D0 Run II Data SAM Project Status

Lee Lueking

Run II Project Review

October 26, 1999

Topics

- Summary of Release Schedule
- Version 1.0 Features
- Code Packages Release
- Who is doing what
- System Functionality
- D0 Integration and Users
- User Documentation
- Moving into Production Operations

SAM Version 1.0

- Infrastructure
  - Development, Integration, and Production Databases on dedicated SUN server (D0ora1,2) machines.
  - Database server process
  - Log and info Server processes
  - Admin tools for monitoring and diagnostics
  - Code management using cvs
  - Product releases and distribution using ups/upd
- User Interfaces
  - Project definition user tools: Web-based GUI, and CLI
  - Database browsing: Web-based GUI
  - Data delivery via Project Master & Stagers: CLI, D0Framework
  - Data storage via File Storage Server and Stager: CLI
    - On-line data logger,
    - User or production processed file,
    - Monte Carlo files

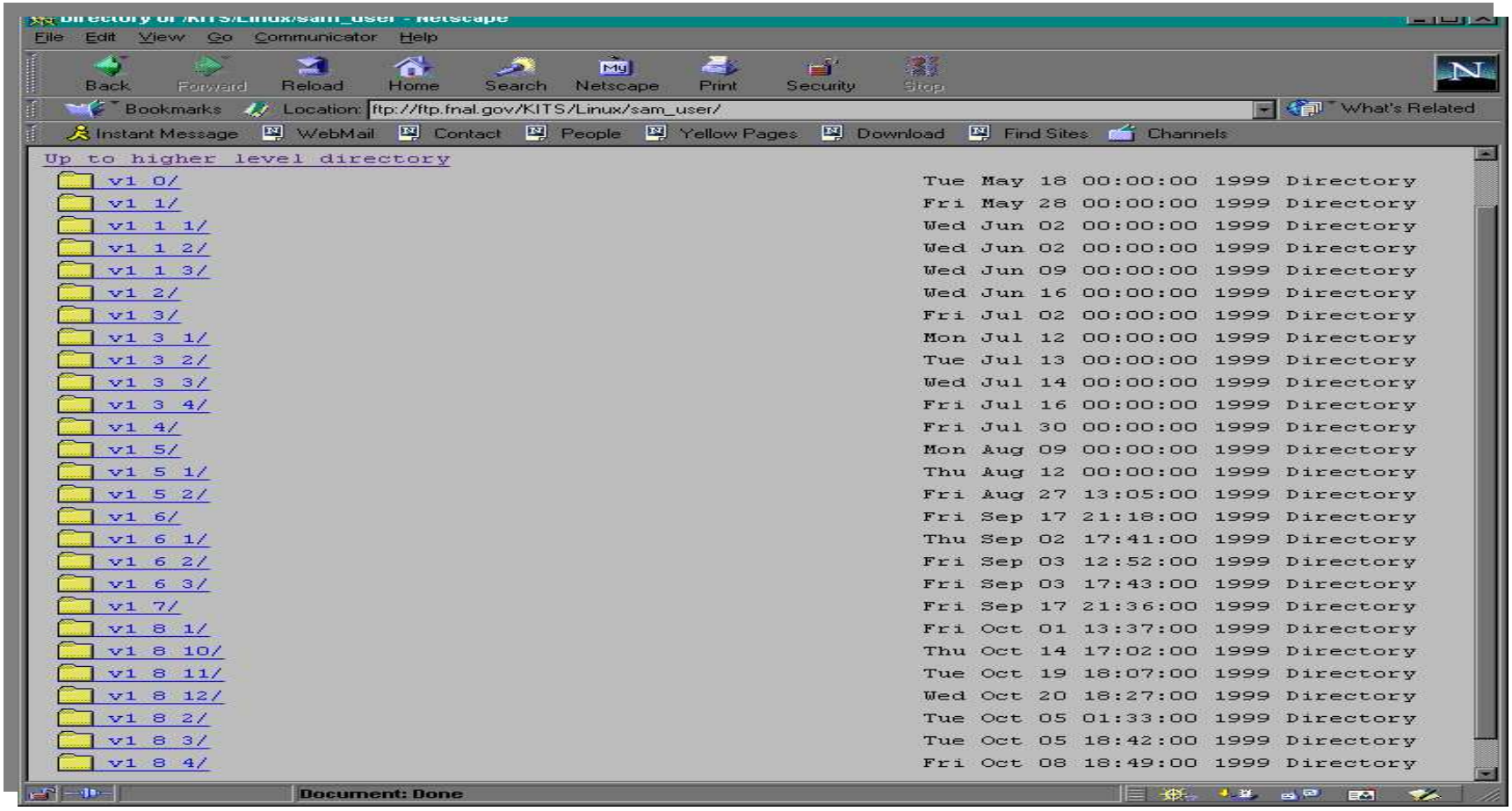
SAM Version 1.0 (cont.)

