Hans Wenzel

Resume:

Current Professional Activities

Apr. 2000 - present Computer Professional VI at FNAL

Currently I am in charge of the operation of the Fermilab Tier 1 computing facility for CMS. I was in charge the design of linux based computing farms and infrastructure, purchasing and installing computing equipment and overseeing installation and commissioning. At Fermilab we are serving a dual mission of supporting physicists and doing R&D towards a fully functional prototype Tier 1 Regional Center. This position requires good relationships with CERN, the US CMS collaborators (e.g. IVDGL) and other departments and experiments at Fermilab. In my opinion Fermilab is in a unique position since we can learn from the experience of the Run II experiments (CDF and D0) which are currently taking data. Therefore I maintained good relationships with Run II experiments. Interestingly many of the products selected for Run II analysis facilities (FBSNG batch system, dCache, Enstore) are also in use by CMS. We benefitted immensely from having access to the hardware owned by CDF.

The main focus of this years R&D work is doing in-depth studies of modern file systems and logical volume management on large scale disk farms and load balancing and dynamic partitioning of cluster computing resources:

Large scale disk farms:

One major research area is related to high performance and reliable data servers for e.g. production style applications (DB services, file transfer etc.) which have to be smoothly integrated between the production farm nodes, the user analysis nodes and the backend tertiary storage services (enstore). The primary intention is to let them act as a transparent cache by making a multi-terabyte server farm look like one coherent and homogeneous storage system. dCache (developed at DESY), available as part of the centrally maintained mass storage services at Fermilab, seems to offer the right functionality and is currently being evaluated by CMS and also by the CDF collaboration. First results look very promising. Among the required functionality the following are the most important:

- rate adapting between the application and the tertiary storage resources,
- optimized usage of expensive tape robot systems and drives by coordinated read and write requests,
- no explicit staging is necessary to access the data,
- the data access method is unique independent of where the data resides,
- even without the backend HSM (enstore) dCache can be seen as a huge data store with a unique namespace and standardized access methods,

- policies can be applied in order to control the data flow between files on disk and magnetic tape,
- High performance and fault tolerant transport protocol between applications and data servers
- files in dCache can be accessed directly from an application via posix compliant function calls (e.g. the ROOT TDCacheFile class).

For user space we investigated and to manage the disk space for the different physics groups we investigated LVM (Logical Volume Manager). LVM supports enterprise level volume management of disk and disk subsystems by grouping arbitrary disks into volume groups. The total capacity of volume groups can be allocated to logical volumes, which are accessed as regular block devices. Further, LVM provides logical separation of storage, the ability to move data from one physical device to another while on-line, and dynamic block device resizing. LVM also enables system administrators to upgrade systems, remove failing disks, reorganize workloads, and adapt to changing system needs, through a minimum amount of time and effort. LVM supports enterprise level volume management of disk and disksubsystems by grouping arbitrary disks into volume groups. The total capacity of volume groups can be allocated to logical volumes, which are accessed as regular block devices. Further, LVM provides logical separation of storage, the ability to move data from one physical device to another while on-line, and dynamic block device resizing. LVM also enables system administrators to upgrade systems, remove failing disks, reorganize workloads, and adapt to changing system needs, through a minimum amount of time and effort. During the last month we have been testing the features and robustness of LVM on a test system (running the RedHat 7.3 version of Linux) with very promising results and we plan to deploy it to manage the disk space of the different physics groups.

load balancing and dynamic partitioning of cluster computing resources:

Another major research area are load balancing and dynamic partitioning of cluster computing resources especially for interactive use. Currently PC's running Linux offer the best price/performance ratio. Therefore several scenarios to make the architecture of a Linux cluster completely transparent to end users (physicists) are being investigated.

