# Farming Data for the HyperCP Experiment

Sirena Boden, Tim Jones, Piotr Zyla (Lawrence Berkeley Laboratory, USA)
Tim Holmstrom, Lan-chun Lu * (University of Virginia, USA)
Durga Rajaram, Howard A. Rubin (Illinois Institute of Technology, USA)

**Abstract**

The main goal of the HyperCP(E871) experiment at Fermilab is to search for CP violation in $\Xi$ and $\Lambda$ decays at the $\sim 10^{-4}$ level. This level of precision dictates a data sample of over a billion events. The experiment collected about 231 billion raw events on about 30,000 5-GB tapes in ten months of running in 1997 and 1999. In order to analyze this huge amount of data, the collaboration has reconstructed the events on a farm of 55 dual-processor Linux-based PCs at Fermilab. A set of farm tools has been written by the collaboration to interface with the Farm Batch System (FBS and FBSNG)[1] developed by the Fermilab Computing Division, to automate much of the farming and to allow non-expert farm shifters to submit and monitor jobs through a web-based interface. Special care has been taken to produce a robust system which facilitates easy recovery from errors. The code has provisions for extensive monitoring of the data on a spill-by-spill basis, as is required by the need to minimize potential systematic errors. About 36 million plots of various parameters produced from the farm analysis can be accessed through a data management system. The entire data set was farmed in eleven months, or about the same time that was taken to acquire the data. We will describe the architecture of the farm, our experience in operating it, and show some results from the farm analysis.

Keywords: HyperCP(E871), Fermilab, CP violation, farm, web-based interface.

## 1 Introduction

Since $CP$ violation was discovered in 1964[2], substantial effort has been dedicated to understand its nature and magnitude. Most experiments, such as KTeV at Fermilab[3] and NA48 at CERN[4], measure the $CP$-violation parameters in neutral-kaon decays. Theory indicates that $CP$ violation not only occurs in the kaon system, but also in the decays of hyperons[5] and neutral $B$ mesons[6].

The HyperCP(E871) experiment was proposed in 1993 and began running in 1997 at Fermilab[7] to study hyperon $CP$-violation by comparing the $\Lambda$ and $\bar{\Lambda}$ distributions in the $\Xi^-$ and $\bar{\Xi}^+$ decays: $\Xi^- \to \Lambda\pi^-, \Lambda \to p\pi^-$ and $\bar{\Xi}^+ \to \bar{\Lambda}\pi^+, \bar{\Lambda} \to \bar{p}\pi^+$. For the decay of unpolarized $\Xi^-$, the angular distribution of the proton in the $\Lambda$ helicity frame is given by

$$\frac{dN}{d\cos\theta} = \frac{1}{2}(1 + \alpha_\Lambda \alpha_{\Xi^-} \cos\theta), \tag{1}$$

where $\cos\theta = \hat{p}_\Lambda \cdot \hat{p}_p$, and $\hat{p}_p$ and $\hat{p}_\Lambda$ are unit vectors in the directions of motion of the proton and $\Lambda$ respectively, and $\alpha_\Lambda$ and $\alpha_{\Xi^-}$ are the decay parameters of the $\Lambda \to p\pi$ and $\Xi^- \to \Lambda\pi^-$ decays.

For $\bar{\Xi}^+$

$$\frac{dN}{d\cos\theta} = \frac{1}{2}(1 + \alpha_{\bar{\Lambda}} \alpha_{\bar{\Xi}^+} \cos\theta). \tag{2}$$

It is convenient to define

$$A_{\Xi\Lambda} \equiv \frac{\alpha_{\Xi^-}\alpha_\Lambda - \alpha_{\bar{\Xi}^+}\alpha_{\bar{\Lambda}}}{\alpha_{\Xi^-}\alpha_\Lambda + \alpha_{\bar{\Xi}^+}\alpha_{\bar{\Lambda}}}. \tag{3}$$

Any non-zero value of $A_{\Xi\Lambda}$ indicates direct $CP$-violation in the $\Xi - \Lambda$ decay sequence.

---

*Presented by Lan-chun Lu at CHEP'01

### 1.1 The HyperCP (E871) Spectrometer

Figure 1 shows the spectrometer used in the HyperCP (E871) experiment. $\Xi$s are generated when an 800-GeV proton beam from Tevatron strikes the target. The $\Xi$s travel through the hyperon magnet channel (collimator) and decay inside the vacuum decay pipe. The trajectories of the $\pi$ from the $\Xi$ decay and the proton and $\pi$ from the $\Lambda$ decay are measured by four proportional wire chambers upstream of the analyzing magnets. The two pions are bent to the same side (SS) of the spectrometer, and the proton is bent to the opposite side (OS) in the field of the analyzing magnets. After four downstream proportional wire chambers the two pions strike the SS hodoscope, and the proton strikes the OS hodoscope before depositing energy in the calorimeter. The muon station is used to study kaon and hyperon rare decays.

