Software, Computing and Analysis Models at CDF and D0

Donatella Lucchesi
CDF experiment INFN-Padova

Outline
Introduction
CDF and D0 Computing Model
GRID Migration
Summary

III Workshop Italiano sulla
fisica di Atlas e CMS
Bari, 20-22 ottobre 2005

Introduction: a thought

"A complex system that works is found to have evolved from a simple system that worked…
A complex system designed from scratch never works and cannot be patched up to make it work. You have to start over, beginning with a working simple system"

G.Booch OO Analysis & Design 2nd ed. Pag. 13

CDF and D0 can be such simple system to start with

Introduction: Some Numbers

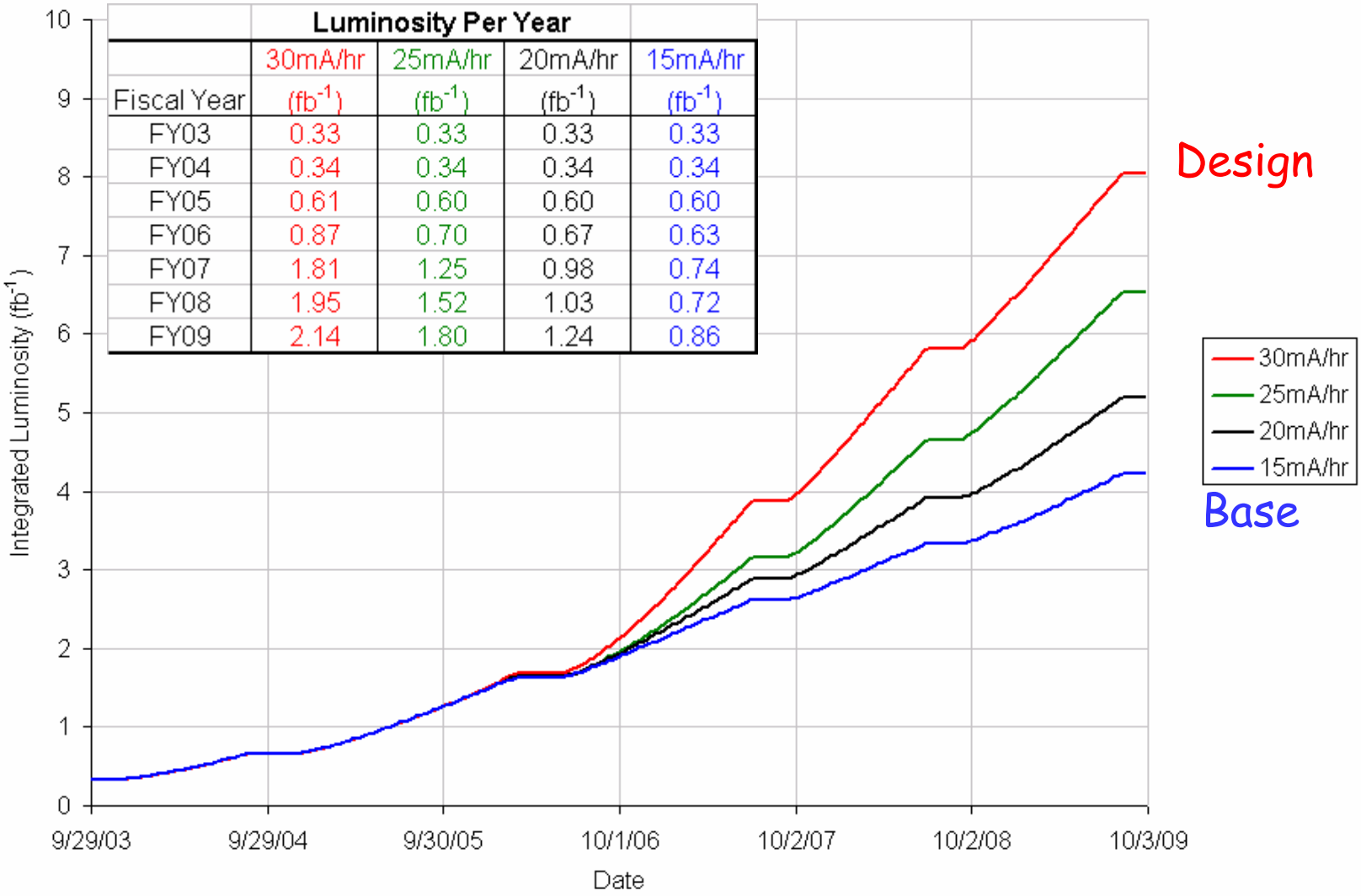
	CDF	D0
Raw data size* (Kbytes/event)	150	250-300
Reco data size (Kbytes/event)	120	200
User format (Kbytes/event)	25-180	20-40
Reco time** (GHz-sec/event)	5(10)	50(120)
User analysis time (GHz-sec/event)	1(3)	1
Peak data rate (Hz)	130(360)	50(100)
Persistent data format	RootIO	TMB,Root

