Offline Data Production

Du Zhizhen, Han Ying, He Ju, Hu Qionghui,
Li Weihua, *Liu Huaimin*, Shao Yuying, Sun Yongzhao, Xiao Xiao,
Zhang Changchun, Zhang Shaoqiang
Data Production Group

June 7, 2000

- Offline Computing Environment
- Data Production Processes
- Data Reconstruction
- Data Processed
- Summary

1 Offline Computing Environment (HP)

- A new HP workstation (HPWS9) came into use since Feb., Y2K, eased the tension of data processing. (See figure)

  Total CPU: ~100 SPECint95

  Hard Disk: ~250 GB

- CPU power & HD space: doubled compared with that for R exp. in 99
- Data reconstruction speed: ~80 runs/day (24 hrs)
- Online data-taking speed: ~40 runs/day (24 hrs)

2 Data Production Processes (see figure)

1. Data Splitting from Online Tape
   RAW (event data), LUM (lum. monitor), ODB (online data base)
2. Bhabha & Dimuon Selection for Calibration
   BB (Bhabha), MU (Dimuon), RB (radiative BB), NU (Endcap BB)

3. Data Quality Check from RAW (see figure)
   Offline feedback for detector performance

4. Data Reconstruction from RAW
   No filter processes any more

5. DST Files Production
   Compressed from REC for physics analysis

6. Pronged Data Sorting
   Not available

3 Data Reconstruction

- BES software version: V102 (developing, not released then)
  V102 released on Apr. 8,2000
  VC-MDC combined reconstruction: NOT used

- 2 Rounds of Reconstruction Have Been Carried Out
  1st round: run13700-15500 (Jan. 11 – Mar. 22)
  Mainly for V102 check

  2nd round: run13500-17018 (Mar. 31 – May 22)
  Improvements made

  1. T0 of MDC wires re-calibrated, fine tuned for the whole run
  2. BSC re-calibrated (new mothed) run by run
  3. Q-dependent effect corrected for reliable MDC error matrix
  4. bugs found in 1st round fixed

- 32 runs crashed during reconstruction:

  1. Some "too dirty" events in these runs
  2. Imperfection of MDC & VC reconstruction codes (most of the bugs fixed)
  3. Events in these runs are not fully reconstructed

4 Data Processed

- Data Taken in the Spring Run
  Run13500-17018 (from Nov. 11, 99 to May 15, 2K)

  J/Psi, Psi', data@3.0GeV (∼3500 runs)
  Largest data set since 1990 (∼1500 runs for R exp. 99)

Number of online tapes: 267
40 runs missing due to online crash or malfunction

- Data Size (30K events/run)

  RAW ~65 MB
  REC ~250 MB (RAW data included)
  DST ~25 MB (new DST format)

- Data on Disk (HPWS8)
  - for more detail, see http://hpws1/~offline or access from BES Web Site http://hptc4/~beswww

run13500-13699.ndst	/data/d53/offline/ndst/
run13700-14400.ndst	/data/d51/offline/ndst/
run14401-15100.ndst	/data/d52/offline/ndst/
run15101-15180.ndst	/data/d56/offline/ndst/
run15181-16500.ndst	/data/d57/offline/ndst/
run16501-17018.ndst	/data/d53/offline/ndst/

- Data Save

  All kinds of data (RAW, REC, DST, ...) tarred to 8mm tapes
  2 copies each
  Tape I/O speed: ~2.5GB/hr