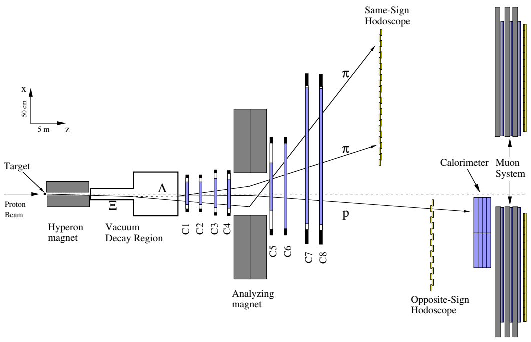


Figure 1: The HyperCP (E871) spectrometer at Fermilab.

### 1.2 Data Acquisition

There were two periods of data acquisition during 1997 and 1999. A total of 231 billion raw events, a data sample of 120 TB, were acquired and stored on 29,401 5-GB Exabyte tapes. This sample contains mainly the decays of $\Xi$s, $\Omega$s, and $K$s. About 22.7 billion of these raw events were acquired at a non-zero production angle, and will be used to measure the polarization of the $\Xi^-(\bar{\Xi}^+)$ and $\Omega^-(\bar{\Omega}^+)$, and their $\beta$ decay parameters. The large unpolarized data sample enables us to search for *CP*-violation in the $\Xi - \Lambda$ decay at the $10^{-4}$ level.

## 2 The E871 Data Processing Farm

In order to analyze this huge amount of data efficiently, a sophisticated tool has been developed to process the data using the Fermilab farms.

### 2.1 Farm Setup

Figure 2 illustrates the computers used in the E871 farm and describes the relations between them. Both the I/O nodes and the PC clusters themselves are components of the Fermilab PC farms. There are actually two SGI I/O nodes used in the E871 farm, FNSFH and FNSFO, which distribute the data in parallel to two independent processing systems. Beauty7 is a PC which submits jobs to the farms and monitors their output. All the PCs run RedHat 6.1 Linux. Beauty7 is responsible for controlling the E871 farm and maintaining the tape database, as well as monitoring the processing of any job on the farm. When a job is launched from beauty7, all databases are updated indicating that a tape has been submitted for processing. Additionally, beauty7 maintains the status of the job and keeps track of which output tape the results are

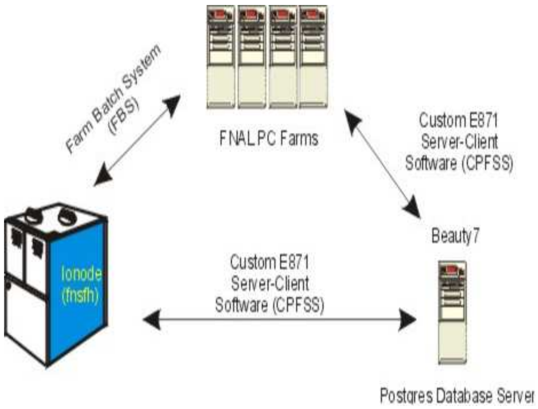


Figure 2: The HyperCP PC farm.

eventually written to. The progress of the job is tracked using custom software (HyperCP Farm Software System CPFSS) and a PostgreSQL database.

Figure 3 illustrates how data moves on the farm. The I/O node contains the I/O hardware for the farming (*i.e.* tape drives and staging disks). The arrows in the diagram show the data flow which proceeds as follows:

- Raw data is read from an Exabyte tape and split into 10 files on a staging disk on the I/O node.
- The 10 input files are copied to local disks on 10 worker nodes (PCs) using `rcp`.
- The analysis job is executed on the 10 PCs.
- Results from the analysis job (up to 8 individual output "streams") are moved from the local disk to output staging spaces on the I/O nodes, again using `rcp`.
- When an output stage is full, it is flushed to tape (Exabyte for most streams, and SONY AIT2 for the primary stream) and the stage is released for reuse.

