The CDF Run 2 Offline Computer Farms

Stephen Wolbers

Fermilab

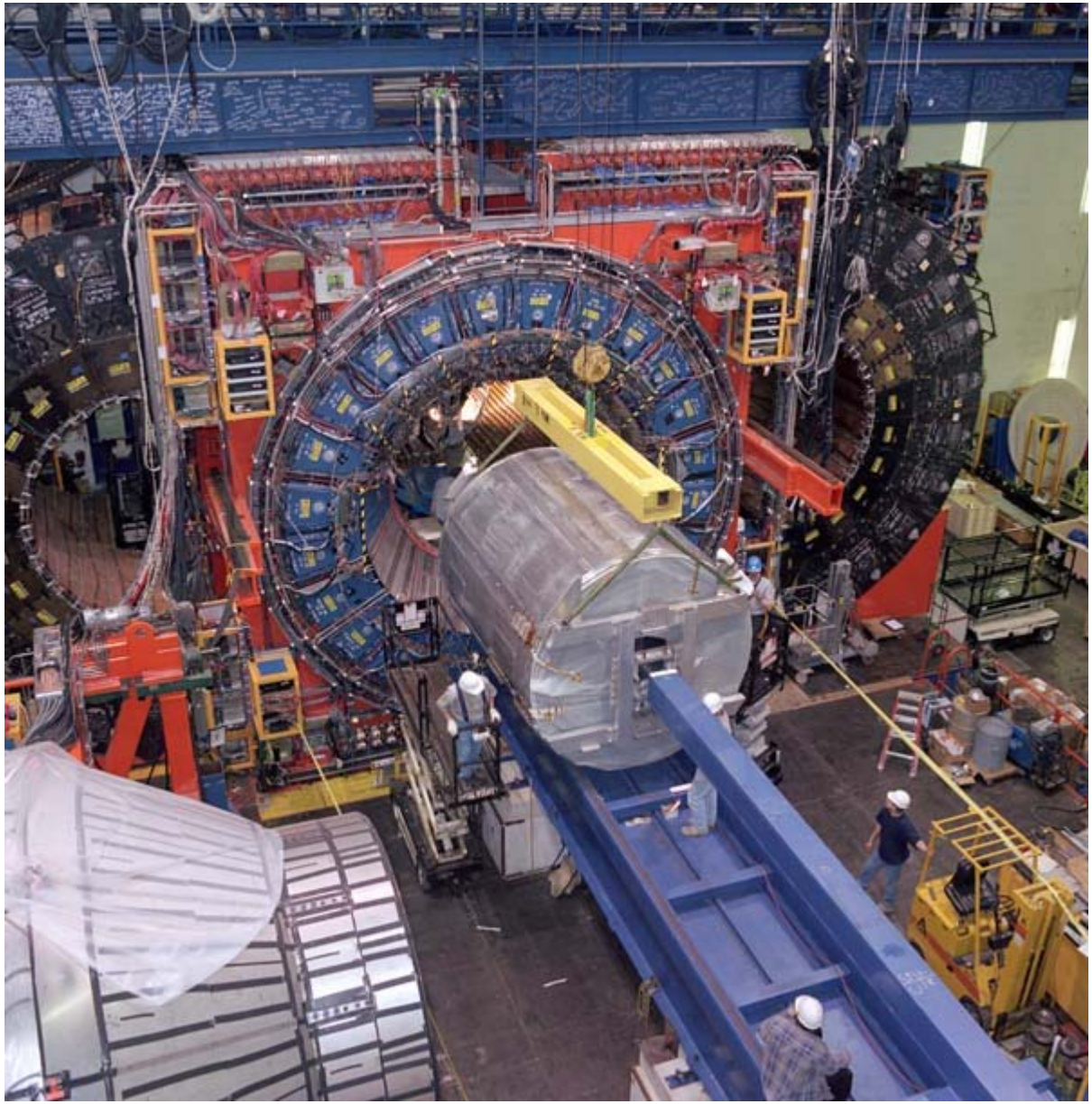
September 3, 2001

Outline

- **Introduction to Run 2 Data Rates/Processing Needs**
- **Architecture of the CDF Run 2 Farms**
- **Experience with the farms in Run2**
- **Future**
  - Run 2a
  - Run 2b