- The Mosix operating system is designed to make a cluster look like an smp machine. MOSIX is a software that was specifically designed to enhance the Linux kernel with cluster computing capabilities. (highly configurable) MOSIX operates silently - its operations are transparent to the applications. The algorithms of MOSIX are decentralized. We deployed Mosix on a test farm and found is easy to install and to configure using the provided installation scripts. In the end we decided not to use Mosix for the following reasons:
  - it is a specialized Kernel and it requires additional effort to compile and enable other Kernel modules required by CMS software (e.g. AFS),
  - some processes (utilizing socket, shared memory) don't migrate. Unfortunately the CMS reconstruction applications belong to this class. This severely limits the use of Mosix for the CMS experiment.

  - in our tests the mosix filesystem which makes the local FS transparently available on all cluster nodes did not perform very reliable resulting in poor performance of IO intensive processes.

- FBSNG (Fermilab Batch System Next Generation) has been used very successfully in CMS Monte Carl production over the last 2 years. In the solution we are currently investigating the user connects via ssh or telnet into a gateway node, consequently the log in process triggers a script which creates an interactive batch job on one of the farm nodes. We developed the scripts and demonstrated that this solution works. If we decide to use FBSNG as unique tool for interactive as well as batch computing this would minimize the administrative overhead allowing us to dynamically assign resources (nodes) to batch, user batch and user interactive queues.

- LVS Linux Virtual Server. Investigation planned in he next few weeks.

My position also required the development, integration and support of physics reconstruction and analysis software tools for distributed computing for the CMS experiment at CERN. It is required to install CMS software packages from CERN onto FNAL computers, making required modifications to adapt to the FNAL environment to be used throughout the collaboration. Contribute in a group studying models of distributed computing architectures for future High Energy Physics experiments. C and C++ programming language and Unix software development.

In Fall 2000 I developed software to run large Monte Carlo productions for the CMS experiment using Linux computing farms as well as sun solaris smp machines. The programs developed at Fermilab were chosen as the base for the production tools used throughout the CMS experiment. Management of object oriented Databases (Objectivity) which currently is the main way to store CMS data. I served as liaison between and the CMS Jets/Met physics analysis group and the CMS MonteCarlo production group and maintained a repository of data cards to control the Monte Carlo production of various physics channels. I developed and presented CMS software tutorials for US physicists.

In addition I served as the librarian for OSCAR which is a geant 4 based simulation of the CMS detector.

Developed a fitting program which allows to combine different data sets statistically correct using object oriented techniques based on the ROOT software package. This work resulted in a Ph.D thesis (Universität Karlsruhe) and has been accepted for publication by physical review.

Jan. 1997-March 2000 wissenschaftlicher Angestellter Universität Karlsruhe
Jan. 1997-Sep. 1998 located at Fermilab

- **Software**

  For the CDF run II upgrade I am working on the online monitoring. CDF will start Run II with a significantly different detector, trigger system, data format and software environment. Therefore, the programs which served to monitor the data stream during

Run I have to be rewritten. For this purpose together with Kaori Maeshima and a small group from Waseda University (Japan) developed a framework based on the ROOT package. The framework uses a client/server architecture and allows real time remote access while data taking. This is realized via Unix sockets or a WWW server and a specialized Browser. The ROOT package provides the C++ classes for developing GUIs, socket-connections, accessing shared memory, histogramming, containers etc.. ROOT is called from within the CDF run II environment (AC++). The code is running on all three platforms supported by CDF namely IRIX/OSF/LINUX and I was in charge of the CVS repository for this package.

Software activities included the porting of CDF vertex fitting code from FORTRAN to C++. In addition,together with a diploma student under my supervision we created C++ Bank classes allowing Run 1 Tracking banks to be read out. This will enable us to do some simple physics analysis within the new CDF run II (C++) framework but using CDF Run I data.

- **Hardware**
  Beginning of 1999 testing of silicon detectors for the ISL run II upgrade of CDF. In 1998 I supervised a Ph.D student on developing a readout system for the SVX3 chip using very cost effective electronics and a PC running Linux.

- **Physics analysis**
  As a member of CDF my main physics interest were in the field of b and top-physics where I could benefit from my knowledge of the SVX tracking. I was mainly working on b-lifetimes, rare b-decays and $B\bar{B}$ - mixing. While performing searches for rare b-decays I became interested in exotic physics since precision measurements of rare b-decays provide tests of physics beyond the Standard Model. In 1998 I concluded three analyses resulting in publications in Physical Review. Currently I am preparing an article summarizing b-physics at hadron colliders.

