DØ Data Management and Grid Computing at CCIN2P3

T. Kurca, P. Lebrun
IN2P3 Lyon

- Setting the Scale - DØ metrics
- SAM - Data Management
- SAM-Grid - Grid Computing
- SAM-Grid/LCG Interoperability
- Summary

IN2P3
Institut National de Physique Nucléaire et de Physique des Particules

Setting the Scale - DØ Metrics 1

- **700 Physicists**
- **78 Institutions**
- **19 Countries**
- **DØ-France: 8 groups**
  80 people

Setting the Scale - DØ Metrics 2

- **Detector - Raw Data**
  - ~1,000,000 Channels
  - ~ 250kB Event size
  - ~ 50+ Hz Event rate
  - ~125 – 250 TB/year
  - Now: 1 B events
- **Total data**
  - raw, reconstructred, simulated

  Now: 760 TB

  By 2008: 3.5 PB

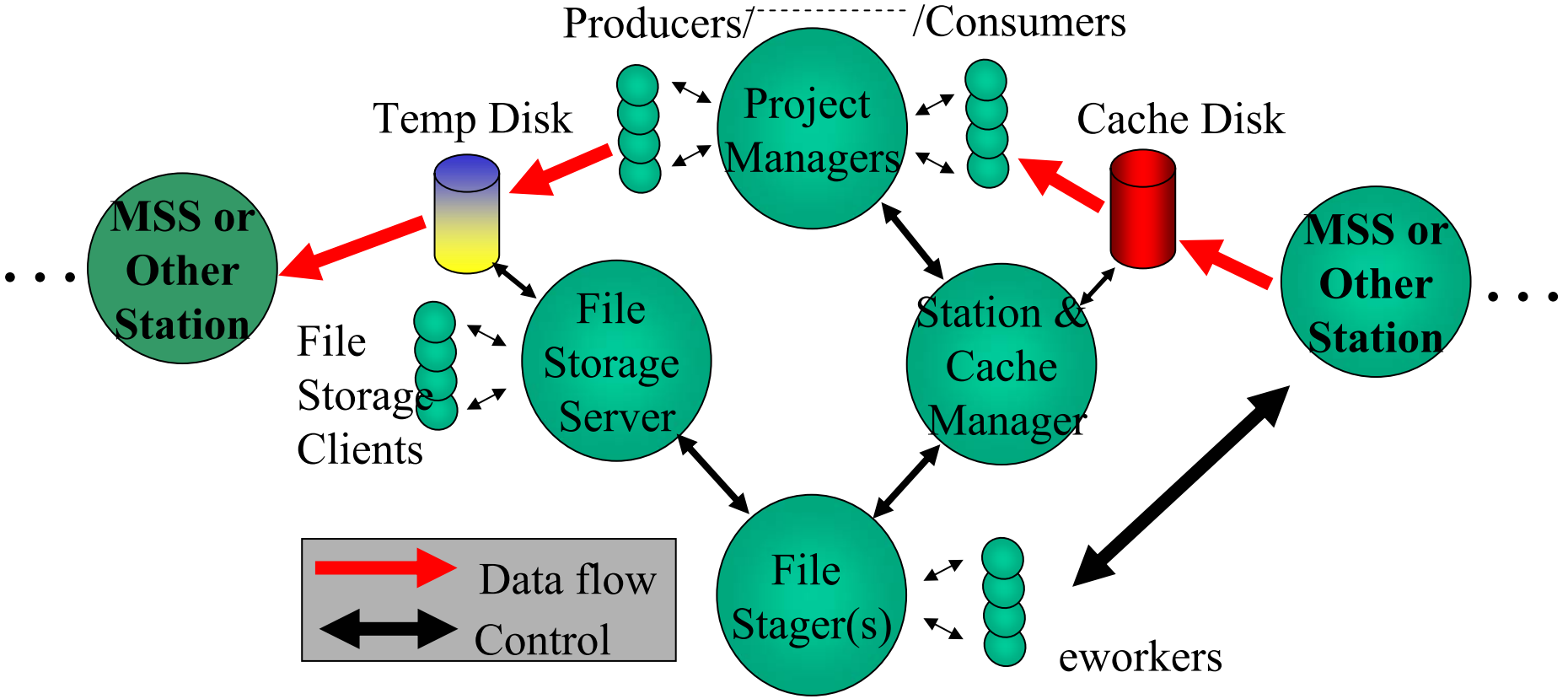
SAM - Data Management System

- **SAM** (Sequential data Access via Metadata)
  - distributed Data Handling System for Run II DØ, CDF experiments
  - set of servers (stations) communicating via CORBA
  - central DB (ORACLE @ FNAL)
  - project started in 1997 by DØ
  - **designed for PETABYTE sized datasets !**

SAM Terms and Concepts

- A project runs on a station and requests delivery of a dataset to one or more consumers on that station.
- **Station:** Processing power + disk cache + (connection to tape storage) + network access to SAM catalog and other station caches
  Example: ccin2p3-analysis
- **Dataset**: metadata description which is resolved through a catalog query to a list of files. Datasets are named.
  Examples: (syntax not exact)
  - data_type physics and run_number 78904 and data_tier raw
  - request_id 5879 and data_tier thumbnail
- **Consumer**: User application (one or many exe instances)
  Examples: script to copy files; reconstruction job

Components of a SAM Station



- SAM: distributed data movement and management service: data replication by the use of disk caches during file routing
- SAM is a fully functional meta-data catalog.