- Event Catalog loading and use
- File preferred locations function
- Python scripts for launching user applications
- Enhanced Database Server features, eg. exception handling
- System monitoring tool

Overview of Release Schedule

Version	Features	Scheduled Date	Actual Date
Prototype	SC 98 Demo, simple db schema, functioning database and project servers	Nov. 98	Nov. 98
V0.0	Improved db schema w/ projects, process, datafiles,…,dev & prod db. Entered data by hand, delivered data to early bird users. Browser, admin db, and monitoring tools. CLI improved. Project tools.	March 15, 99	July 15, 99
V1.0	Automated data entry, data store for online, processed and MC w/ FSS. Maturing db schema w/ events table. Int db, , admin charts, SAMManager. Python user script improved. Preferred locations. Web based GUI for project tools.	Oct. 1, 99	Oct. 15, 99
V1.5	V1 Fixes. Station Design and implementation. Improved reliability, recovery mechanisms. 2nd generation project tools. Meta data creation in D0Framework. Data import facility, tape injest/sync with SAM. Details, details, details.	Dec. 15, 99	

SAM Code Packages in KITS



Who is Doing What

- Lee Lueking - Organize meetings. MCC99 Data entry and coordinating. D0 User liaison. Write and massage documentation. Give review talks.
- Vicky White - Schedule (including WBS), Web page coordinator. Design and shepherding. Presentations et cetera.
- Igor Terekhov - Overall system designer. Developer for station, project master, stager, and file storage servers. Logger. Trouble shooter. D0 code liaison.
- Julie Trumbo (50%)- Database schema, database tools. Oracle expert. Helps maintain and augment user CLI's. Tests user interfaces.
- Heidi Schellman - Testing sam on farm, breaking everything, uncovering every conceivable problem possible, but still optimistic. D0 User liaison.

Who is Doing What (cont.)

- Sinisa Viseli - SAMManager and integrating with D0Framework. Will be performing stress testing on all platforms. D0 code liaison.
- Matt Vranicar - db_server developer and maintainer. upd/ups release procedures. Oracle expert. Helps organize all the docs. Troubleshooter.
- Rich Wellner - Project tools including CLI and Java GUI, Sam admin monitoring tool as Java GUI. Developer for info_server. Testing overall system. Interfaces closely with Enstore. Troubleshooting.
- Steve White(50%) - db_server maintenance and development, web server on Ora2, project server updates. Also, works with D0 Calibration and RCP database efforts.

DB Support + D0 Systems Contributions

- Diana Bonham - Add new db user accounts. Backups on Ora1 and 2. Install and configure Legato backup. Plus other dba type activities.
- Steve Kovich - D0Ora1 & 2 system manager. Install and test new hardware, Clariion RAID box and Exabyte robot. Install and configure Legato backup software.
- Nellie Stanfield - Helps D0 when needed.
- Julie Trumbo (50%)- Maintainer for Oracle Designer repository. Testing Oracle 8.1.5 and db server RAID box. Oracle Enterprise Manager for monitoring systems.
- Peilei Zhang - Helped coordinate configuration for RAID system, Moving db to new RAID system. Will help plan, install, and configure Legato backup.