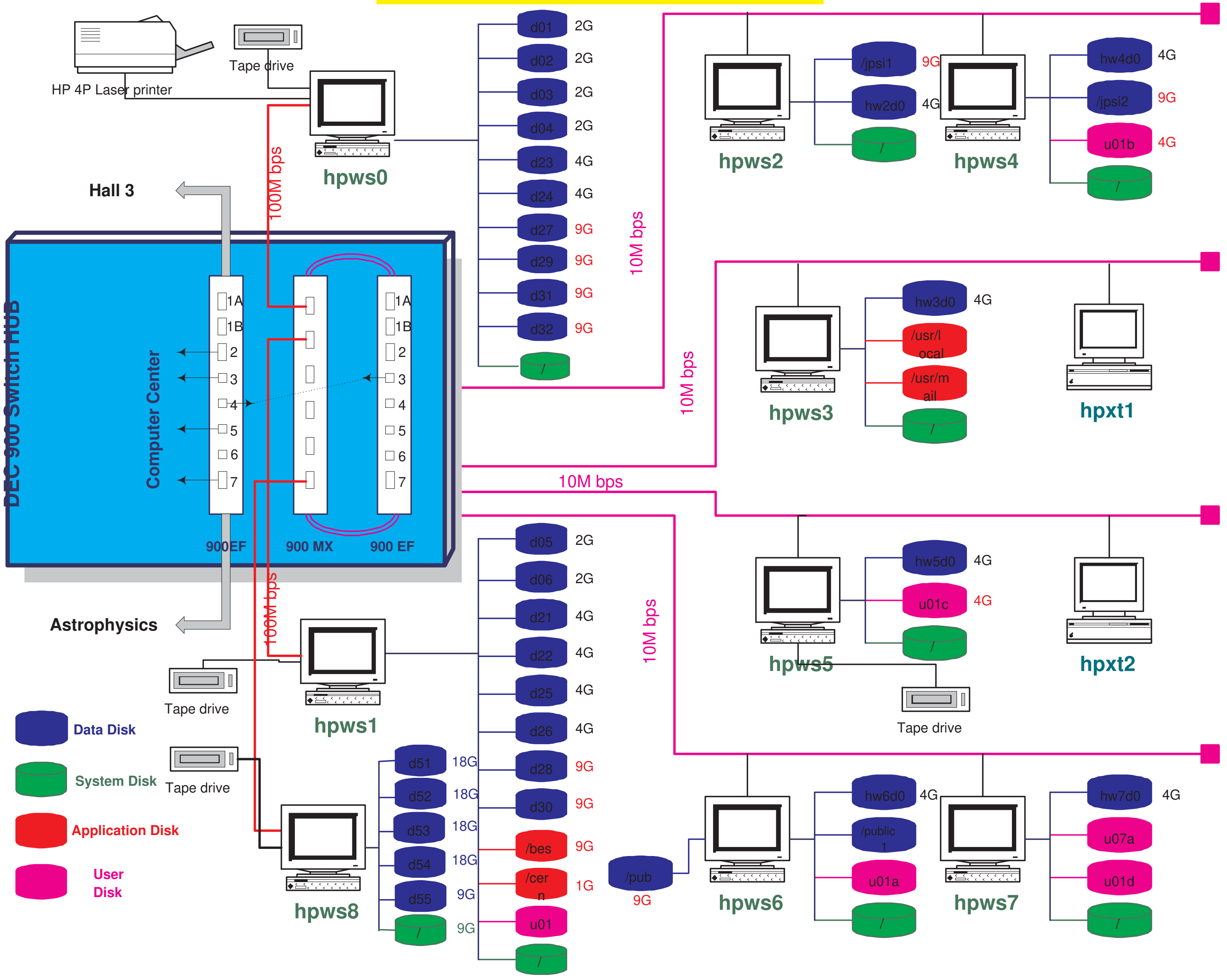
  Number of 8mm tapes (4.5GB/tape):
  RAW $2 \cdot 56$
  REC $2 \cdot (171+86)$ (2nd+1st)
  DST $2 \cdot (18+11)$
  Others $2 \cdot 20$

  Data saved to tape: ~3TB

5 Summary

- Current Computing Power is Capable of Offline Data Processing
- Offline data processing can cope with online data taking
- Raid Disk and Tape Robot Needed for Data Save and Fast I/O
- SAN project proposed by Computer Center under approval