- **Teaching**
  Organized the tutorial groups for the nuclear and particle physics course. Supervised students in laboratory courses.

  Supervising one Diploma student (topic: "Search for single top "), and one Ph.D student his (topic: "Measurement of the b-lifetime using exclusive B-decays").

- **Committees**
  Representing CDF in the Physics Analysis Software Joint Project which evaluates software tools for Run II (http://runiicomputing.fnal.gov/runiiweb/pas.html). Computing representative for the Karlsruhe CDF group.

Dec. 1993 - Dec 1996 Post doctoral fellow at Lawrence Berkeley Laboratory

As part of the CDF validation and calibration group I was responsible for Valplots which is a software package which checks the quality and for possible bugs (e.g. use of wrong calibration constants) of the offline reconstruction of data. The package looks for physics objects like electrons, muons, Jets, J/ψ or detector oriented quantities like tracks in the

vertex or central tracking detectors. The same batch job resulted in the calculation, with high precision, of the beam profile and position using the silicon vertex detector (SVX). In addition, the Luminosity of each file was calculated. Using data provided by Valplots one could also track the long term performance of the detector and understand e.g. the degrading tracking performance with higher integrated and instantaneous luminosities. Besides developing monitoring programs one had to develop DEC/DCL scripts to run large jobs on thousands of tapes.

Together with Richard Kadel I worked on developing a photo-detector which would be able to detect light in the vacuum ultra violet with high efficiency and which would be able to work in high magnetic fields. One application for this device would be a Cerenkov detector system which would improve the CDF particle identification capabilities for run II. We tested different designs and analyzed various photo-cathode materials (KCl, CsI) and different gases. This study required the building of a test set up consisting of vacuum system, UV-spectrometer, computer control and DAQ.

Until November 1995 I served as convener of the CDF B mixing group. Together with Manfred Paulini I organized a small workshop at LBL to bringing together all groups within CDF intending to work on $B\bar{B}$-mixing and to organize the effort.

I volunteered to serve as a so called 'ACE' (person who runs the DAQ during shifts). In addition I served as God Parent for three different analyses.

I was mainly working on b-lifetimes, rare b-decays and $B\bar{B}$ - mixing leading to publications in 1998.

July 1991 - Nov. 1993 Fermilab Guest scientist. Thesis experiment CDF

Feb. 1990 - June 1991 INFN Pisa

For CDF Run 1 I developed software to reconstruct tracks and vertices in the Silicon Vertex detector (SVX). The SVX detector was the first silicon vertex detector operating successfully at a hadron collider. The SVX played a central role in the discovery of the top quark and in developing a very successful b-physics program which competes with and complements the b-physics programs at e^+e^- machines.

The tracking uses a progressive method which starts with tracks from the outer tracker and updates the fit with hits found in the Silicon Tracker. The developed code worked fast and reliably through the entire run I. This work also included documenting the algorithm and bank structure as well as developing and running test suites to authenticate the functionality of the software. Besides that I was involved in other aspects of SVX software like simulation, defining and validating the geometry, hit generation and the bank structure.

I also developed several programs to measure various performance parameters of the SVX detector (e.g. alignment, signal/noise ratio, Landau distributions, tracking efficiency, Hall effect shifts etc.). This work was documented in various NIM articles.

In addition I developed code to determine the position and other parameters of the Tevatron beam and maintained a data base where this information was stored. This information was used in many physics analyses as an estimate of the primary interaction vertex.

All the code was mainly developed on VAX/VMS platforms using FORTRAN. I served as librarian for the SVX CMS code repository.

Besides doing the tracking and SVX software as service work for CDF I did several Monte Carlo physics studies regarding top and b-physics for CDF and LHC.

My Ph.D thesis was the first physics result of Run 1a fully utilizing the SVX detector and the first b-lifetime measurement at a hadron collider.