SAM Functionalities

- ✓ **file storage** from online and processing systems
  → MSS - FNAL Enstore, CCIN2P3 HPSS…
  disk caches around the world
- ✓ **routed file delivery**
  - user doesn't care about file locations
- ✓ **file metadata cataloging** → datasets creation based on file metadata
- ✓ **analysis bookkeeping** → which files processed succesfuly by which application when and where
- ✓ **user authentication**
- ✓ **local and remote monitoring capabilities**
  SAM TV@ DØ , SAM TV@ CDF

SAM station at CCIN2P3

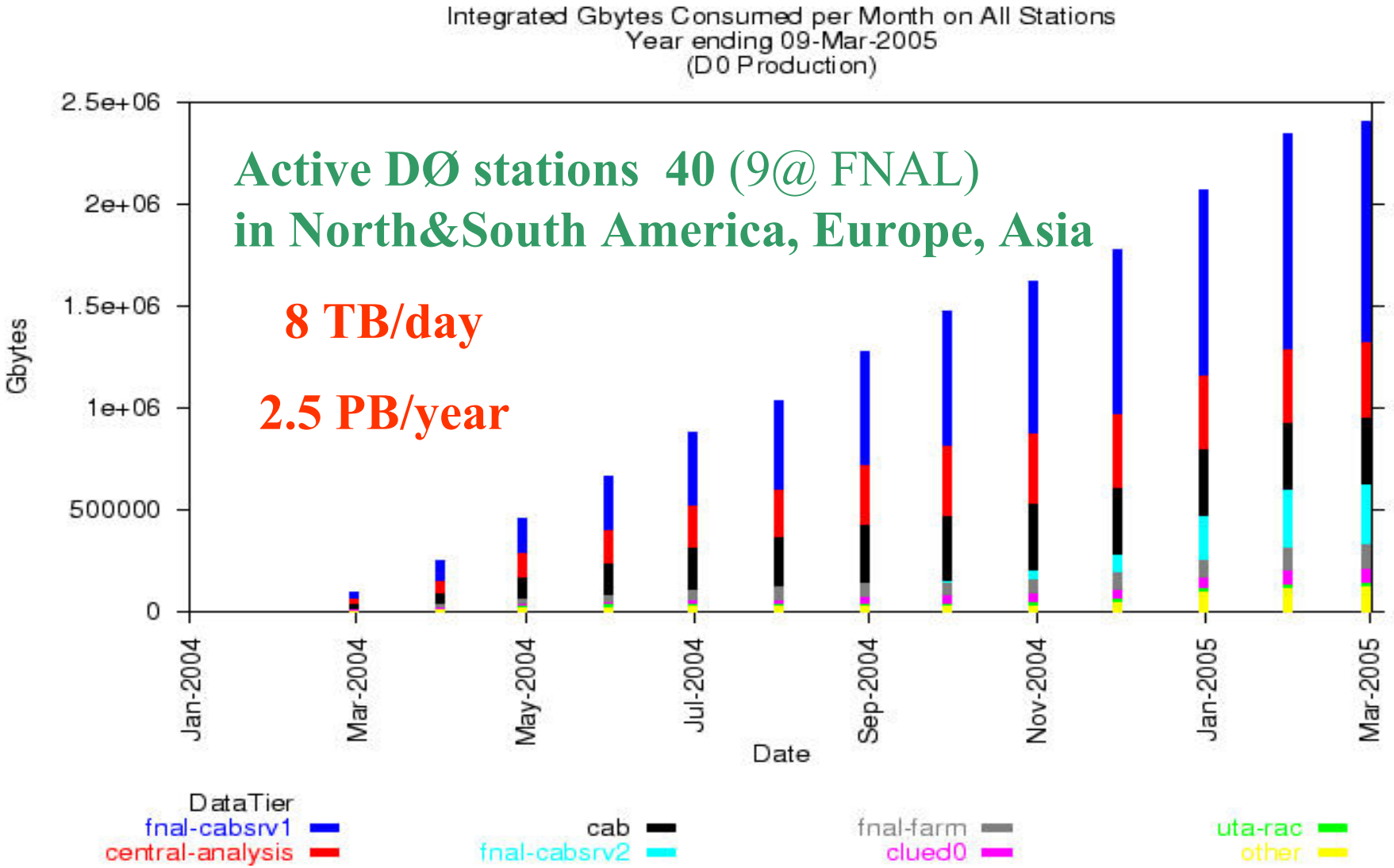
Station name: ccin2p3-analysis

- gateway node →network access to SAM catalog and other station caches + other MSS
- **2 caches** :
  - SAM cache 100 GB ccd0.in2p3.fr:/samgrid
  - HPSS 'disk' 150 TB (seen as a SAM cache!)
    rfio://in2p3.fr:cchpssd0.in2p3.fr/hpss/in2p3.fr/group/d0
    sam/ccin2p3 specific interface !

Node: ccd0.in2p3.fr under Scientific Linux 3.0.3

- PC IV Linux bi-proc 2.8 GHz
- Memory 2 GB
- local disk 40 GB ccd0.in2p3.fr:/d0products/
  260 GB ccd0.in2p3.fr:/other

SAM Data Consumption



2003 Data Reprocessing

- ~5,5 months preparation → done in ≤ 7 weeks
- 100 M events done remotely (20%)
- ~25 TB data transferred
- **CCIN2P3:** local bookkeeping (ORACLE DB)

→very efficient resubmission of failed jobs + fabric stability

…. *but not « real Grid computing » yet …*

Institute	# Events (millions)	# CPUs (1 GHz PIII)
CCIN2P3	**36**	**160**
UK GridPP	23	270
GridKa	21	200
WestGrid (CA)	12	1000
NIKHEF	7	400