Introduction

- **CDF Run 2 Data Rates are substantially larger than Run 1 (factor 20 higher).**
  - 20 Mbyte/sec to tape peak
  - Well over 100 Tbyte/year to tape
- **This data must be processed as quickly as it is collected (with a short time delay for preparing constants or code).**
- **In addition, reprocessing and simulation needs are also required.**



CDF Run 2a Farm Computing

- **CPU for event reconstruction of about 5 sec/event on a PIII/500 MHz PC (Each event is 250 Kbyte).**
- **Assuming 20 Mbyte/sec peak (approx. 75 Hz)**
  - Requires 375 PIII/500 processors to keep up
  - Faster machines -> Fewer processors required
- **Requirement is degraded by accelerator/detector efficiency and upgraded by farms efficiency (or inefficiency)**

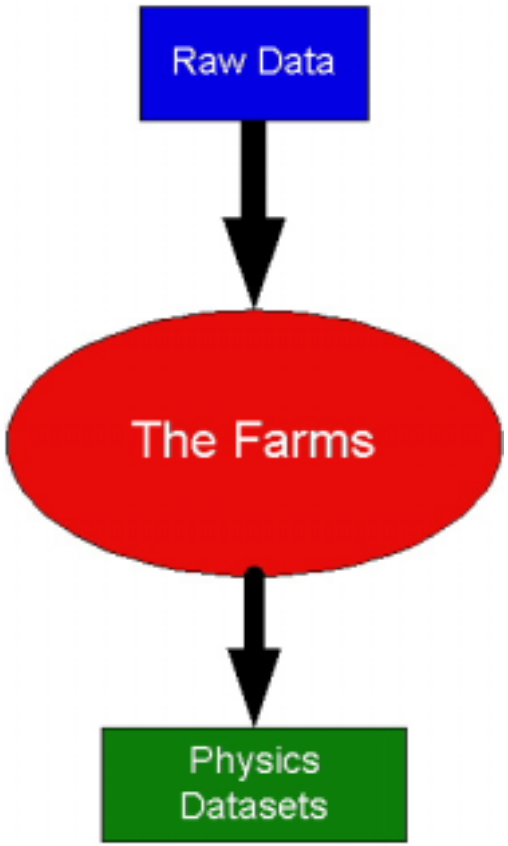
RunII: Data Volume

Category	Parameter	D0	CDF
DAQ rates	Peak rate	53 Hz	75 Hz
	Avg. evt. Size	250 KB	250 KB
	Level 2 output	1000 Hz	300 Hz
	maximum log rate	Scalable	80 MB/s
Data storage	# of events	600M/year	900 M/year
	RAW data	150 TB/year	250 TB/year
	Reconstructed data tier	75 TB/year	135 TB/year
	Physics analysis summary tier	50 TB/year	79 TB/year
	Micro summary	3 TB/year	-
CPU	Reconstr/event	25 - 65 SI95xsec	30 SI95xsec
	Total Reconstruction	2000-4000 SI95	2000-4000 SI95
	Analysis	2000-4000 SI95	2000-4000 SI95
Access for analysis	# of scientists	400 - 500	400 - 500

