Experience at the OSU T3

Waruna Fernando, Harris Kagan
Ohio State University

What we needed/had

- Grid/Pathena
  - Submit grid jobs to get official datasets in D3PD form
- Athena for full simulation
  - Generate custom datasets
- Proof cluster for analysis
- Cadence/Hspice for chip design

- limited resources
  - for hardware
  - No tech support people
- Access to department network
- LDAP/NFS

What we built

- A proof cluster with 46 workers
- MC farm which can make ~10K events/day (full chain final D3PD)
- 25TB disk space as a network disks (NFS)
- 6TB disk space distributed (/data1)
- Grid access and pathena submission

How we built - Hardware

- One Dell PowerEdge 2950 (8 cores 8 threads 2.5Ghz, 16GB RAM) also disk server
- One Core2duo Desktop (2 cores 2 threads)
- Three Intel Core2quad based Desktops (4 cores 4 threads)
- Three Intel Corei7 based Desktops (4 cores 8 threads)
- Total 34 cores 46 threads

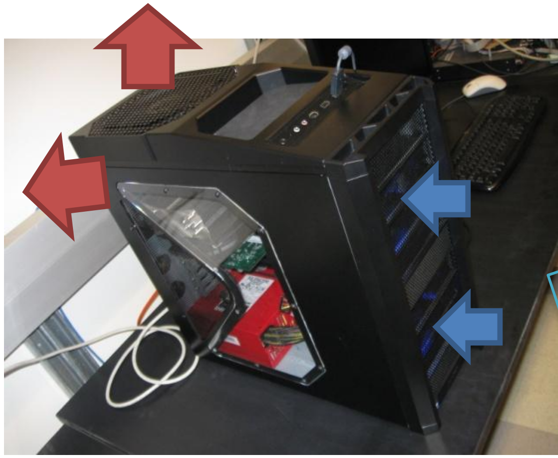
How we built - Hardware

- High performance desktops
  - Custom built
  - All the parts individually selected for best performance for lowest price
  - Require stability @100% load for > 30 days
  - Two models
    - Q6600 based (old discontinued)
    - Corei7 920 based

Hardware configuration

Part	Price ($)
CPU - corei7 920	230
Motherboard ASUS P6T Deluxe	300
2X OCZ Platinum 6GB DDR3 1600	150
2X Western Digital WD6400AAKS 640GB	140
PC Power & Cooling S75CF 750W	90
Heat sink and fans	80
9500GT video card	45
Case (Antec Nine Hundred or PowerSpec C5)	40-100
total	$1075-$1135

Cooling

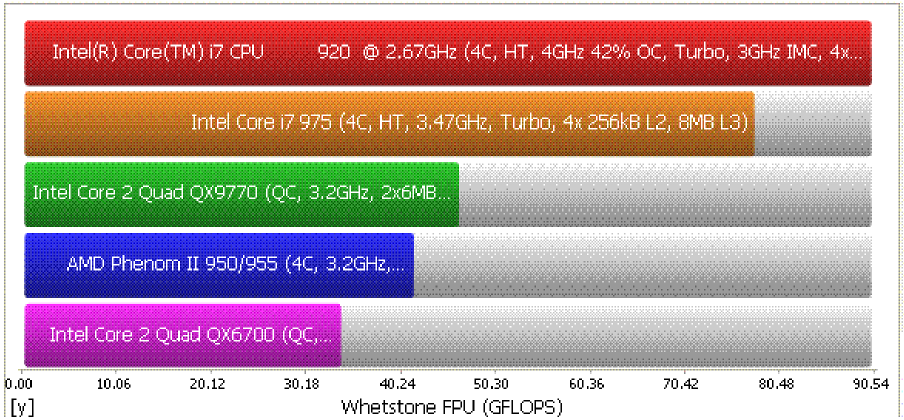


5 fans for cooling
Big heat sink

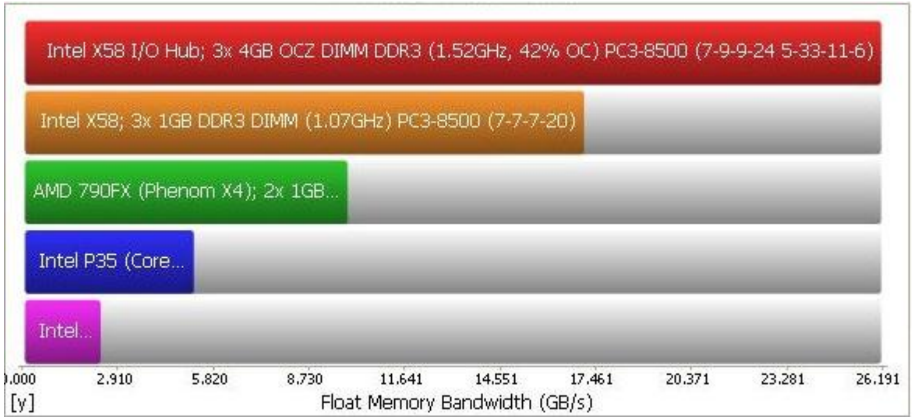
How we built –over clocking

- Core i7 920/Q6600 stock speeds 2.66/2.4GHz
- Tuned the bios settings for over clocking
- Memtest for memory stability
- Used Mprime for stress test all threads for 17hr
- Final speeds
  - i7920: 4.02GHz,3.95GHz,3.83GHz
  - Q6600: 3.7GHz,3.5GHz,3.48GHz
  - E6600: 3.4GHz