The SAM-Grid Project

- **Goal**:

  enable globaly distributed computing for DØ and CDF

- **SAM-Grid = SAM + JIM**

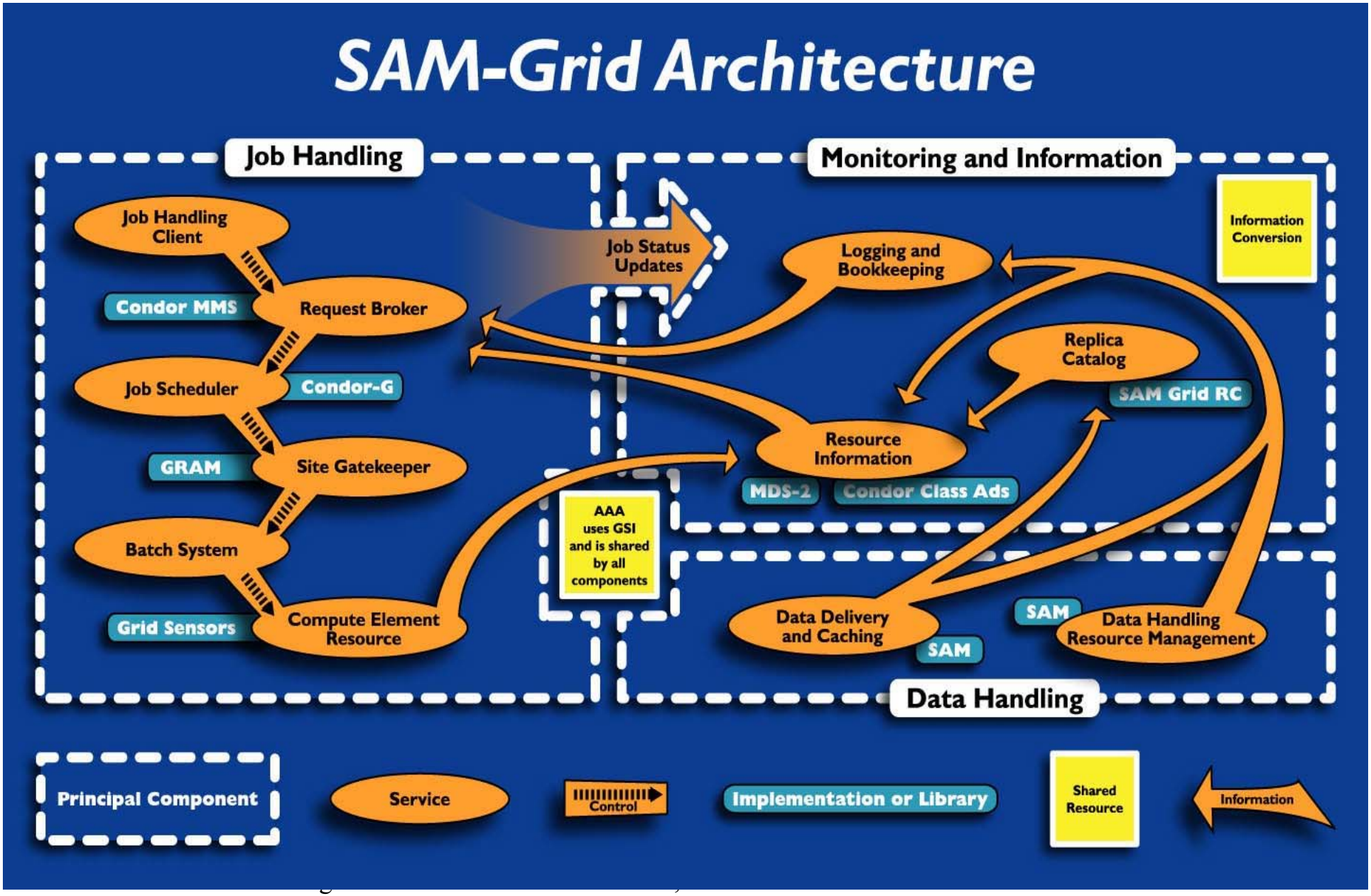
  - JIM (Job & Information Manager) started end of 2001

  enhance SAM, the distributed data handling system

  → *integrating standard Grid tools and protocols*

  → *developing new solutions for Grid computing (JIM)*

- **Funds**: Fermilab, PPDG (US) and GridPP (UK)


SAM-Grid Architecture
Job Handling
Monitoring and Information
Job Handling Client
Job Status Updates
Logging and Bookkeeping
Information Conversion
Condor MMS
Request Broker
Replica Catalog
Job Scheduler
Condor-G
SAM Grid RC
Resource Information
GRAM
Site Gatekeeper
MDS-2
Condor Class Ads
AAA uses GSI and is shared by all components
Batch System
Grid Sensors
Compute Element Resource
Data Delivery and Caching
SAM
SAM
Data Handling Resource Management
Data Handling
Principal Component
Service
Control
Implementation or Library
Shared Resource
Information

SAM-Grid & Grid Services

- Distributable **sam_client** provides access to:
  - VO **storage service** (sam store command, interfaced to sam_cp)
  - VO **metadata service** (sam translate constraints)
  - VO **replica location service** (sam get next file)
  - Process **bookkeeping services**
- **JIM components provide**:
  - **Job submission service** via Globus Job Manager
  - **Job monitoring service** from remote infrastructure
  - **Authentication services**

SAM-Grid at CCIN2P3

- **SAM station:** ccin2p3-analysis
  sam_station v4_2_1_77

- **SAM-Grid** installed in summer 2003 as a
  - client (very light-weight) &
  - submission &
  - monitoring &
  - execution site