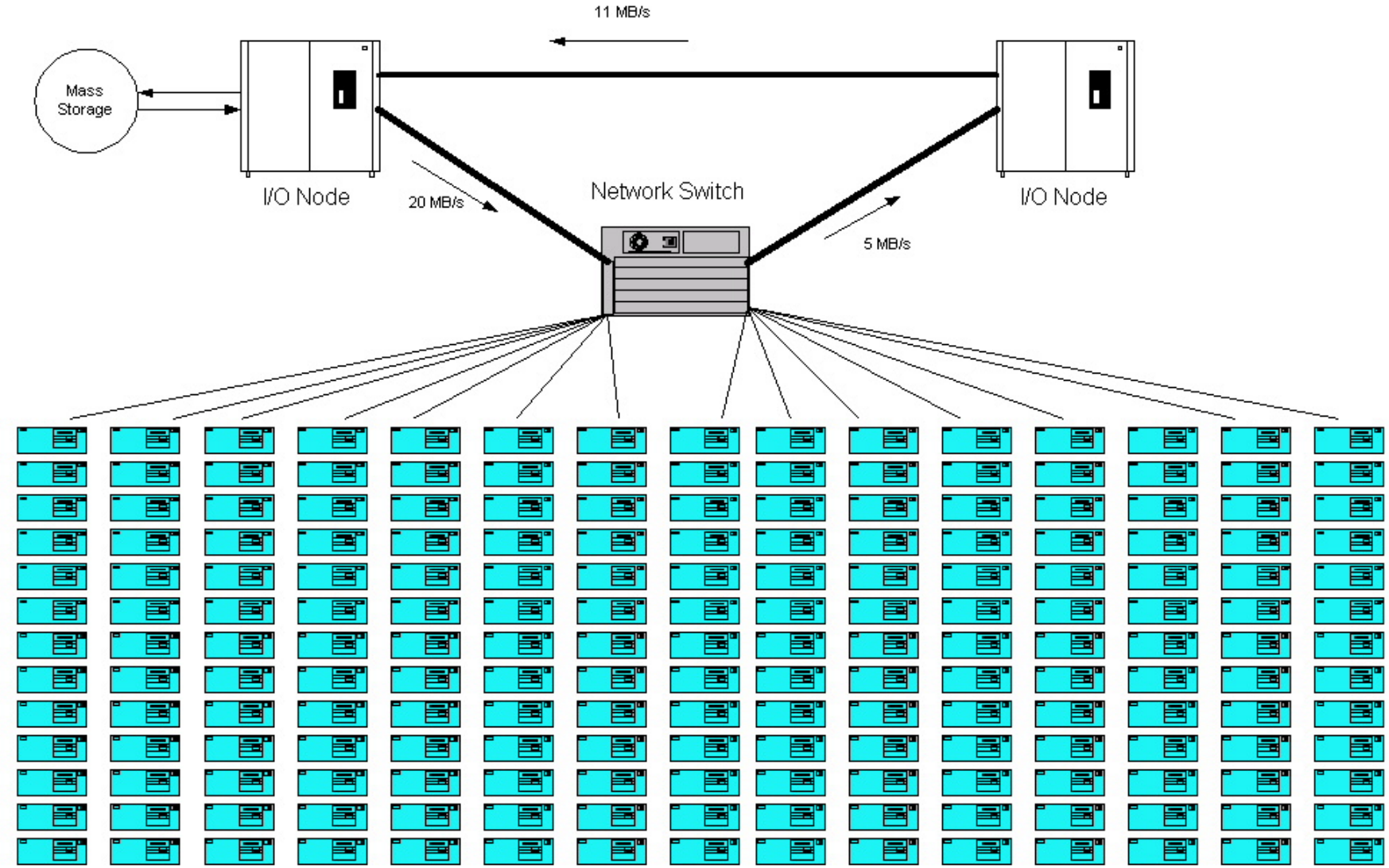
CDF/D0 Offline Production Farms for event reconstruction

- **The CDF/D0 farms must have sufficient capacity for Run 2 Raw Data Reconstruction.**
- **The farms also must provide capacity for any reprocessing needs.**
- **Farms must be easy to configure and run.**
- **The bookkeeping must be clear and easy to use**
- **Error handling must be excellent.**

Simple Model



Run II CDF PC Farm



Design/Model

- **Hardware**
  - Choose the most cost-effective CPU's for the compute-intensive computing.
  - This is currently the dual-Pentium architecture
  - Network is fast and gigabit ethernet, with all machines being connected to a single or at most two large switches.
  - A large I/O system to handle the buffering of data to/from mass storage and to provide a place to split the data into physics datasets.