Functionality

- Storing Data
- Browsing
- Creating Projects
- Accessing Data
- Monitoring
- Diagnostics

File Store Server

- Files are stored into sam using the File Storage Server
  - Command :

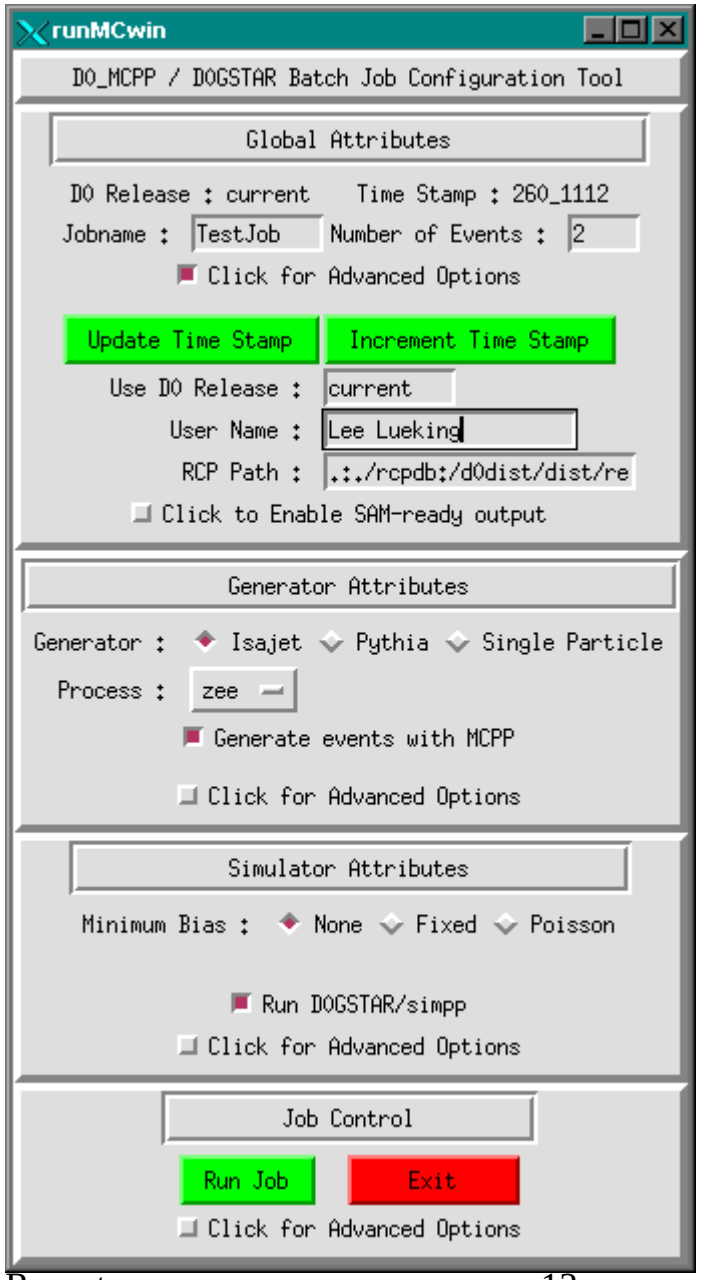
    sam store --descrip=meta_data_file --source=source_directory --destination=destination_directory
  - Meta_data_file:

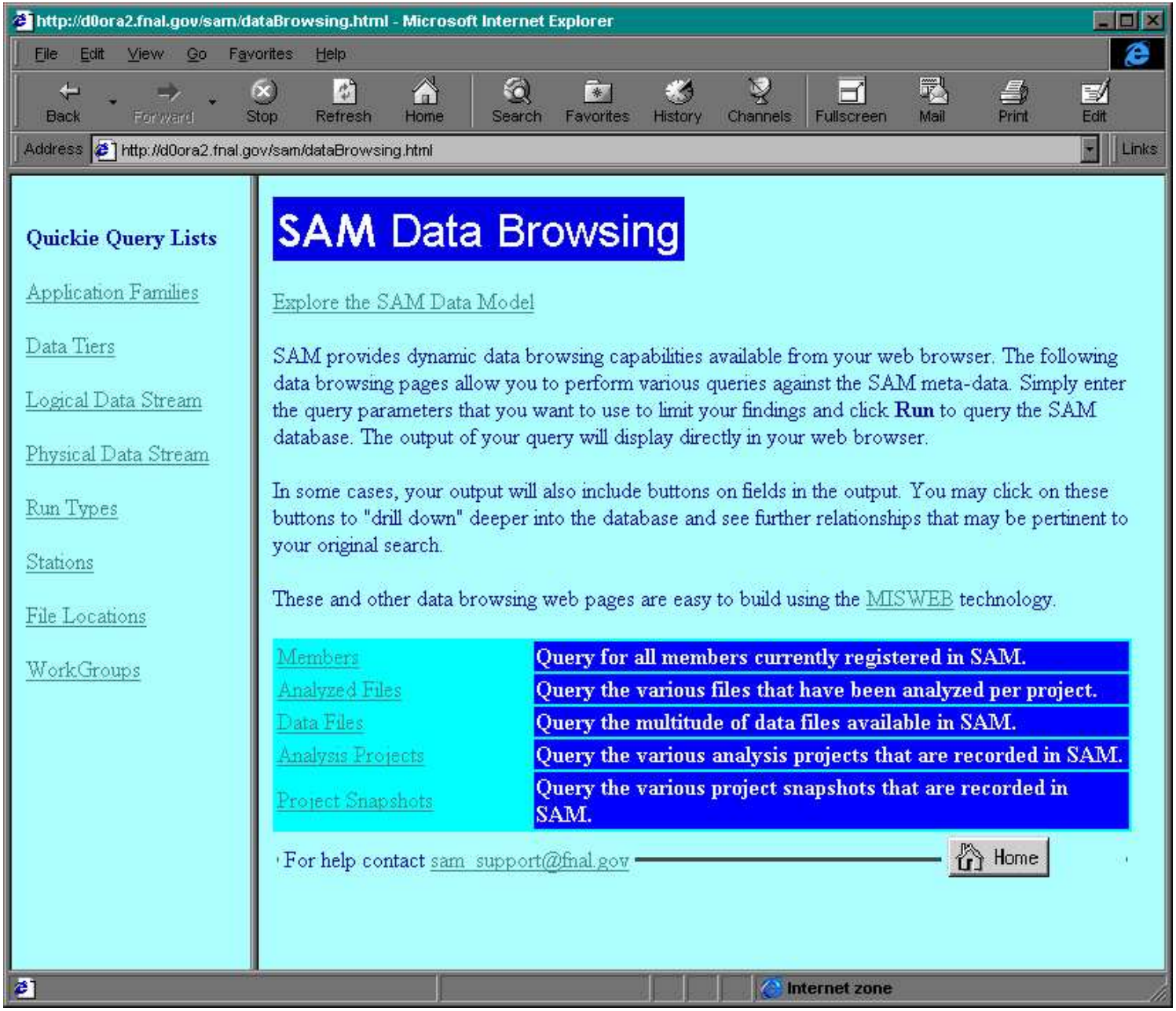
```
from import_classes import *
    TheFile = ReconstructedFile(name='fifo.ace',
      size=12345,events=Events(1, 100, 70),
      tier='reconstructed', start_time='08/01/1998 17:00:00',
      end_time='08/01/1998 18:00:00', parent_name='ssh-
      kurino')
```