→ full grid functionality

→used for official MC-production

→tested for reprocessing from raw data

- production & merging individual thumbnails

CCIN2P3 specifica

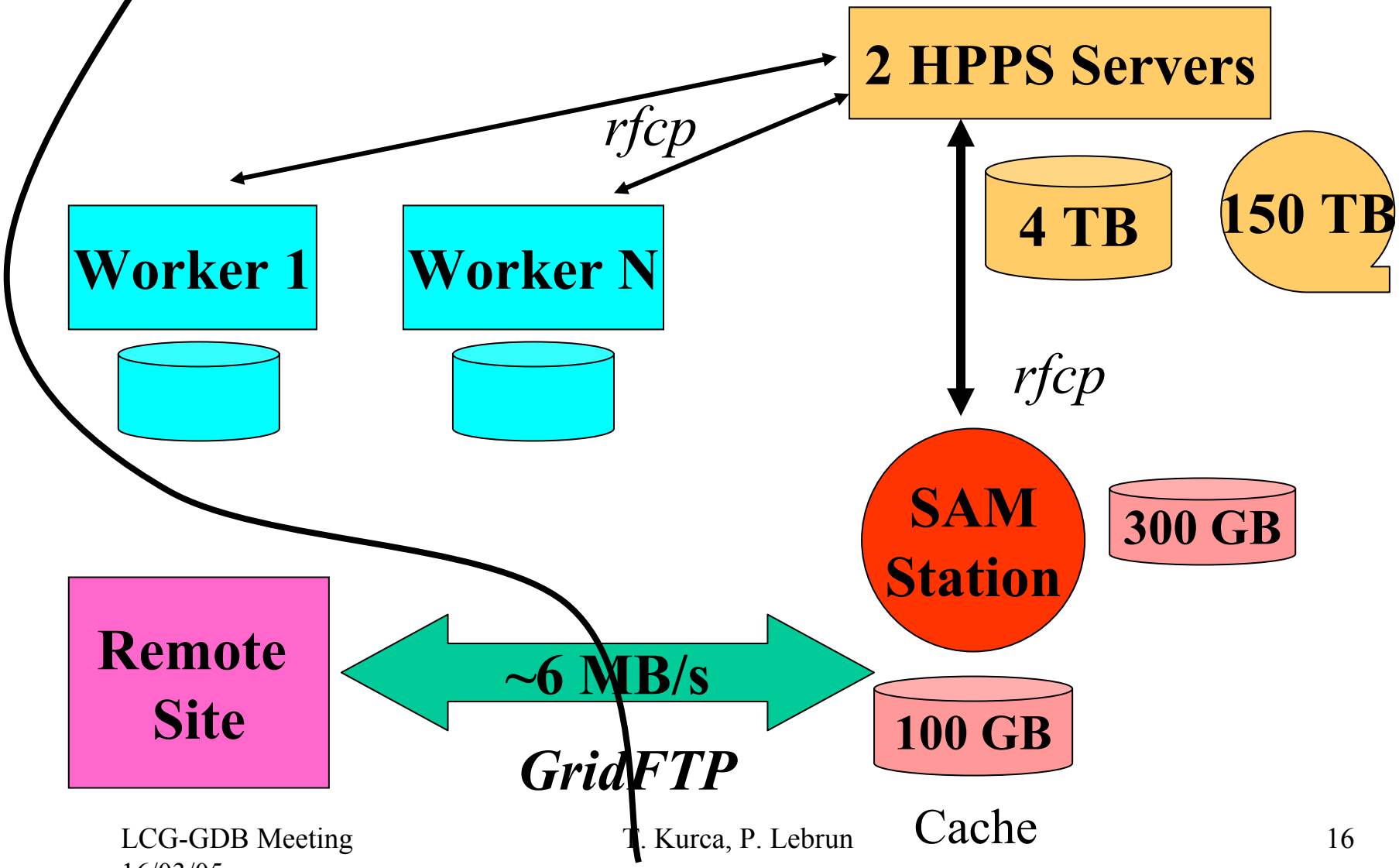
- **HPSS data files accesed with RFIO** (rfcp or API)
  - data transfer from outside via SAM cache
  - local data access not using SAM cache

  (files in HPSS are seen as local files)

  ---> D0 framework had to be adapted (-->using rfcp)
  - data transfer works fine
  - had problems with framework jobs (files reported ‘ not available ’)
- **Batch system** : local product BQS

  → batch adapter/idealizer was developped

SAMGrid @ CCIN2P3



Status & installation of DØ SAM-Grid

- **Active execution sites: 10 DØ (1 @ FNAL)**
  - **Active Monte Carlo production at multiple sites**
  - **Prepared for DØ reprocessing from raw data**
  - **Installation**
  - **via ups/upd FNAL products**
  - **No specific requirements on environment**
  - **Non invasive system , very flexible**

    →Drawback : non trivial configuration

    requires good system understanding

SAM GRID INFORMATION & MONITORING SYSTEM

Launching the Monitoring System:

Please click at the map to monitor the execution sites.
Get information about the submission sites
Get information about the advertised sites.

Participating Experiments:

● D0 ● CDF

2005 Data Reprocessing

- ***The Goal:***
  reprocess remotely all 10^9 DØ events via SAM-Grid
- **~250 TB** of raw data to move
- This time from **raw data** → DB access needed :
  calibration proxy DB-servers installed & tested at remote sites
- Merging and final storage from remote sites
- Time scale: 6 – 8 months
- CPU's foreseen: 2700 (1 GHz PIII)
  - CCIN2P3 ~15%
- **Real full scale Grid computing!**