*Raw event size depends upon trigger type and luminosity

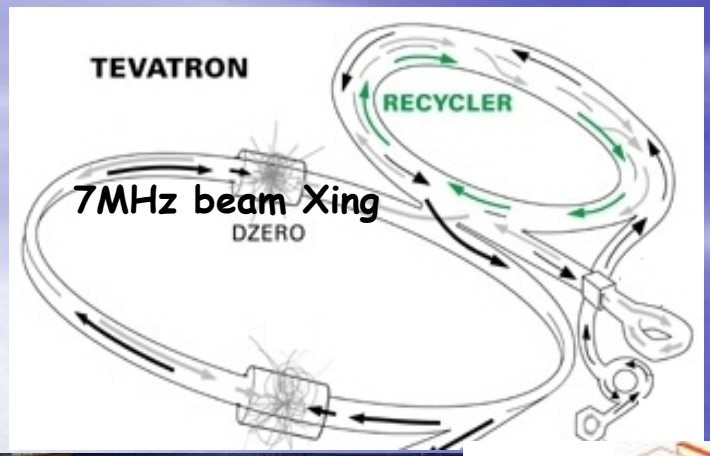
**Reconstruction time depends upon raw data size

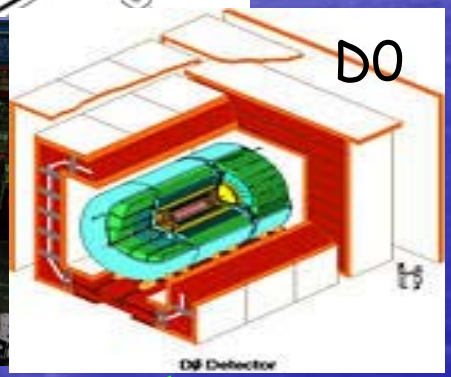
Introduction: Integrated Luminosity

Luminosity Per Year				
	30mA/hr	25mA/hr	20mA/hr	15mA/hr
Fiscal Year	(fb^{-1})	(fb^{-1})	(fb^{-1})	(fb^{-1})
FY03	0.33	0.33	0.33	0.33
FY04	0.34	0.34	0.34	0.34
FY05	0.61	0.60	0.60	0.60
FY06	0.87	0.70	0.67	0.63
FY07	1.81	1.25	0.98	0.74
FY08	1.95	1.52	1.03	0.72
FY09	2.14	1.80	1.24	0.86