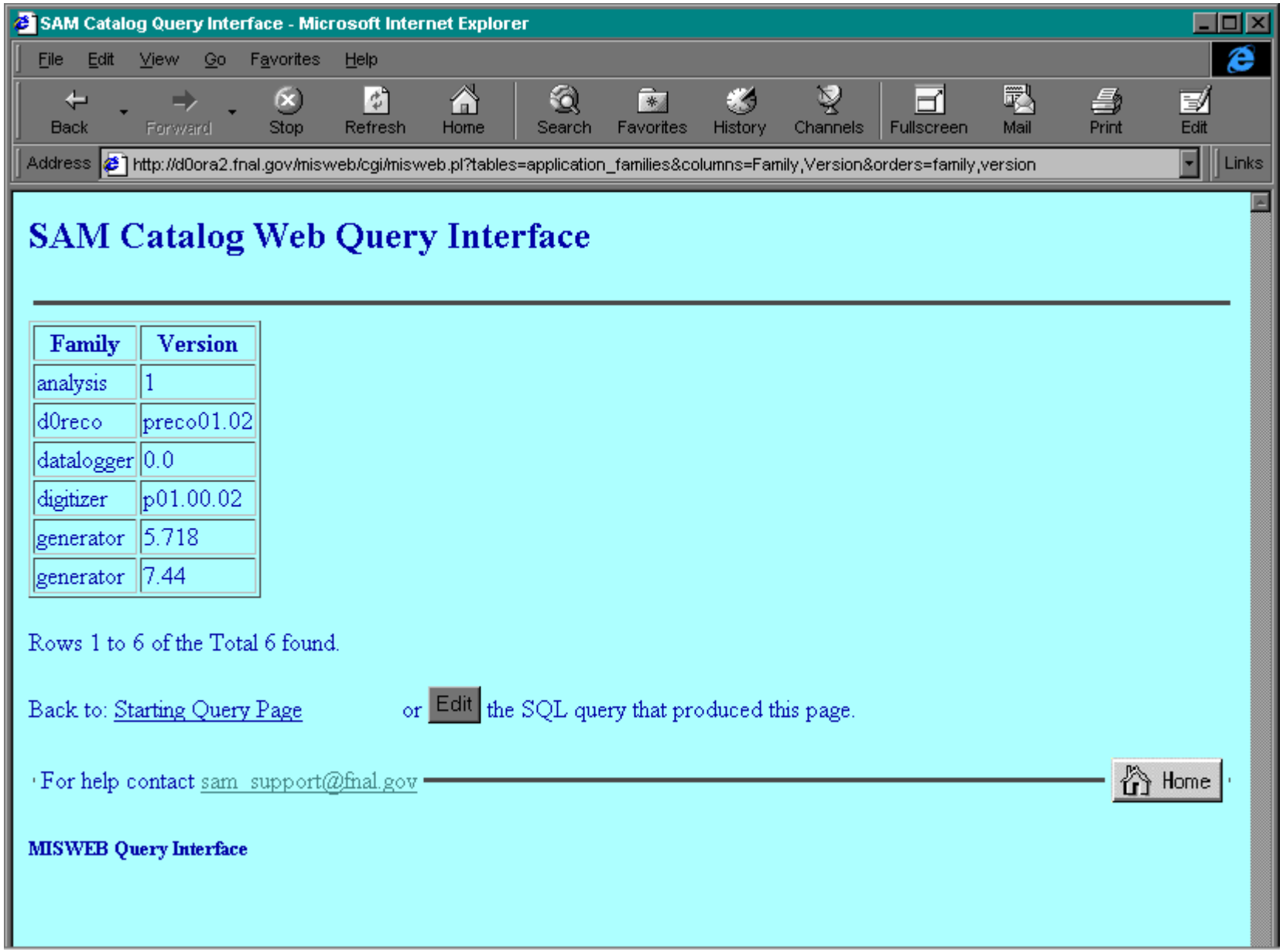
  - Metadata formats for data logger, MC data, etc.