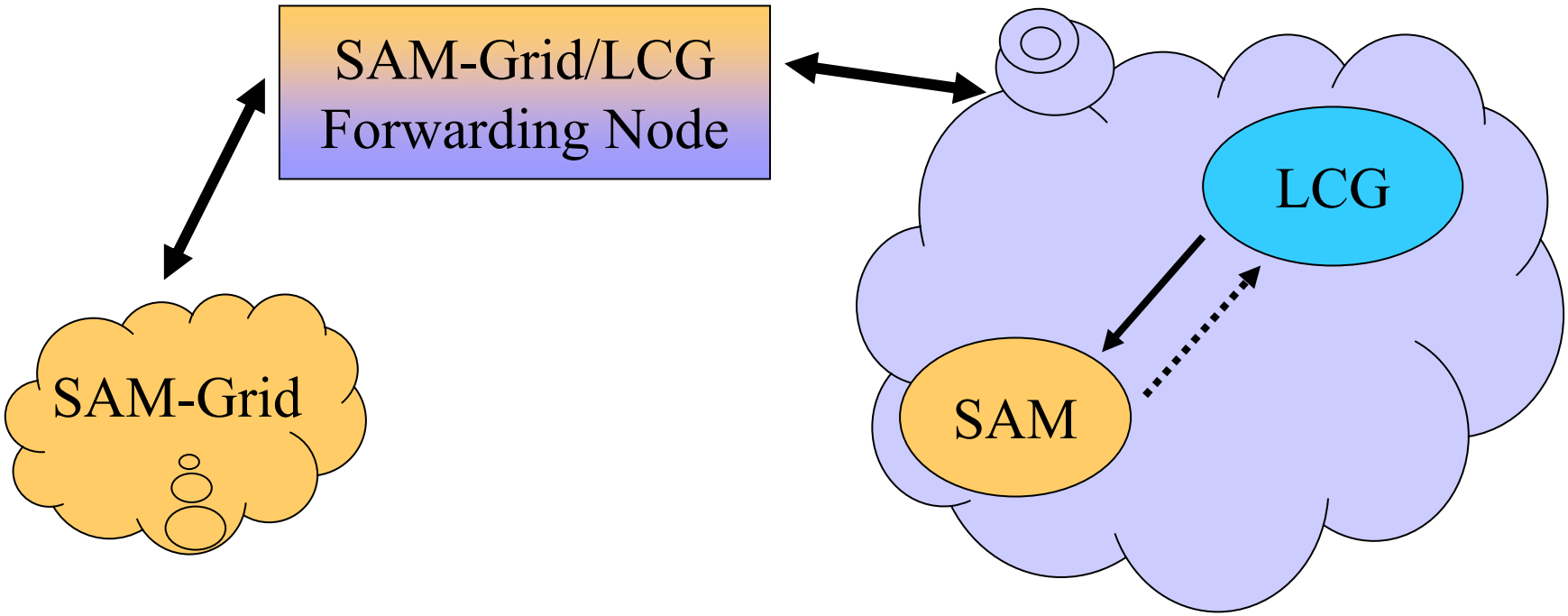
The SAM-Grid/LCG Interoperability

- **Motivation & Goals**
  - resources and manpower drifting towards LHC
  - make LCG resources available to DØ via SAM-Grid
  - integration project, no massive code changes expected
- **Limitations & Problems**
  - most of the LCG resources w/o SAM-Grid gateway node
  - firewall problems : station interfaces use callbacks
  - SAM/LCG batch adapter to be developped
  - security : authentication → agreement on a set of CA
    authorization to use LCG resources

SAM-Grid/LCG Phase-1

SAM station installed within the LCG site
to overcome accessibility problems

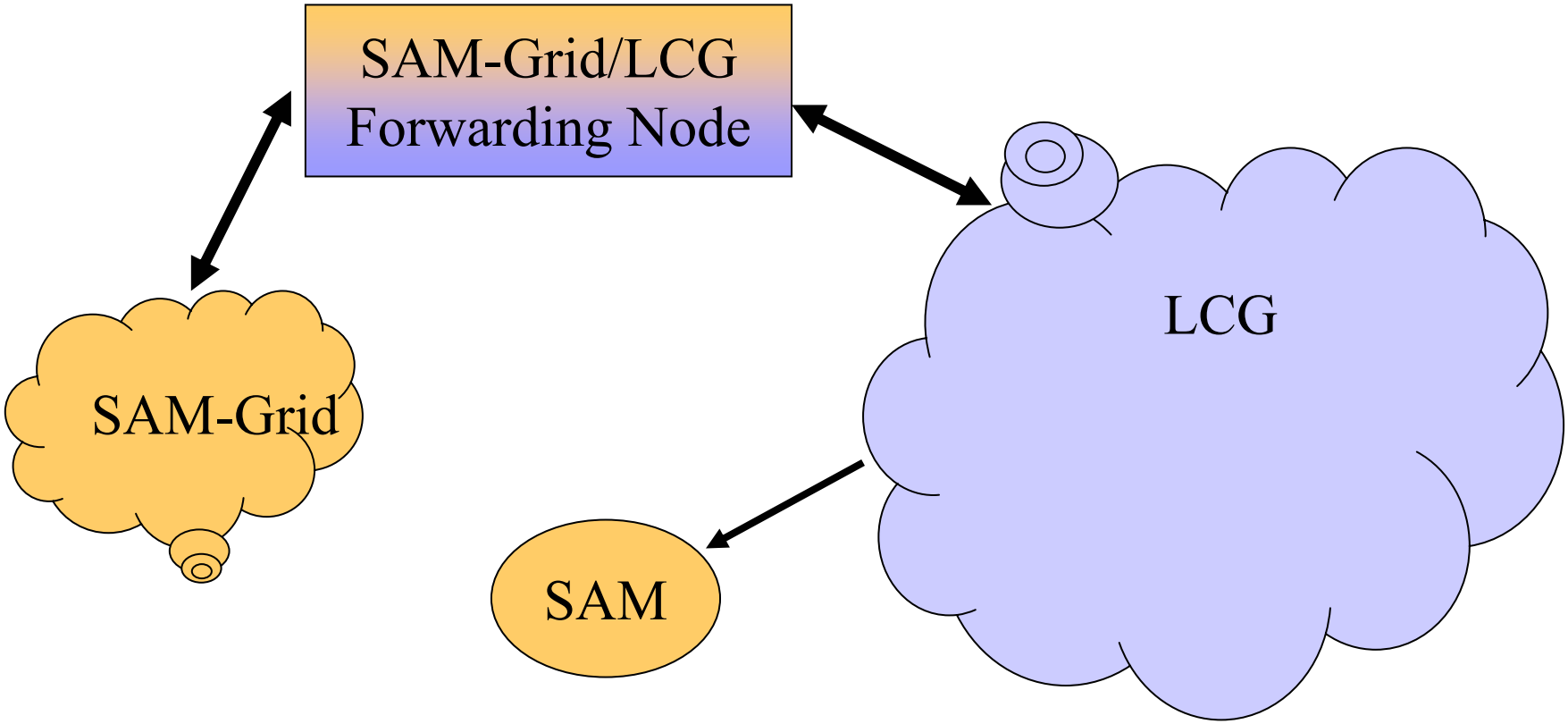
- application executable & input data files
  are downloaded from SAM