CDF & D0 Computing Model: Data Production flow

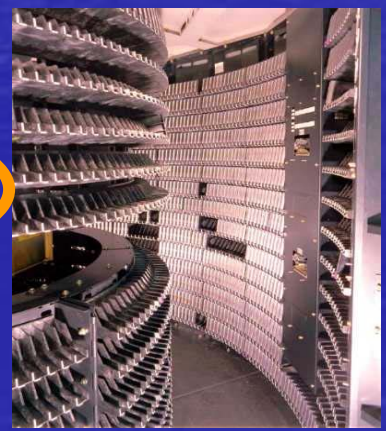




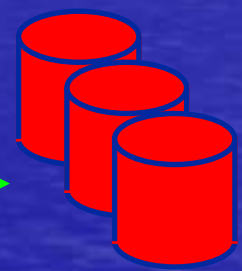
Production Farm

DH

DH



DH



Disk cache

Level 3 Trigger
CDF~100 Hz
D0~50 Hz

Data Handling Services
DH

Robotic Tape Storage

Data Handling Services

SAM: Sequential Access via Metadata

CDF: used dCache then moved to SAM.
D0: use SAM since several years
SAM provides:

- Local and wide area data transport
- Batch adapter support (PBS, Condor, lsf, ..)
- Caching layer
- Comprehensive meta-data to describe collider and Monte Carlo data
  - Simple tools to define user-datasets
- File tracking information
  - Location, delivery and "consumption" status
- Process automation

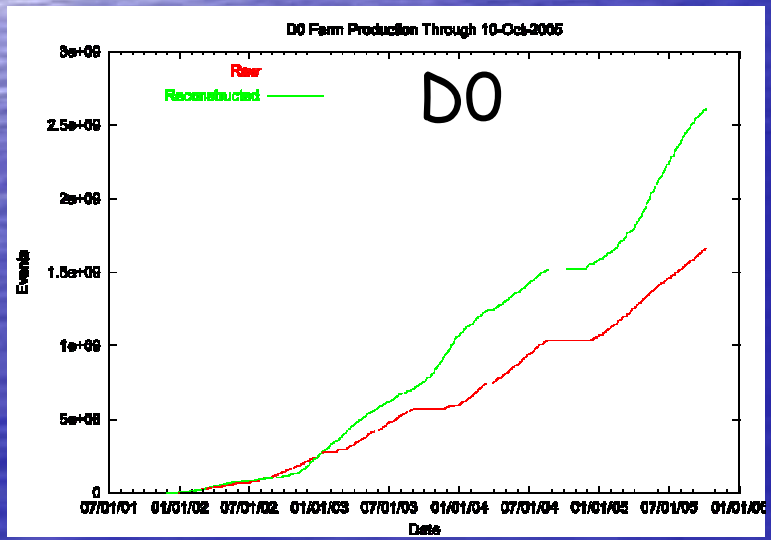
Production Farms

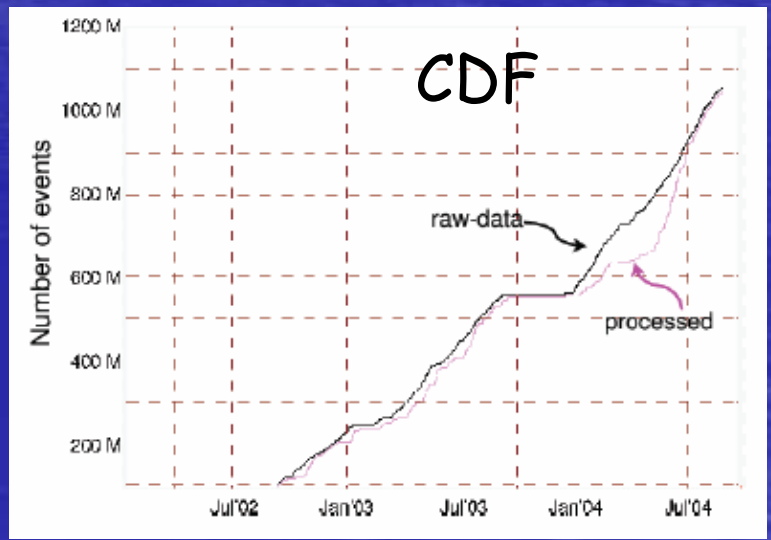
D0:

Data: compute capacity 1550 GHz ⇒ ~2390 GHz soon
Efficiency ~80% ~24 M events/week
MC: 14M events/month ⇒ 2 M events if data reprocessing

CDF:

New SAM based farm, standard CDF code
Data: compute capacity 1200 GHz
~78 M events/week





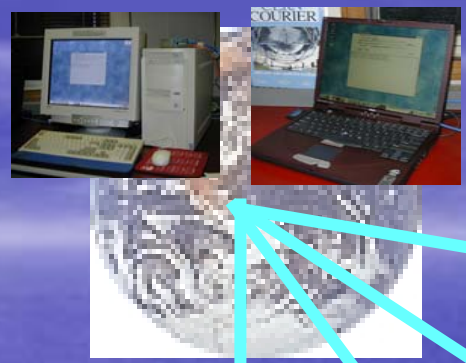
CDF & DO Computing Model: Analysis Data flow

Robotic
Tape Storage

DH

Disk
cache

DH



User Desktop

DH

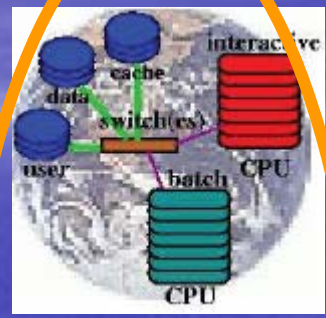
DH

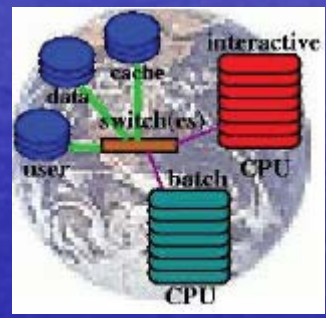
DH

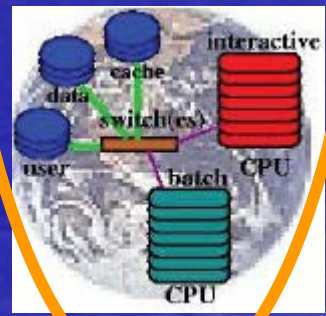
DH

Central Analysis System

Remote Analysis System
CDF=CAFs DO=Remote Farms







(CDF=Central Analysis Facility CAF, DO=Central Analysis Backend,Clued0)

Central Analysis System

CDF: CAF

- Primary analysis platform
  - User analysis: Ntuple creation and analysis
  - Semi-coordinated activities: secondary and tertiary dataset, Monte Carlo Production
- User experience:
  - Monitoring
  - Control: hold, resume job, copy output to any machine
  - Quasi-interactive feature: Look a log file on worker node

D0: CAB

- CPU intensive analysis job
- Direct analysis
- Fixing and skimming
- Group organized translation to root format

D0: clued0

- Desktop cluster for interactive work and user analysis

Remote Analysis System

Job Information Monitoring

D0

- Farms based upon SAMGrid =SAM+JIM
- SAM: file storage, delivery and metadata cataloging, analysis bookkeeping.
- JIM : Job Manager (D0 specific installations: SAM station, DB proxy servers, job manager)
- 10 remote sites in operation:

CCIN2P3 (Lyon) , CMS-FNAL, FZU (Prague), GridKa (Karlsruhe), Imperial OSCER (Oklahoma), SPRACE (Sao Paolo), UTA (Texas, Arlington), WestGrid (Canada), Wisconsin