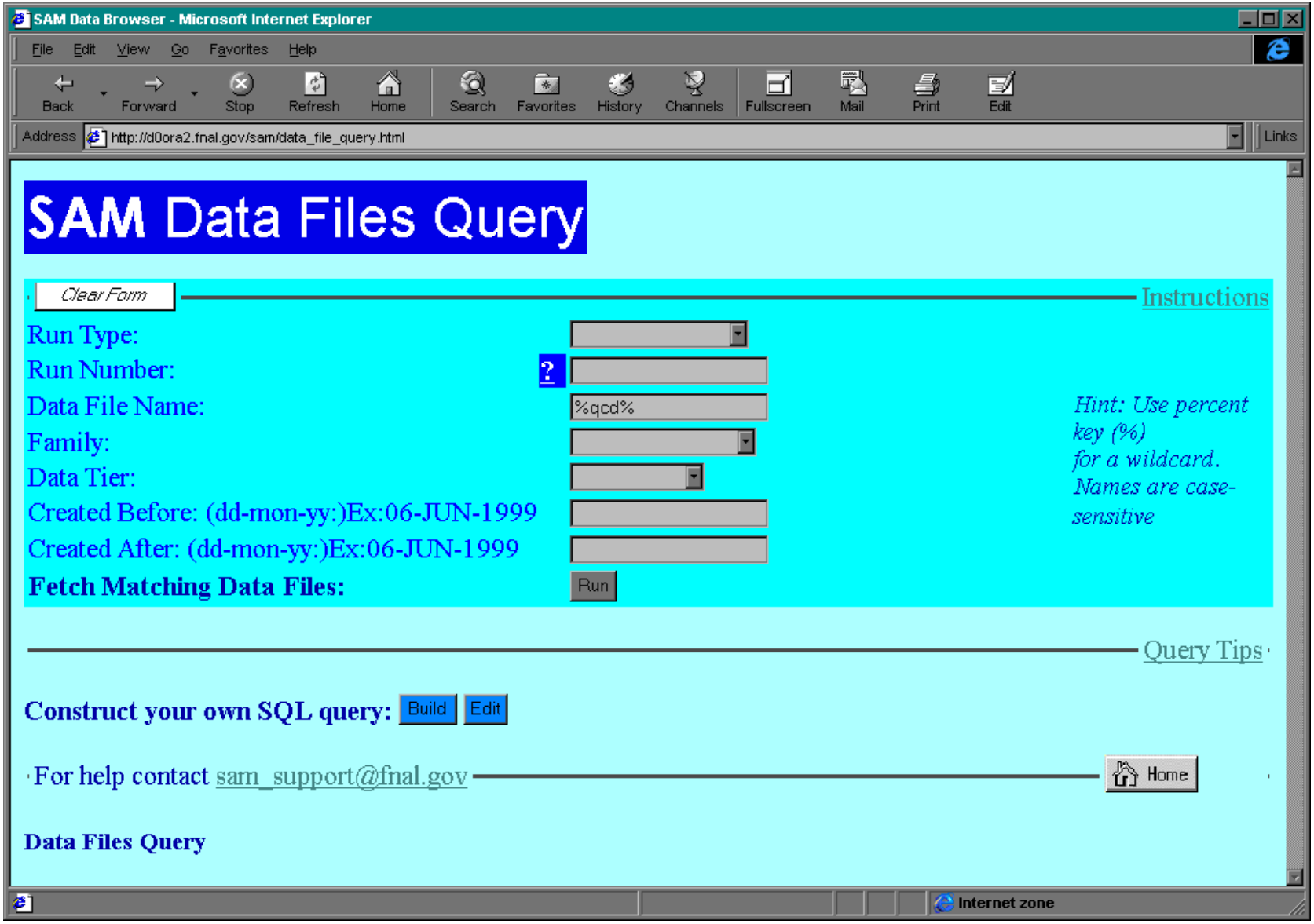
Adding data to SAM: MC example

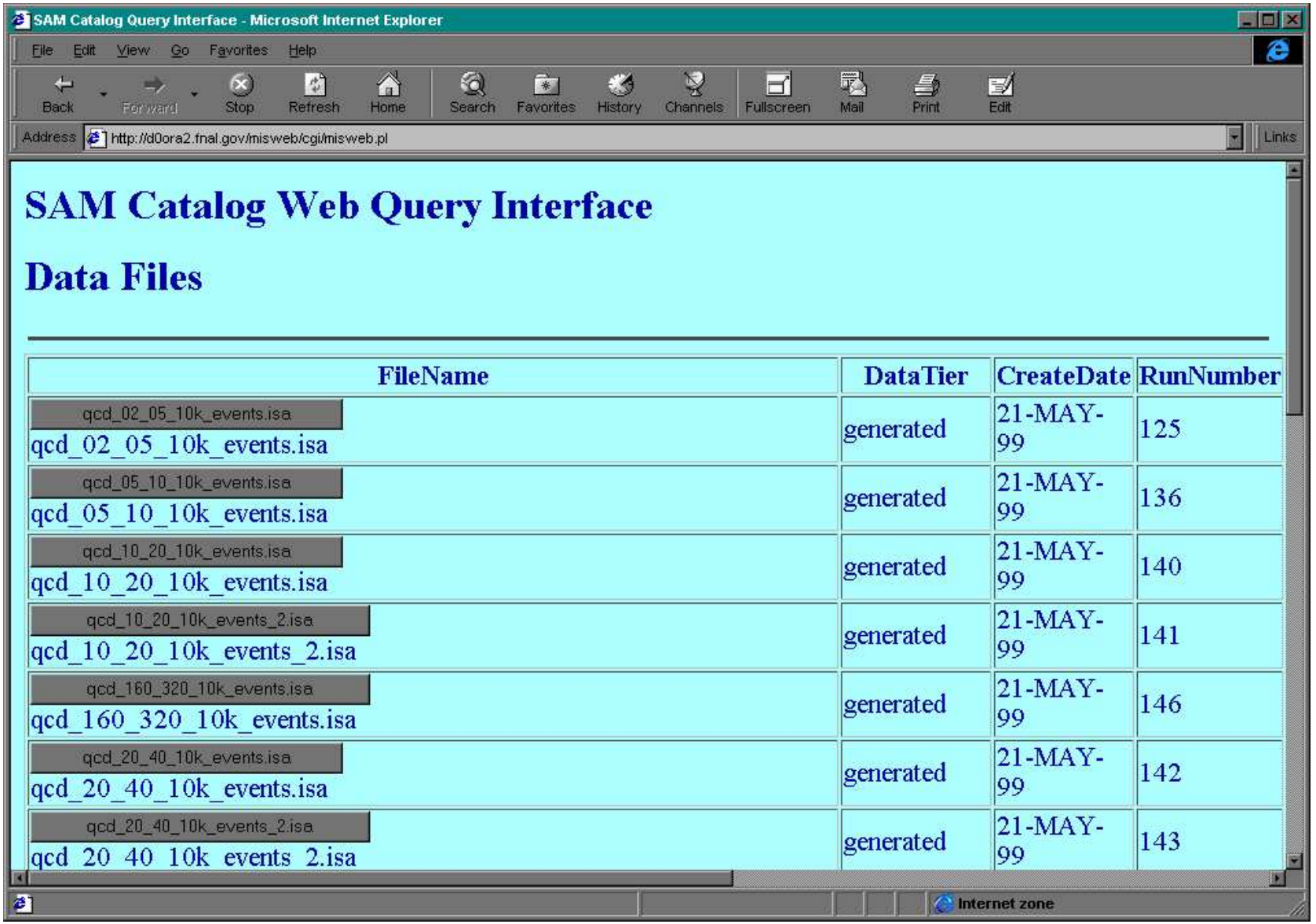
- Greg Graham has created Python GUI tool to assist in MC processing.
- Automatically generates the metadata (description files) to put data into the SAM system.
- Can use
  - sam store --descrip=<desc_file> --source=<in_dir> --dest=<outdir>
    ... for FNAL data
  - sam declare --descrip=<desc_file> --loc=<location>
    ... for data shipped to us on tape.