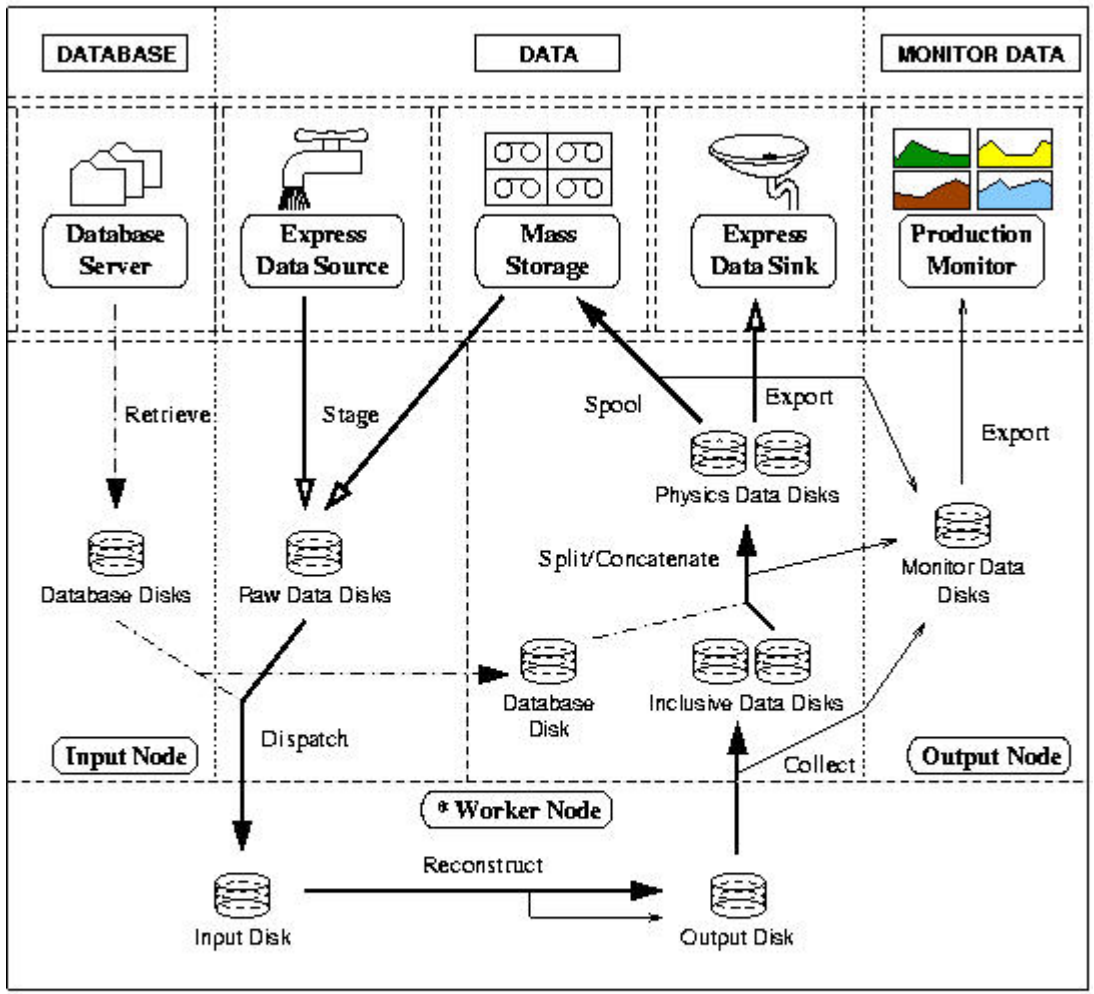

EXIT>

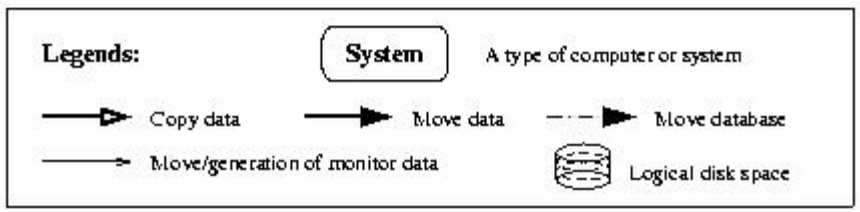

1/24/2001

Software Model

- **Software consists of independent modules**
  - Well defined interfaces
  - Common bookkeeping
  - Standardized error handling
- **Choices**
  - Python
  - MySQL database (internal database)
  - FBSNG (Farms Batch System)
  - FIPC (Farms Interprocessor Communication)
  - CDF Data Handling Software