SAM-Grid/LCG Phase-2

Requires to change the sam_client :

replace callback by polling mechanism



Summary 1

- **DØ – running HEP experiment:**
  - produces PB-sized datasets
  - computing resources distributed around the world
- **SAM – Distributed Data Handling System**
  - reliable data management & worldwide file delivery to the users
- **SAM-Grid – full Grid functionality**
  - standard Grid middleware + specific products
  - MC-production running
  - Remote reprocessing → to start very soon

Summary 2

- **Future:**
  - **work on Grids interoperability SAM-Grid/LCG**

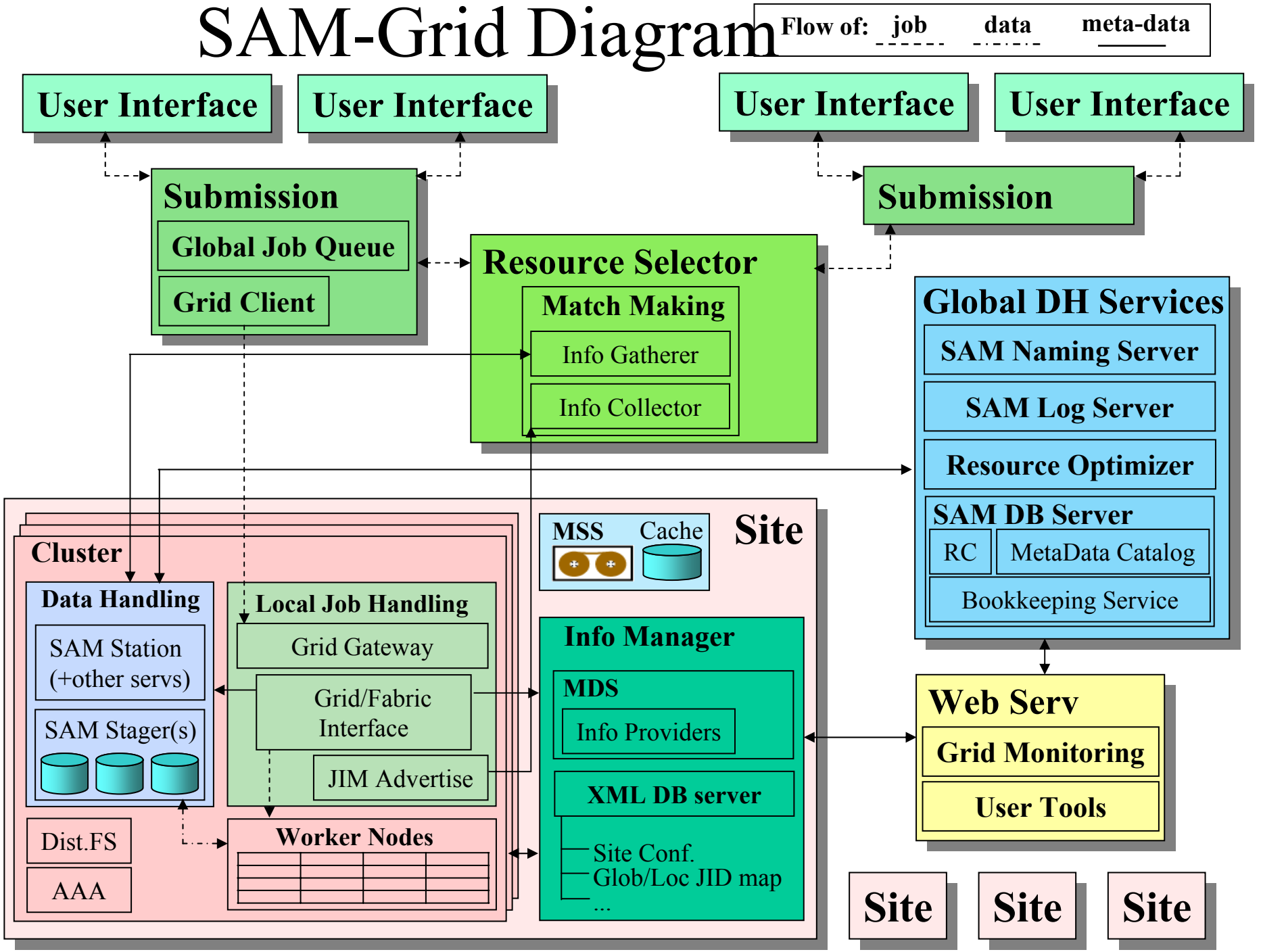
    → continuation of a global vision for the best use of available resources
- **Remote computing – huge profit to DØ experiment**

- **CCIN2P3: major contribution to the D0-Grid computing**
  - reliable fabric & know-how