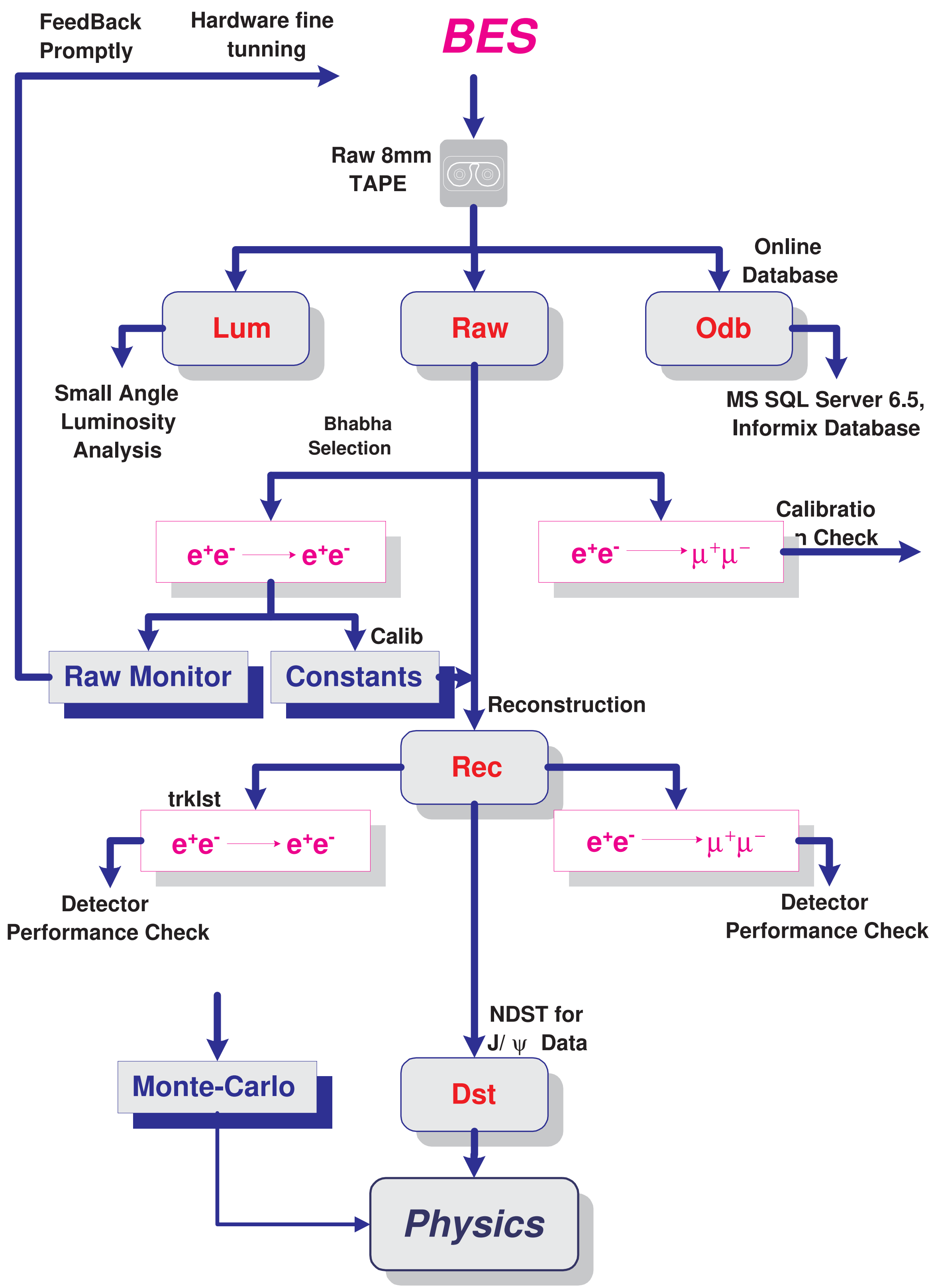
BES HP/Unix Environment



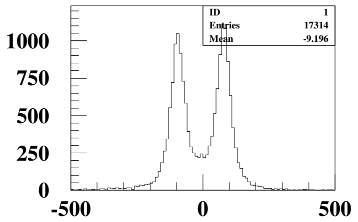
Offline Computing Environment

HOST	TYPE	CPU (SPCINT95)	RAM (MB)	HD (GB)	TAPE DRIVE	NETWORK (FDDI)
HPWS0/1	HP735/ 125	2x4	2x144	2x(4+9)	2	100 Mbps
HPWS8	HP9000/ J2240	34 (dual)	1024	6x18	1	100 Mbps
HPWS9	HP9000/ J5000	$\sim$60 (dual)	2048	7x18	2	10 Mbps Ethernet

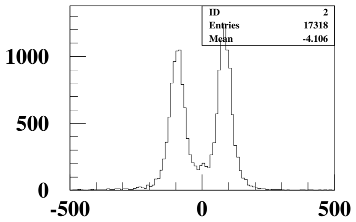
```
*******************************************
Run #      Total Good   Good Rate  Hadron
-----------------------------------------

 RUN  15331     12497      41.7%      62.5%

KEPT HADRON   =>       7807   26.0%
KEPT BHABHA   =>        924    3.1%
KEPT R BHAB   =>        118    0.4%
KEPT MU-MU    =>        552    1.8%
KEPT NEUTRL   =>       3096   10.3%
TESTING EVENT =>      29998  100.0%
COSMIC REJECT =>        426    1.4%
GARBAG REJECT =>        210    0.7%
UNKNOWN       =>      16865   56.2%
-------------------------------------------
TOTAL EVENTS  =>      29998

*******************************************
Run #      Total Good   Good Rate  Hadron
-----------------------------------------

 RUN  15332     12462      41.5%      61.8%

KEPT HADRON   =>       7705   25.7%
KEPT BHABHA   =>        917    3.1%
KEPT R BHAB   =>        126    0.4%
KEPT MU-MU    =>        597    2.0%