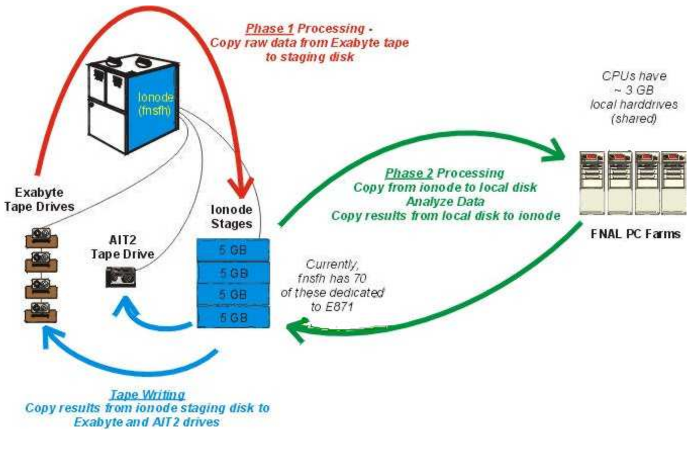


Figure 3: The flow of farm data.

The E871 Farm includes:

1. Hardware:

   (a) I/O Nodes: Two SGI I/O nodes, FNSFH and FNSFO, are used in the E871 farms to distribute the raw data. Each node has its own tape drives, staging disks, and its own farm processing system. The software on the two systems is identical (See Figure 2 and Figure 3).
   (b) Worker Nodes: 55 dual-processor Linux-based PCs are used in the E871 farms. All the data are analyzed on these PCs, with each CPU running two separate analysis jobs. The output data is copied from the local disk on the PC to the output stages on the I/O nodes.
   (c) Stages: There are 70 I/O stages on FNSFH and 27 on FNSFO. A few of the stages are used to store log files. Each stage is 5 GB (matched to the capacity of an Exabyte tape) and used to store data temporarily. When a stage is full, the data is scheduled for output to tape. The stage is cleaned up and unlocked for next job after the data is copied to output tape.
   (d) Tape Drives: Exabyte tape drives are used to load input raw data tapes and output data from several output streams. Two AIT2 tape drives are used on FNSFH and FNSFO to copy output data to 50-GB AIT2 tapes.
2. E871 Farm Software (CPFSS): CPFSS is custom software built for E871 to launch jobs, monitor the progress of jobs, manage I/O tapes, and control the flow of data on the Fermilab PC-farms. The CPFSS Client-Server Software consists of a Postmaster server, a Farmer server, and a Stream server. All servers run on beauty7.
3. Job Structure: There are three phases for a job, denoted Phase1, Phase2, and Phase3.
   (a) Phase1:
      i. Allocate an input tape drive on the I/O nodes.
      ii. Mount a tape on the tape drive.
      iii. Read data from the input tape drive, splitting the data into 10 worker files, on a disk mounted on the I/O node.
   (b) Phase2:
      i. Copy a file from the input stage on the I/O node to the local disk on the PC.
      ii. Set up the environment needed by the E871 analysis code.
      iii. Run the analysis code.
      iv. Copy the stream output to the appropriate staging disks on the I/O node.
   (c) Phase3:
      i. Cleanup the job.
4. Database: A PostgreSQL Database is used on the farm to store job information, log files, and accumulate statistics. There are three main databases: e871farmdb, e871tapedb, and phase2sumdb. e871farmdb contains the detailed messages of a job for the three phases: for example, the time the job was submitted, when it started on which I/O node, the tape label for this job, and 'fail' or 'succeed' for each phase. e871tapedb contains the tables of input tapes and output tapes. Database phase2sumdb contains a summary for each job submitted.
5. Farm Monitoring Web Page: A web page is used to monitor the farm status, allowing non-expert farmers to monitor the processes. To build this web page, Perl CGI scripts were written to grab the job information from the databases and to check the status of the servers. Then the scripts display messages on the web page. This self-refreshing web page displays the status of all servers, detailed messages about stages on I/O nodes, the history and status of submitted and staged jobs, and statistics from the output. If a job fails, users will be told what the error was and how to resubmit the job. The e-log book on the web page enables a farmer to record the important farm messages.

## 3 Data Processing

After processing began and became stable, non-expert shifters took farm shifts and actually submitted jobs and monitored the farm status. The experts generally only needed to debug and fix any problems which occurred during the farming that a shifter couldn't fix. The web page-based interface enabled shifters to take shifts from off site. Once scheduled by a shifter on beauty7, all the jobs were automatically submitted. The only manual task was that operators at the Fermilab Computing Division had to mount and dismount tapes from tape drives.

1. Tasks for Shifters: Schedule a job for submission. Monitor the job status through the web page. If a job fails, a shifter determines the error from the information on the web page. If the job is re-submittable, resubmit the job by typing a resubmit command following the instructions on the web page. A job may fail several times due to a bad tape drive or network-busy although the tape is still good. All jobs that need to be resubmitted stay in a resubmit-table and wait for resubmittal.
2. Tasks for Experts: Farm experts resolve any problems that shifters are not able to handle. They also have to check whether the output written onto the output tapes is consistent with the database. Any difference between output tapes and the database may cause errors in interpreting the results.

