  - My latest “laptop”, i7-750, 4 cores, 3Ghz each, 6GB memory, 1TB disk space
  - LHCb software, ~5GB, ~1M events < 100GB, local storage not a problem (Default AFS user space 50MB !)
- LHCb had for a long time the policy to also provide software + binaries on Windows, mainly for debugging, superior Visual studio. Recently added Mac OS
  - Allows, for example, very efficient running of the event display, however, integration with Grid, Castor, etc. problematic
- For linux, only slc flavours supported, but many private laptops run Ubuntu, OpenSuse, Fedora, …
- CERN VM is a ideal way to run LHCb software also on these machines.

  O(10-20) people are using it in LHCb, Windows/Mac/Linux
  expect more people in future


My Computer
CERN Virtual Machine
VMware Player
Devices
xterm
triggerConfigurationKey : 2152005641
timeAlignmentEventWindow : 2
calibrationType : A
Visualize Velo clusters
7.6. 2009 7:14:56
Run 50439 Event 1519
Coin info in (): Velo_TED_50439_1244351709722368_1519_0.jp
True Velo_TED_50439_1244351709722368_1519_0.jpg with forr
Coin info in (): Velo_TED_50439_1244351709722368_1519_1.jp
True Velo_TED_50439_1244351709722368_1519_1.jpg with forr
createTTLiteClu...WARNING RawBankToSTLiteClusterAlg:: lost
{ runNumber : 50439
eventType : 0
orbitNumber : 2119957
eventNumber : 1977
gpsTime : 1244351750522944
detectorStatus : 0
errorBits : 0
bunchId : 2687
triggerType : 6
readoutType : ZeroSuppressed
forceBit : T
bunchCrossingType : NoBeam
bunchCurrent : 0
version : 5
calibrationStep : 0
triggerConfigurationKey : 2152005641
timeAlignmentEventWindow : 2
calibrationType : A
Visualize Velo clusters
7.6. 2009 7:15:37
Run 50439 Event 1977
Coin info in (): Velo_TED_50439_1244351750522944_1977_0.jp
True Velo_TED_50439_1244351750522944_1977_0.jpg with forr
Coin info in (): Velo_TED_50439_1244351750522944_1977_1.jp
True Velo_TED_50439_1244351750522944_1977_1.jpg with forr
Panoramix
File Detector Event Examples Online Scene Camera GUI Viewers Gaudi Plotters Help
Tree actions
Detector
dd
Event
stat
Viewer
VELO RZ
Motion X
Motion Y
Motion Z
Interpreter
Python
COMPAT v1r0 from /opt/lhcb/lhcb/COMPAT/COMPAT_v1r0)
[truf@pb-d-128-141-225-226 ~]$ emacs privateReleases.bat
[truf@pb-d-128-141-225-226 ~]$
Events from TED run,
~3 weeks ago,
raw data copied from Grid,
and then reconstructed
A Happy User of CERNVM on Windows
08:46
Wednesday
24/06/2009

Requirements

- Parallel processing should become fully supported for interactive and batch / Grid processing
  - Only way to gain from processor developments
- Virtualization seems to be the most efficient way to become independent of the local operating system.
  - Should be very attractive for the Grid, where changing from SLC3 to SLC4 already caused a lot of pain
  - From some LHCb institutes came the feedback that their system administrators don't want to know about it for shared machines.