- Monte Carlo production and data reprocessing

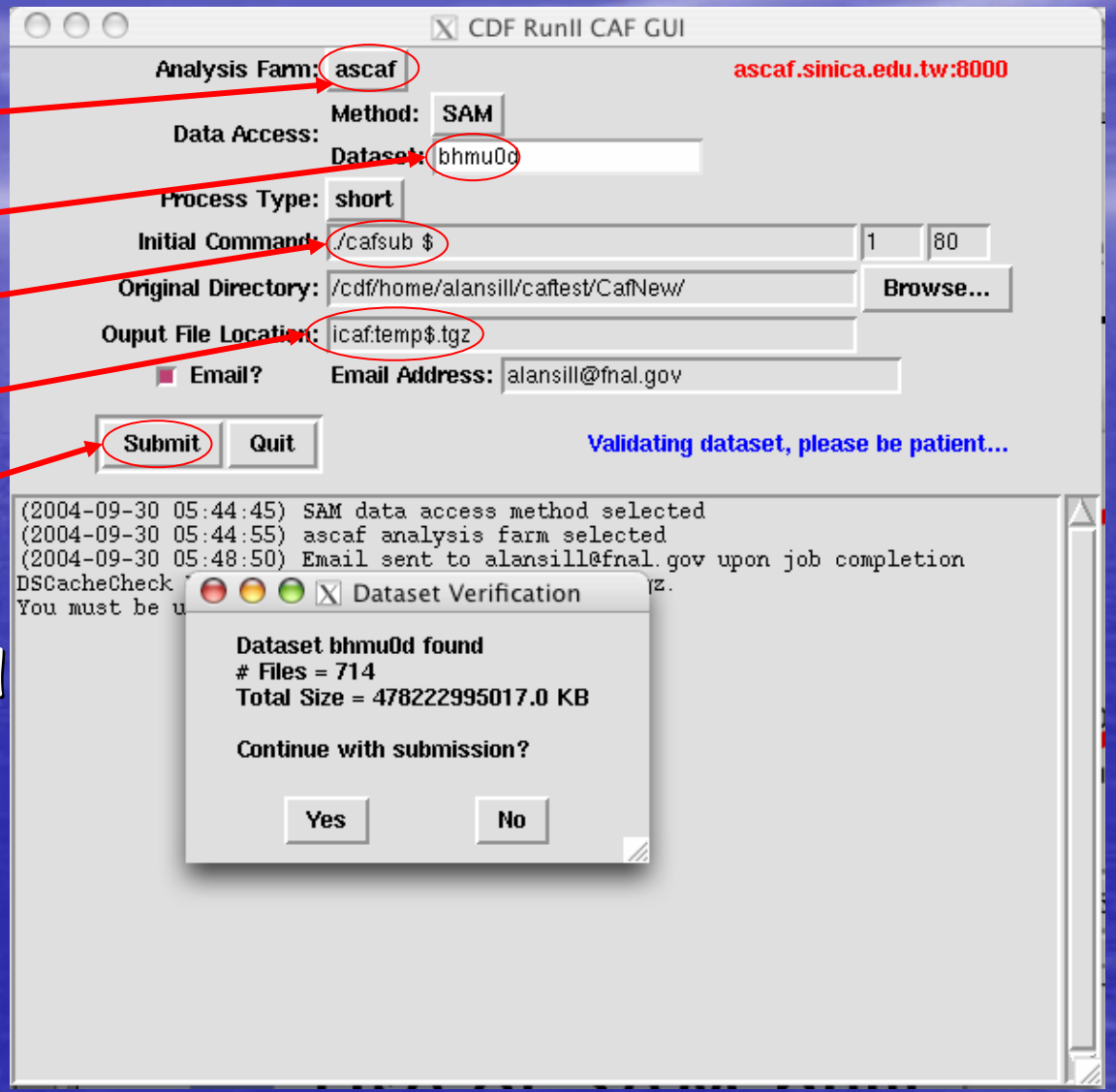
CDF

- dCAF a replica of CAF (specific CDF installation)
- Monte Carlo production and now data analysis

CDF Submission GUI

User's experience

- Select site
- Specify dataset
- Startup script
- Output location
- Press "submit"



User's context tarballed sent to execution site

Same interface used for GRID submission

CDF & D0 Migration to GRID

Reason to move to a grid computing model

- Need to expand resources luminosity expect to increase by factor 8
- Resource are in dedicated pools, limited expansion and need to be maintained
- Resource at large:
  - in Italy access to ~3X dedicated resources
  - ~30 THz in USA fro GRID