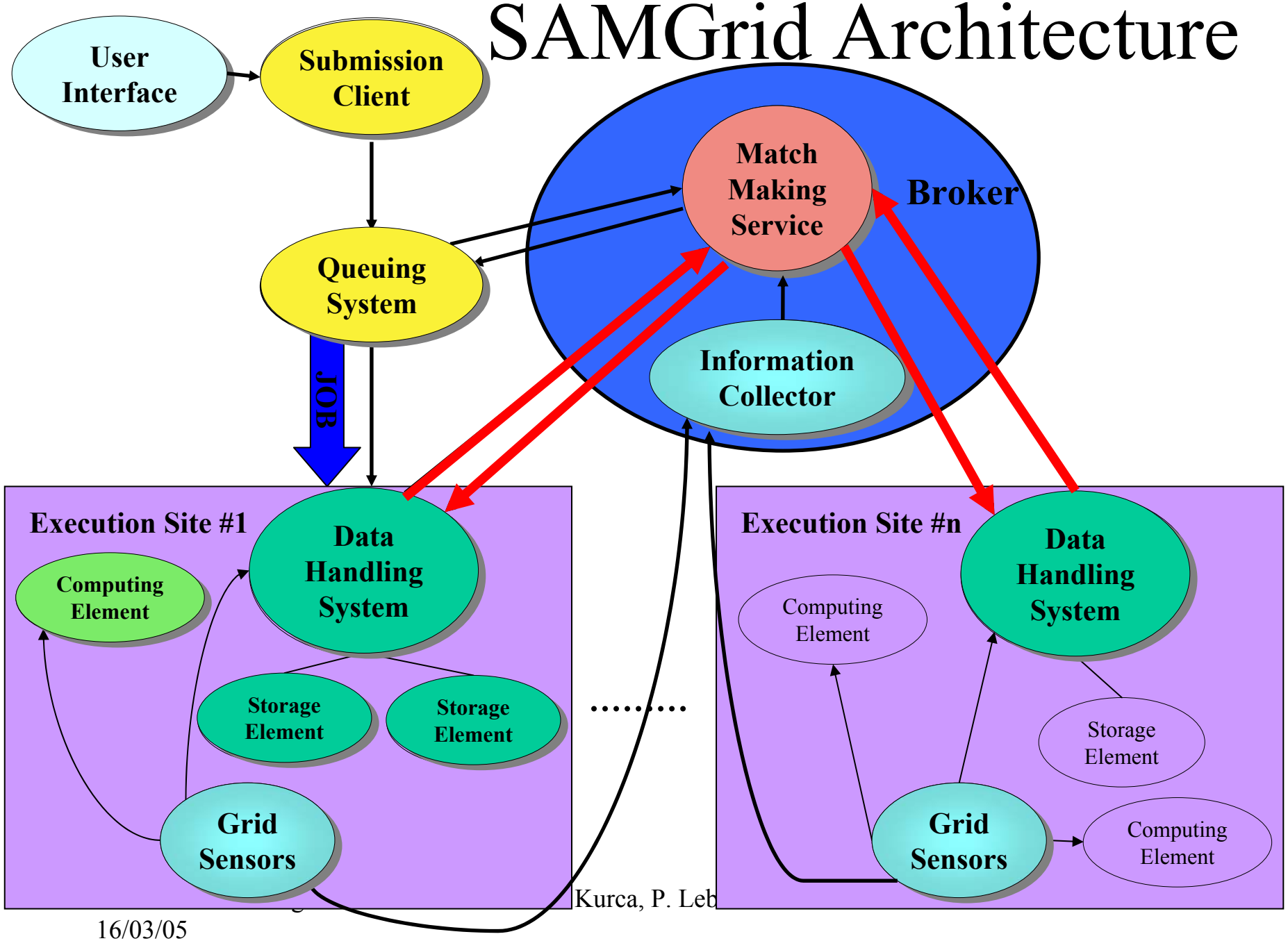
    ⊕ close collaboration with SAM-Grid team