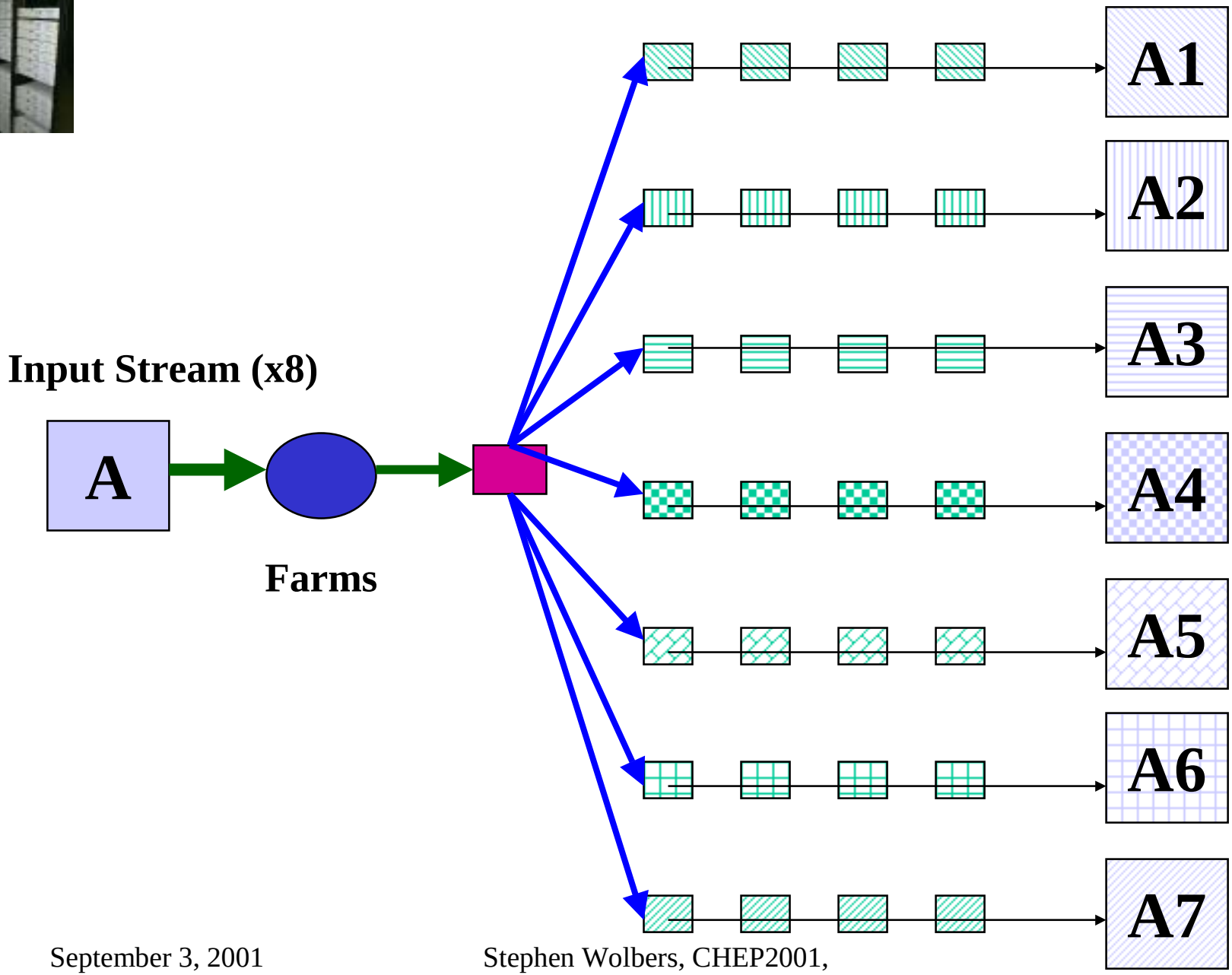
Conceptual Model of Run 2 Production System





Physics Analysis Requirements and Impact

- **Raw Data Files come in ~8 flavors, or streams**
  - 1 Gbyte input files
- **Reconstruction produces inclusive summary files**
  - 250 Mbyte output files
- **Output Files must be split into ~8 physics datasets per input stream**
  - Target 1 Gbyte files
  - About 20% overlap
- **Leads to a complicated splitting/concatenation problem, as input and output streams range from tiny (<few percent) to quite large (10's of percent)**