KEPT NEUTRL   =>       3117   10.4%
TESTING EVENT =>      29993  100.0%
COSMIC REJECT =>        414    1.4%
GARBAG REJECT =>        198    0.7%
UNKNOWN       =>      16919   56.4%
-------------------------------------------
TOTAL EVENTS  =>      29993

*******************************************
Run #      Total Good   Good Rate  Hadron
-----------------------------------------
 RUN  15334     11822      39.4%      62.2%

KEPT HADRON   =>       7358   24.5%
KEPT BHABHA   =>        816    2.7%
KEPT R BHAB   =>        120    0.4%
KEPT MU-MU    =>        634    2.1%
KEPT NEUTRL   =>       2894    9.6%
TESTING EVENT =>      29990  100.0%
COSMIC REJECT =>        589    2.0%
GARBAG REJECT =>        223    0.7%
UNKNOWN       =>      17356   57.9%
-------------------------------------------
TOTAL EVENTS  =>      29990
```


FeedBack
Promptly
Hardware fine
tunning
BES
Raw 8mm
TAPE
Online
Database
Lum
Raw
Odb
Small Angle
Luminosity
Analysis
MS SQL Server 6.5,
Informix Database
Bhabha
Selection
$e^+e^- \rightarrow e^+e^-$
$e^+e^- \rightarrow \mu^+\mu^-$
Calibratio
n Check
Calib
Raw Monitor
Constants
Reconstruction
Rec
trklst
$e^+e^- \rightarrow e^+e^-$
$e^+e^- \rightarrow \mu^+\mu^-$
Detector
Performance Check
Detector
Performance Check
NDST for
J/ψ Data
Monte-Carlo
Dst
Physics