…backup slides

SAM-Grid Diagram



SAMGrid Architecture



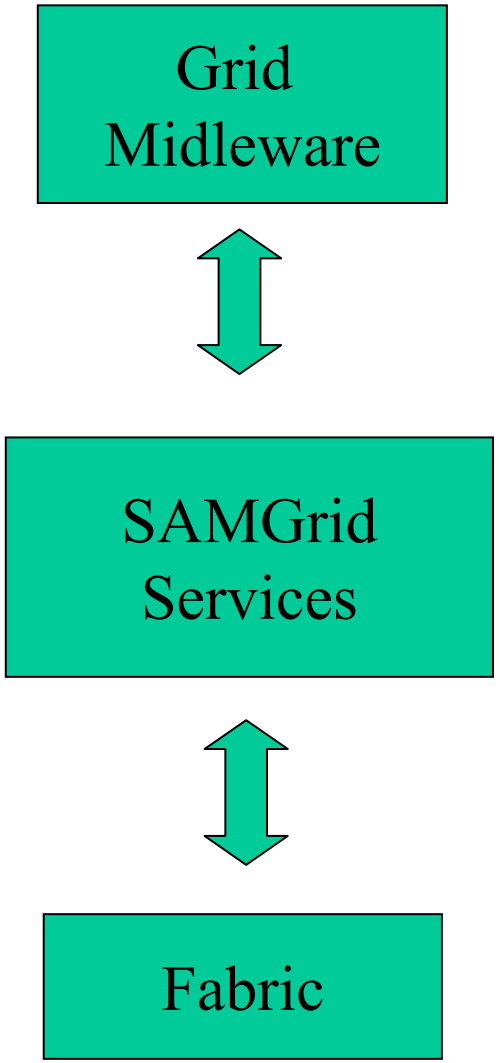
SAM-Grid at CCIN2P3

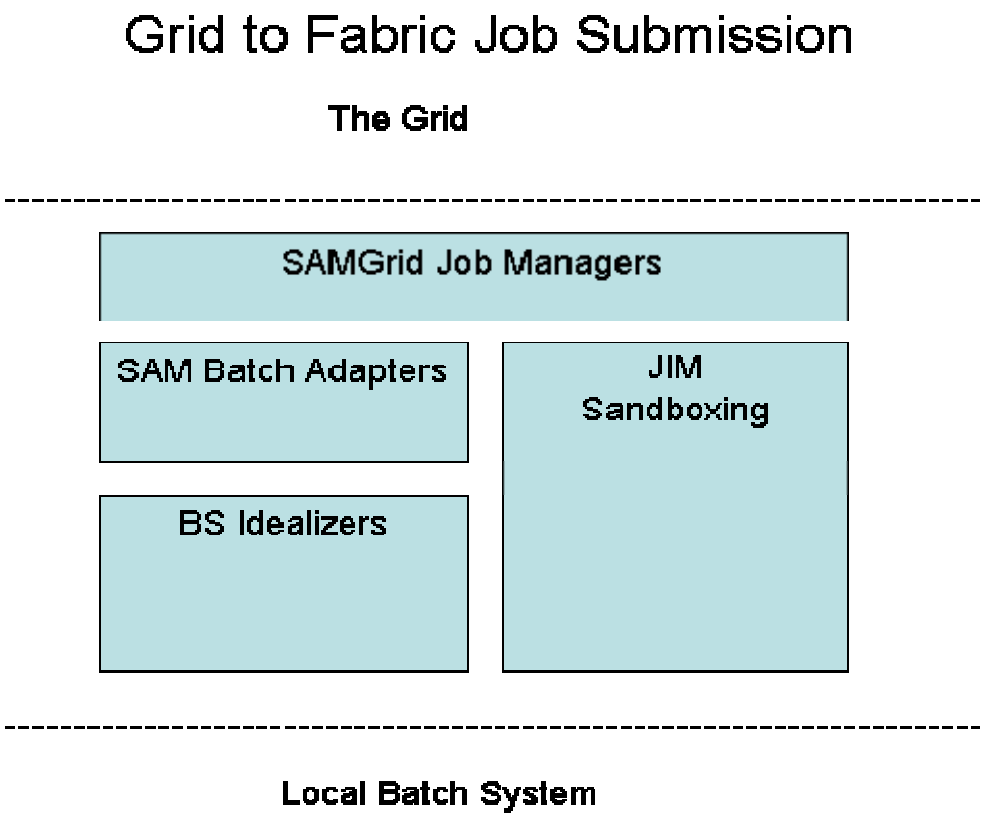
- **Open ports for Incoming TCP connections:**
- **2119 - grid-gatekeeper**
- **61001-61500 - job-managers**
- **61501-61700 - condor_scheduler**
- **2135 - grid MDS**
- **7080, 7081 - tomcat**
- **4501-4505 - sam**
- 4567 - sam_gridftp server
- 60001-61000 - sam_gridftp client

Forwarding Node

- ✓ Globus Gatekeeper
- ✓ Sandboxing services
- ✓ XML database
- ✓ SAM-Grid advertisement
- ✓ LCG job submission & monitoring software
- ✓ SAM-LCG batch adapter

Grid to Fabric Job Submission

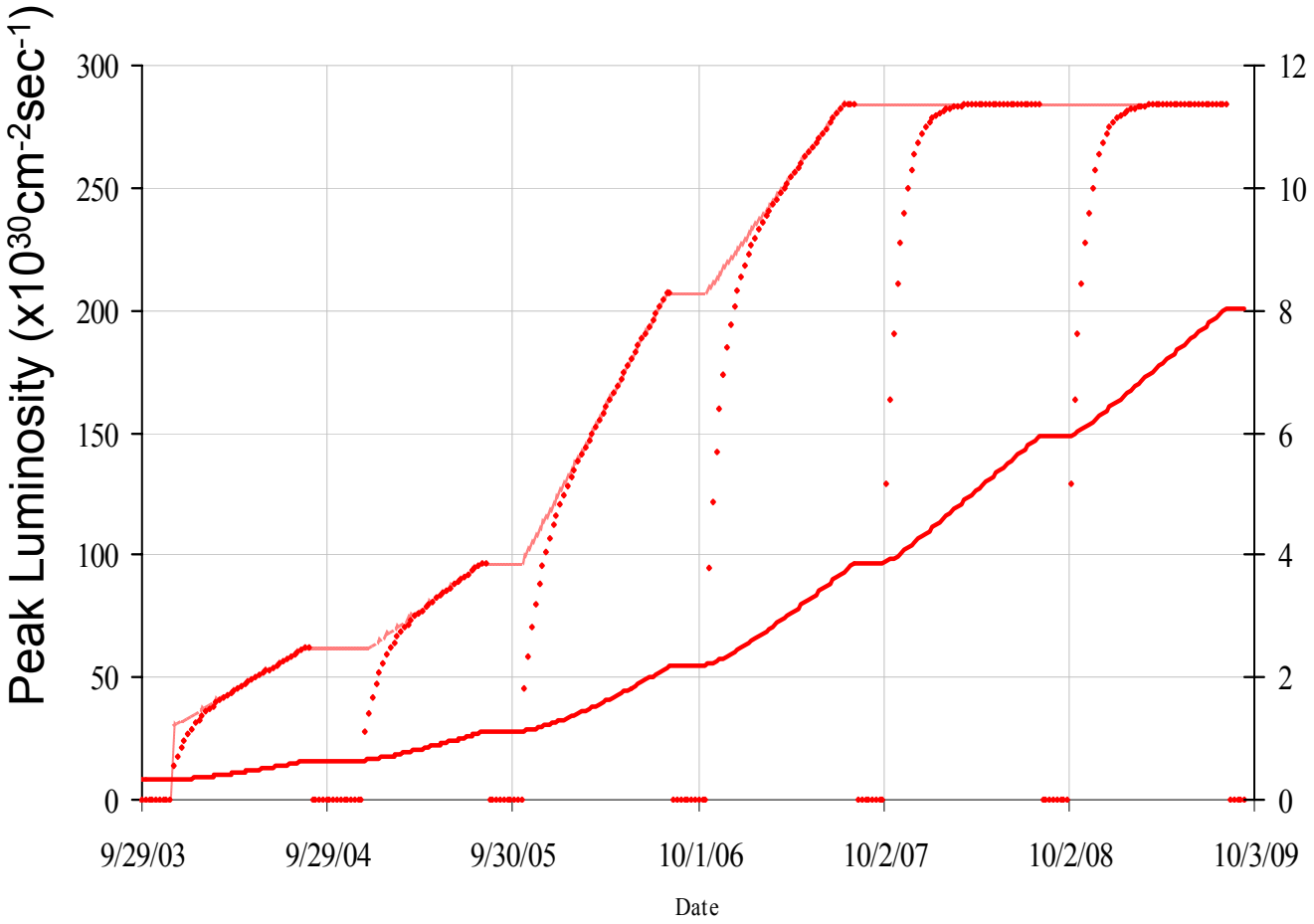




Luminosity Prospects

Accelerator draft plan:
Peak luminosities

~2.8e32 peak
~8 fb^{-1} integrated



SAM

A Station is a collection of resources controlled by the SAM system.