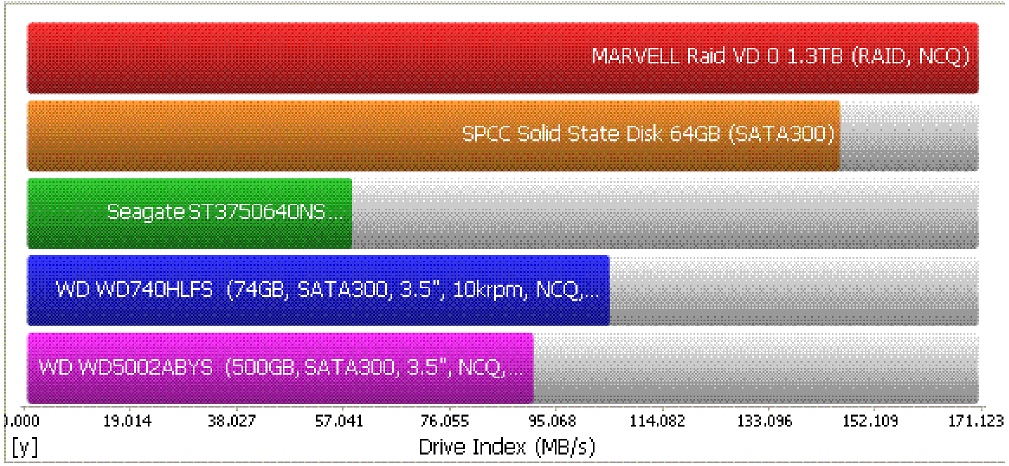
Benchmarks - synthetic



Processor Arithmetic



Memory Bandwidth



Drive index (Read)

Benchmarks Athena job transformations

- Z→tautau job used for benchmarking
- Measured the real time take for full chain of 10 events for all threads. (number of threads =number of parallel jobs)

Platform	Speed (GHz)	Threads	Total time (s) /10 events	Speed/thread/event
Dell PowerEdge 2950	2.50	8	6600	82.5
Q6600	3.70	4	3125	78.1
Corei7 920	4.02	8	3420	42.7