Input Stream (x8)
A
Farms
A1
A2
A3
A4
A5
A6
A7

Status of CDF Farms

- **154 PC's are in place.**
  - 50 PIII/500 duals
  - 40 PIII/800 duals
  - 64 PIII/1 GHz duals
- **I/O nodes are ready.**
- **The CDF Data Handling System has sufficient capacity to handle the I/O to/from tape.**

Experiences so far in Run 2

Early Processing Experience

- **Commissioning Run (October, 2000)**
  - Tevatron Collider luminosity was small
  - CDF detector was not yet complete
  - Data was taken and processed on the farms

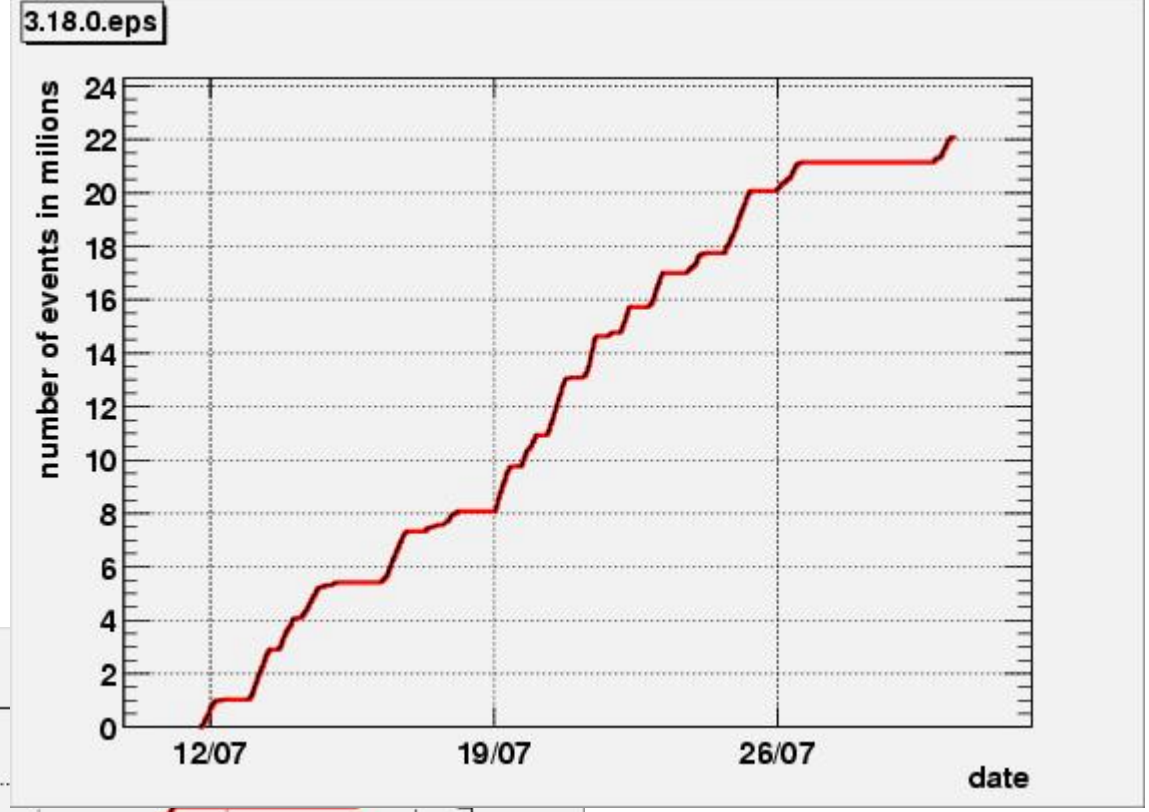
Early Processing Experience

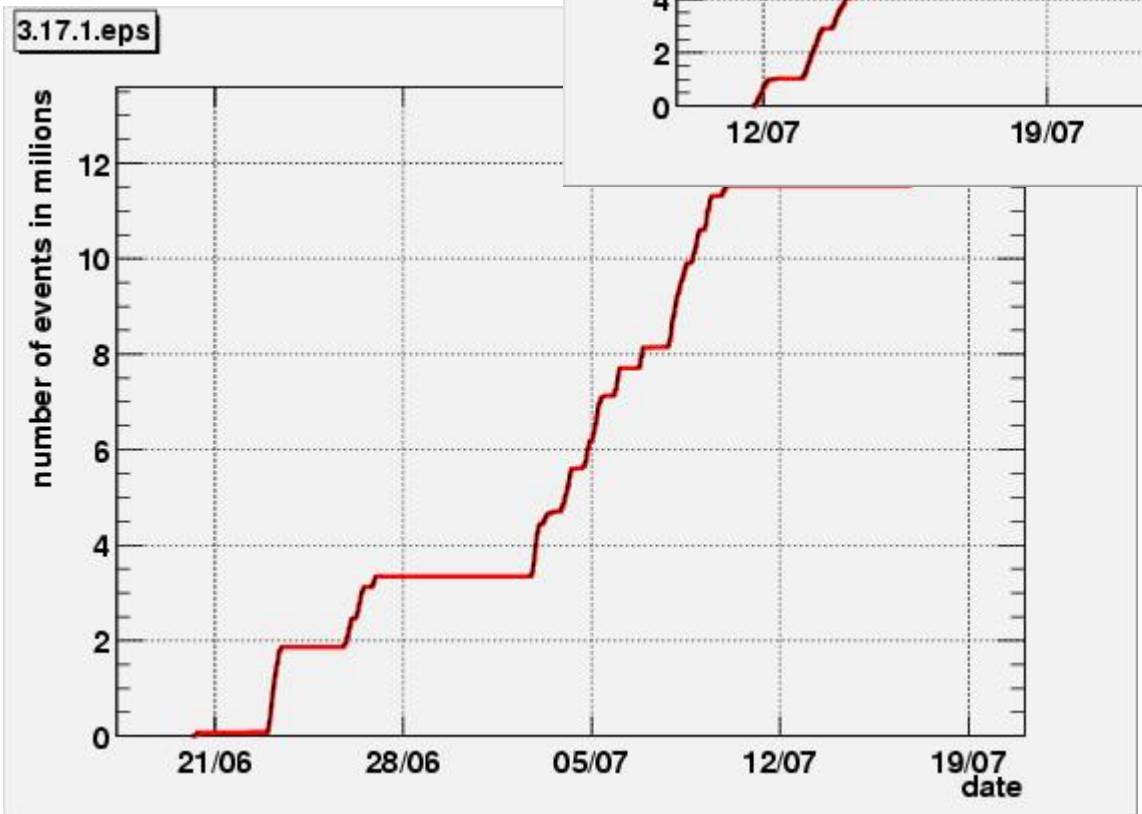
- **April 2001 Data**
  - First 36x36 bunch collisions
  - Most of CDF detector was complete
  - Small amount of data, additional experience

June-July 2001 Data

- **First substantial data taken in Run 2**
  - Approximately 34 million events (“good runs”)
  - Processed with two versions of the code
  - Long downtime (unplanned) allowed the farms to catch up with the backlog of data

June/July Processing





Run 2b at Fermilab

- **Run 2b will start in 2004 and will increase the integrated luminosity to CDF and D0 by a factor of approximately 8 (or more if possible).**
- **It is likely that the computing required will increase by the same factor, in order to pursue the physics topics of interest:**
  - B physics
  - Electroweak
  - Top
  - Higgs
  - Supersymmetry
  - QCD
  - Etc.

Run 2b Computing

- **Current estimates for Run 2b computing:**
  - 8x CPU, disk, tape storage.
  - Expected cost is same as Run 2a because of increased price/performance of CPU, disk, tape.
  - Plans for R&D testing, upgrades/acquisitions will start next year.
- **Data-taking rate:**
  - May be as large as 80 Mbyte/s.
  - About 1 Petabyte/year to storage.

Summary

- **CDF Production Farms are commissioned, tested and have processed tens of millions of events.**
- **Run2a will be a major task for the farms.**
- **Run2b is potentially substantially larger than Run2a, and some changes to the farms will likely be needed to address this.**