## 4 Farm Results

The official data processing started in June 2000 and completely finished in May 2001. A total of 427 jobs failed from 29,843 jobs submitted (1.4%) due to bad raw data tapes. About 36 million histograms were generated on a spill-by-spill level and on a run level. The output DST data are stored on 333 50-GB AIT2 tapes for the primary output stream, 638 5-GB Exabyte tapes were written for auxiliary output streams. The followings are two highlights from farming.

1. Statistics: Figure 4 shows the statistics of all jobs and the number of $\Xi$s and $\Omega$s reconstructed. 'Pol. = $-$' refers to particles while 'Pol. = $+$' refers to anti-particles. Most of the failed jobs occurred in Phase 1 which means the raw input tapes could not be read and split.

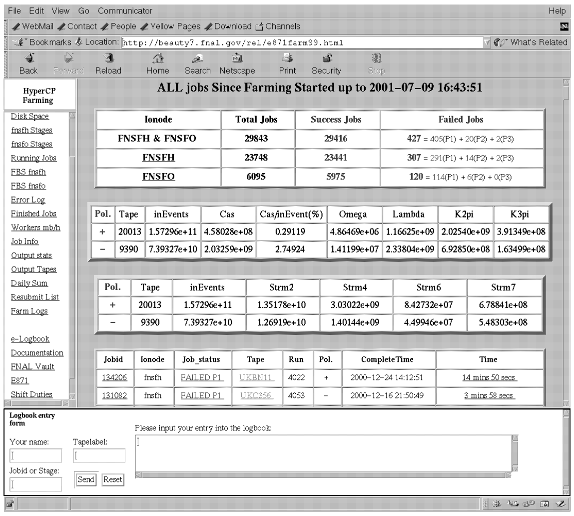


Figure 4: Farm statistics.

2. Histograms from the Farm: Figure 5 shows the mass spectrum of $\Xi^-$ and $\bar{\Xi}^+$ as an example of one of the histograms accumulated directly from the farm. The circle histogram is for $\bar{\Xi}^+$.

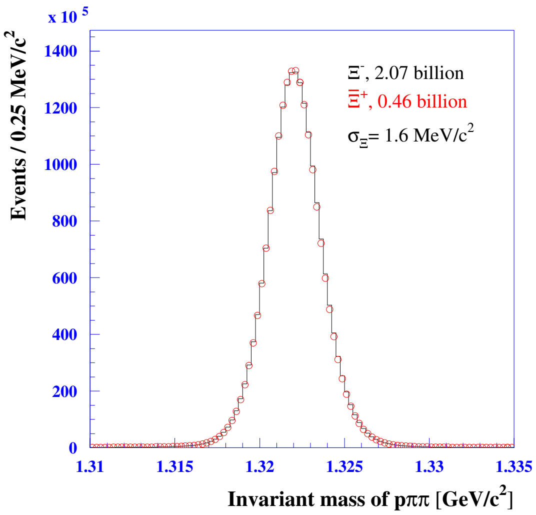


Figure 5: Mass distribution of $\Xi$ and $\bar{\Xi}^+$.

## 5 Conclusion

The powerful HyperCP farm has successfully processed all the raw data taken in both the 1997 and 1999 runs in only eleven months. The results indicate that the HyperCP farm is very efficient, well-organized, and easy to monitor. All the log information from the farm processing is easily accessible. The reconstructed DST data and 36 million histograms are now ready for the *CP*-violation study as well as a number of other physics studies.

## 6 Acknowledgment

We are grateful to the Fermilab Computing Division for their support of the farm.

## References

[1] See `http://www-isd.fnal.gov/fbs/`.
[2] J.H. Christenson, J.W. Cronin, V.L. Fitch, R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
[3] KTeV Collaboration (A. Alavi-Harati *et al.*), Phys. Rev. Lett. **83**, 22 (1999).
[4] NA48 Collaboration (V. Fanti *et al.*), Phys. Lett. B **465**, 335 (1999).
[5] S. Pakvasa, Proc. Third Internat. Conf. on B Physics and CP Violation, ed. by H.-Y. Cheng and W.-S. Hou, World Scientific, 2000, p. 423.
[6] R. Fleischer, hep-ph/0011323.
[7] For further information, see `http://ppd.fnal.gov/experiments/e871/public/documentation.html`.