Software

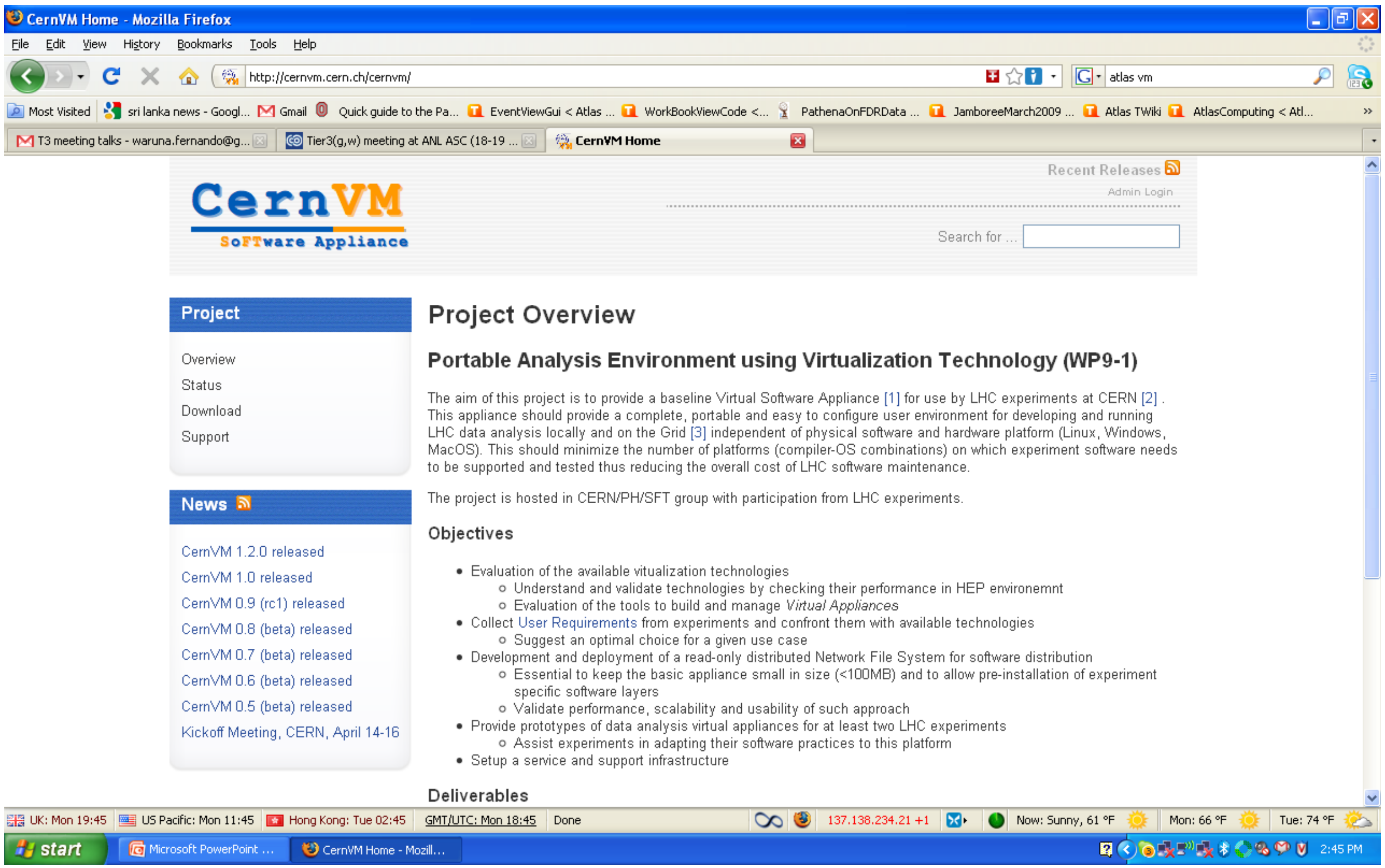
- E6600 machine had stable drivers for RHEL and able to run athena/pathena/grid access form that
- No stable drivers for RHEL for the other new hardware
- SUSE has drivers but hard to set it to use old compilers for athena
- This problem solved by using Virtual Machines (VM)
- Machines built with SUSE 11.1 (64 bit)
- PROOF setup on SUSE
- Athena job transformations through VM’s
- No Job scheduler: replaced with nomachine

Software- Virtual machines

- canadianvm:
  - Had issues that we could not fix in job transformations:
    - Algorithm of type Pythia is unknown
    - Algorithm of type MboyDigiEmptyLoop is unknown
- cernvm:
  - Had a issue with DBreleases in digitizing - fixed.
  - Fast (except the very first time)
  - Extremely easy to setup, almost no maintenance
  - Change : setup KDE,LDAP, use 2cores,2.5GB RAM, 20GB for /, home, temp, 25TB NFS
  - Have 22 VM running… (44 parallel jobs)

Atlas OS built with CernVM tool

- http://cernvm.cern.ch/cernvm/



Software- PROOF cluster setup

- Download the ROOT source for the latest release (5.22.0: ftp://root.cern.ch/root/root_v5.22.00.source.tar.gz) to a network disk and untar it
- Go to the machine you want it install as a super user
  - export ROOTSYS=/usr/local
  - export LD_LIBRARY_PATH=/usr/local/lib
  - cd to root source
  - make clean
  - ./configure linux --with-f77=/usr/bin/gfortran --enable-xrootd --enable-soversion --enable-ssl --enable-roofit --enable-python
  - may need to add few missing lib's or/and add simlinks to soversions
  - make
  - make install
- Repeat this for all the machines you want in the cluster

Software- PROOF cluster setup

- Setup cluster in proof
  - Need to setup few scripts
      - http://www.physics.ohio-state.edu/~waruna/proof.conf
      - http://www.physics.ohio-state.edu/~waruna/xpd.cf.sample
      - http://www.physics.ohio-state.edu/~waruna/xpd.groups.sample
    - Change the master(cadence105→your master)
    - Change the workers (cadenceXX→your workers)
    - Set the user ids in xpd.groups.sample
  - Replace the original files at /usr/local/etc/proof/ (as su)
  - Add the following to your .bashrc
      - export ROOTSYS=/usr/local
      - export LD_LIBRARY_PATH=/usr/local/lib
  - /usr/local/bin/xrootd -c /usr/local/etc/proof/xpd.cf.sample
      - If xrootd successfully initialized ☺ run this as a background job
  - Do this for all the machines

Analysis

- Most of analysis codes that we are developing with root →proof
- Currently we are using egammaD3PD maker (NAEgamma) to convert all the interesting official datasets through pathena
- Analysis on proof
- Hope to move to Sframe soon

Summary

- We built a T3 system for low cost, low maintenance
- Very easy to expand due to modern operating system, standardized atlas VM software and simple hardware
- Can serve what we need