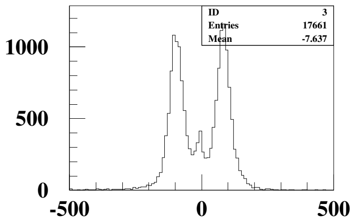
bb15516-15525.hbk



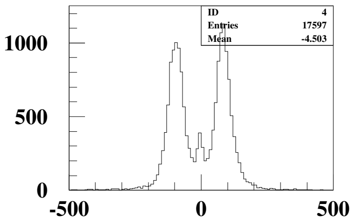
MDC-DI-PEAK-1



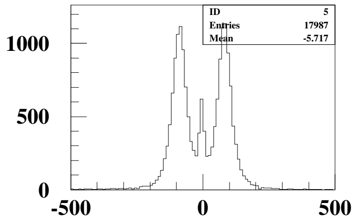
MDC-DI-PEAK-2



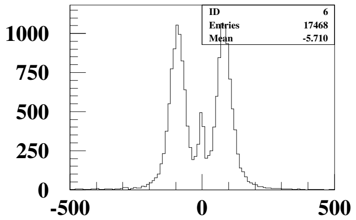
MDC-DI-PEAK-3



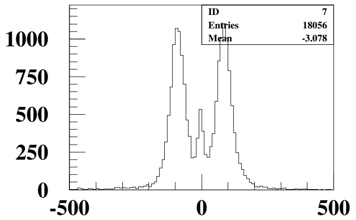
MDC-DI-PEAK-4



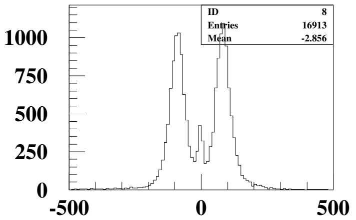
MDC-DI-PEAK-5



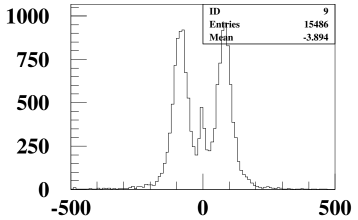
MDC-DI-PEAK-6



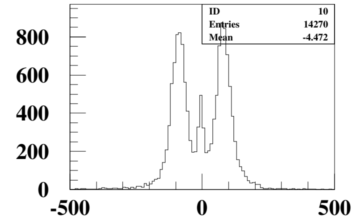
MDC-DI-PEAK-7



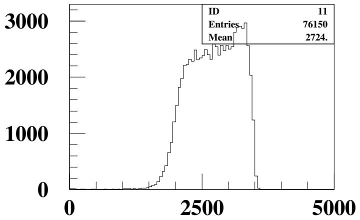
MDC-DI-PEAK-8



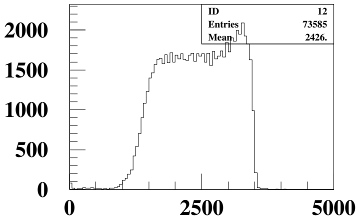
MDC-DI-PEAK-9



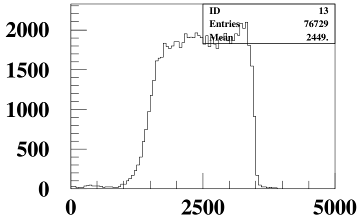
MDC-DI-PEAK-10



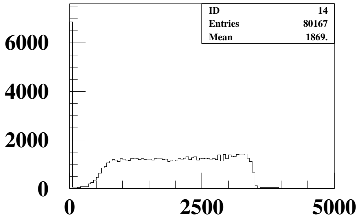
MDC-TIME-1



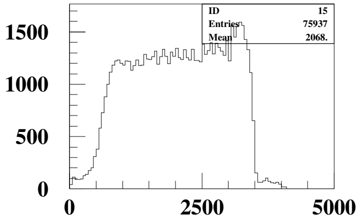
MDC-TIME-2



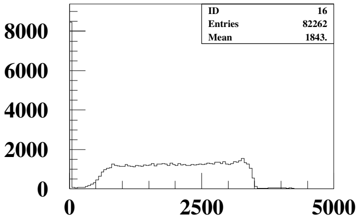
MDC-TIME-3



MDC-TIME-4



MDC-TIME-5



MDC-TIME-6