CDF has three projects:

✓ GlideCAF: MC production on GRID, data processing at T1

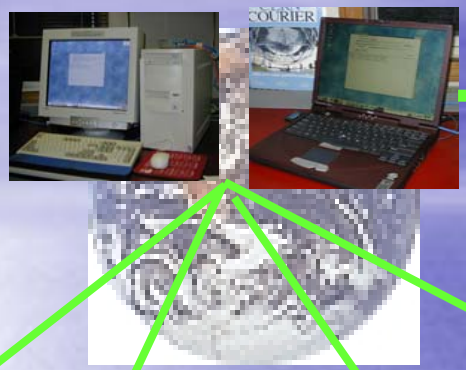
➢ gLiteCAF: MC production on GRID

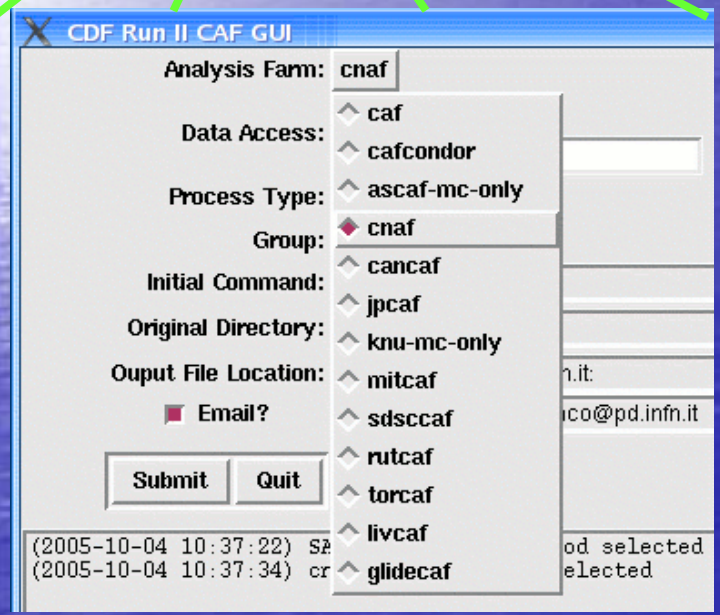
➢ OSGCAF: " (USA)

D0:

➢ SAMGrid: MC production

CDF Migration to GRID: GlideCAF





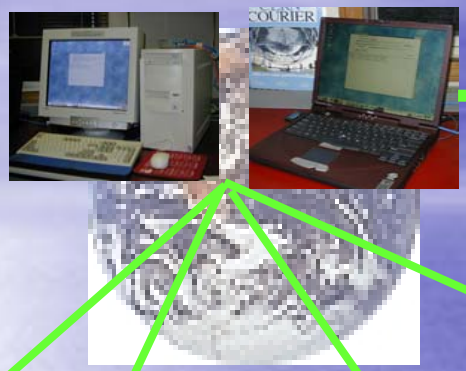
Condor protocol

Dedicated resources

Gatekeeper

Dynamic resources

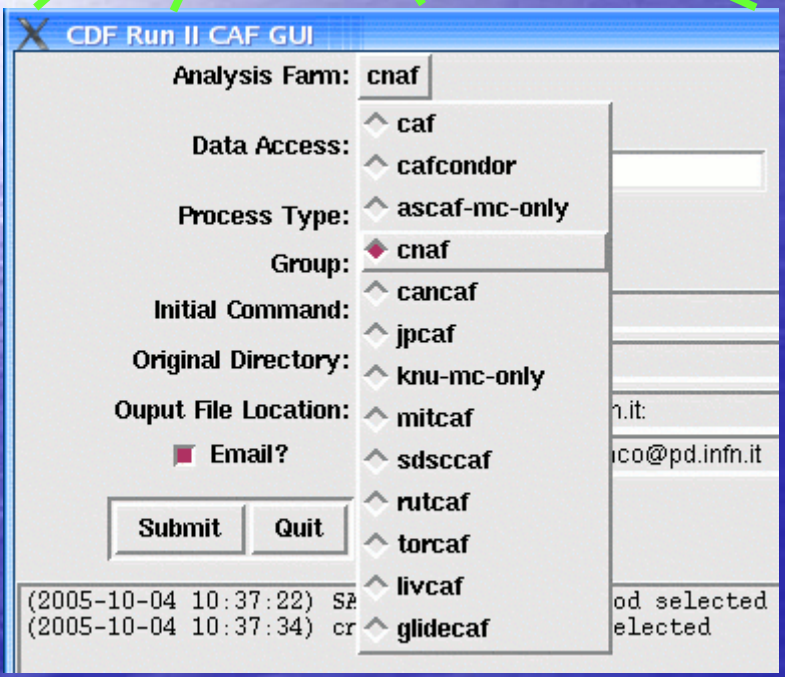
CDF Migration to GRID: gLiteCAF

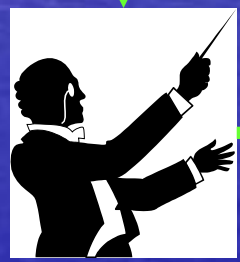


Secure connection

via Kerberos

Head node ≅
User Interface

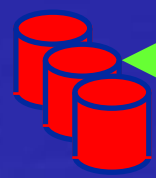




Resource Broker

Gatekeeper

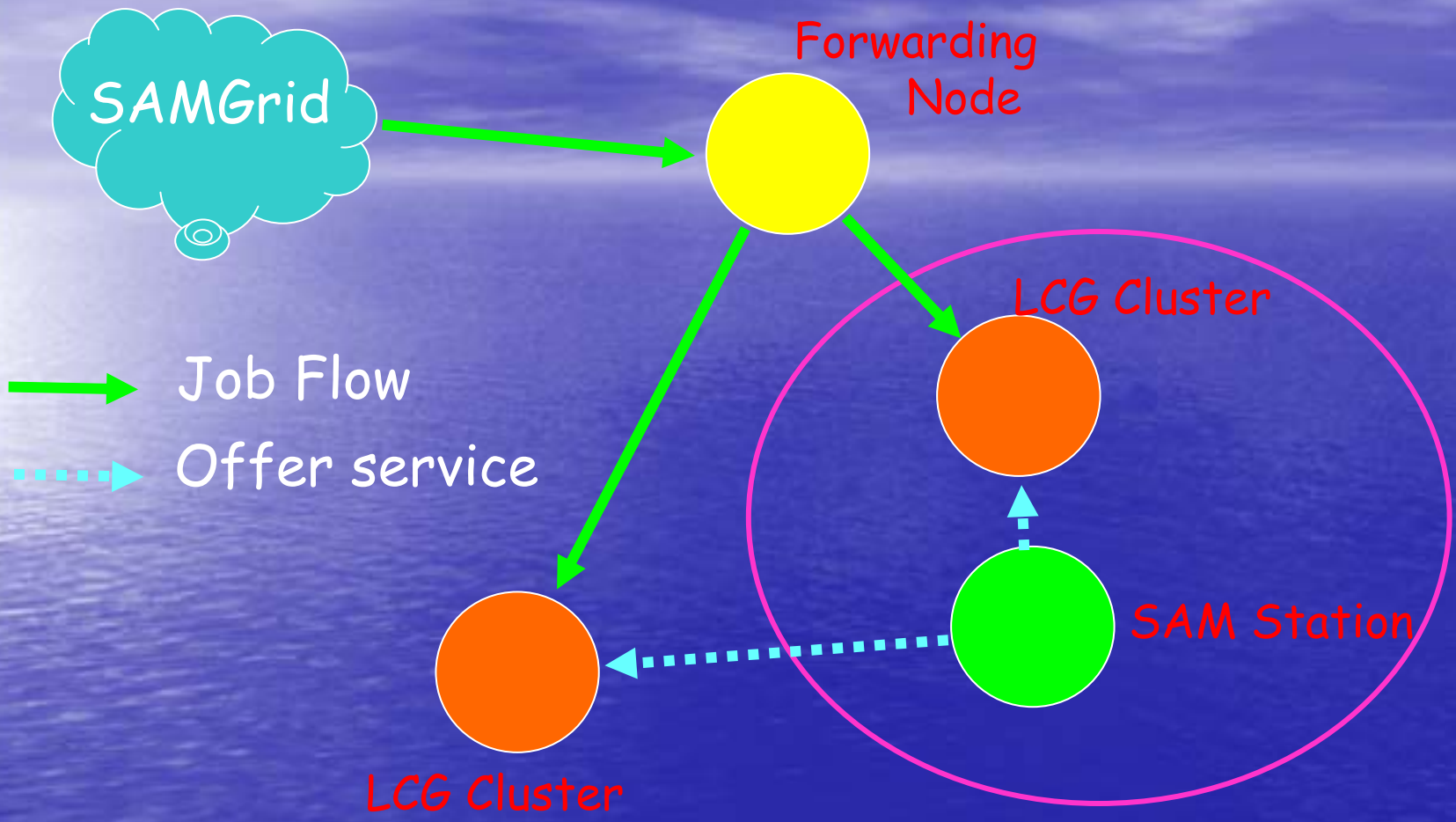
Gatekeeper



CDF Storage Element

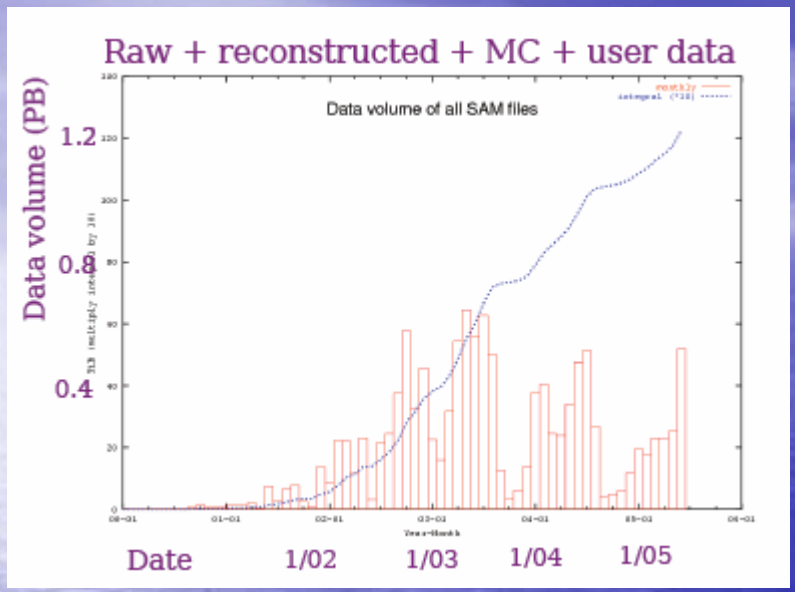
Computing Elements

DO Migration to GRID: SAMGrid ⇒ LCG



Backup

Introduction: Data volume



Data volume vs. time
Total 1.2 PB
Estimated volume of about 5 PB by 2009

	FY	Int L. (fb^-1)	Evts (10^9)	Peak rate (MB/s)	Peak rate (Hz)
Actual	2003	0.3	0.6	20	80
Actual	2004	0.7	1.1	20	80
Estimated	2005	1.3	2.4	40	220
Estimated	2006	2.2	4.7	60	360
Estimated	2007	3.9	7.1	60	360
Estimated	2008	6.0	9.5	60	360
Estimated	2009	8.2	12	60	360

Data logging rate triples from 2004 to 2006
Event rate quadruples due to
Increase compression
Computing problem not static
More difficult with time