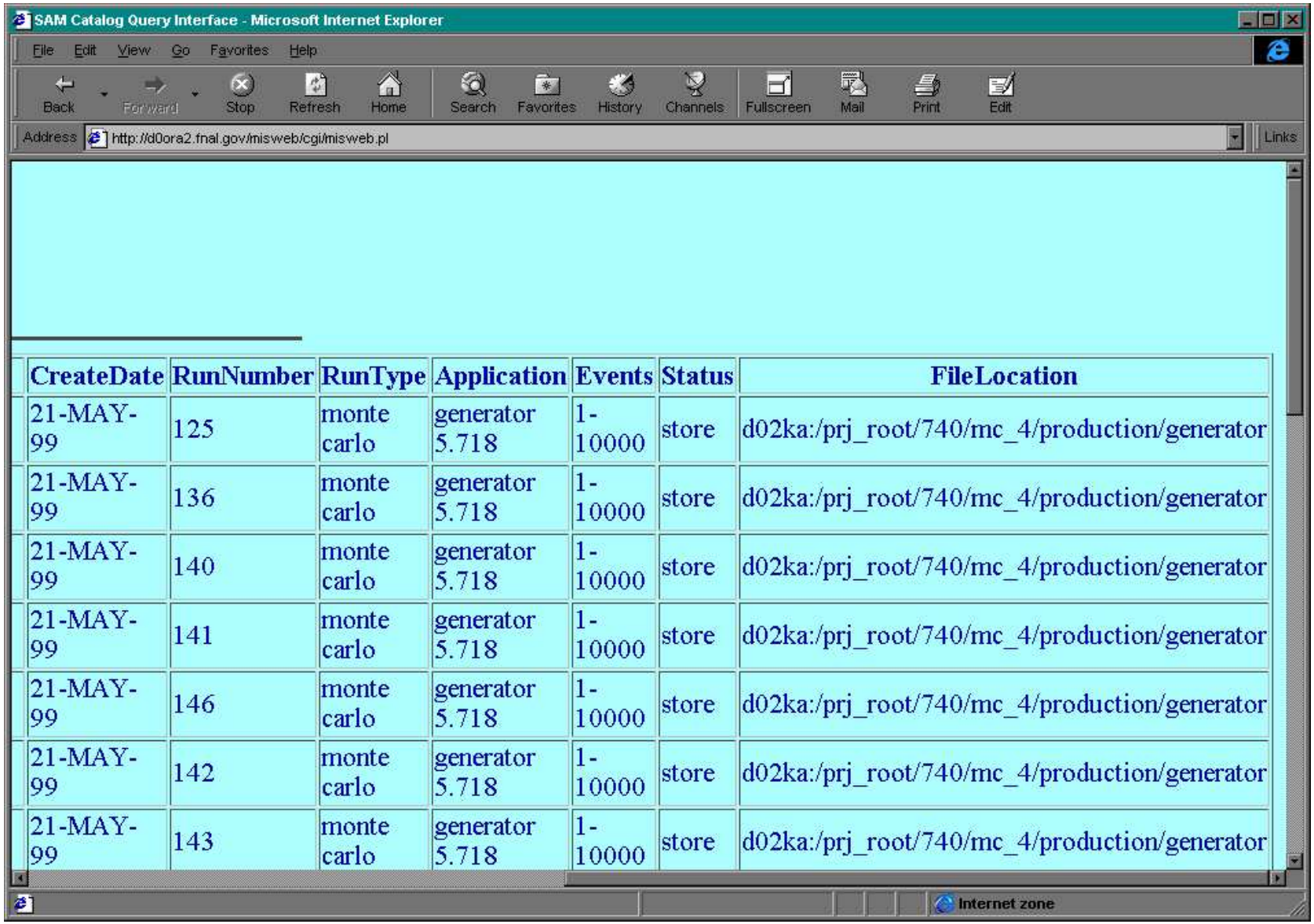



http://d0ora2.fnal.gov/sam/dataBrowsing.html - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit
Address http://d0ora2.fnal.gov/sam/dataBrowsing.html
Links
Quickie Query Lists
Application Families
Data Tiers
Logical Data Stream
Physical Data Stream
Run Types
Stations
File Locations
WorkGroups
SAM Data Browsing
Explore the SAM Data Model
SAM provides dynamic data browsing capabilities available from your web browser. The following data browsing pages allow you to perform various queries against the SAM meta-data. Simply enter the query parameters that you want to use to limit your findings and click Run to query the SAM database. The output of your query will display directly in your web browser.
In some cases, your output will also include buttons on fields in the output. You may click on these buttons to "drill down" deeper into the database and see further relationships that may be pertinent to your original search.
These and other data browsing web pages are easy to build using the MISWEB technology.
Members
Query for all members currently registered in SAM.
Analyzed Files
Query the various files that have been analyzed per project.
Data Files
Query the multitude of data files available in SAM.
Analysis Projects
Query the various analysis projects that are recorded in SAM.
Project Snapshots
Query the various project snapshots that are recorded in SAM.
For help contact sam_support@fnal.gov
Home
Internet zone


SAM Catalog Query Interface - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit
Address http://d0ora2.fnal.gov/misweb/cgi/misweb.pl?tables=application_families&columns=Family,Version&orders=family,version
Links
SAM Catalog Web Query Interface
Family Version
analysis 1
d0reco preco01.02
datalogger 0.0
digitizer p01.00.02
generator 5.718
generator 7.44
Rows 1 to 6 of the Total 6 found.
Back to: Starting Query Page or Edit the SQL query that produced this page.
For help contact sam_support@fnal.gov
Home
MISWEB Query Interface

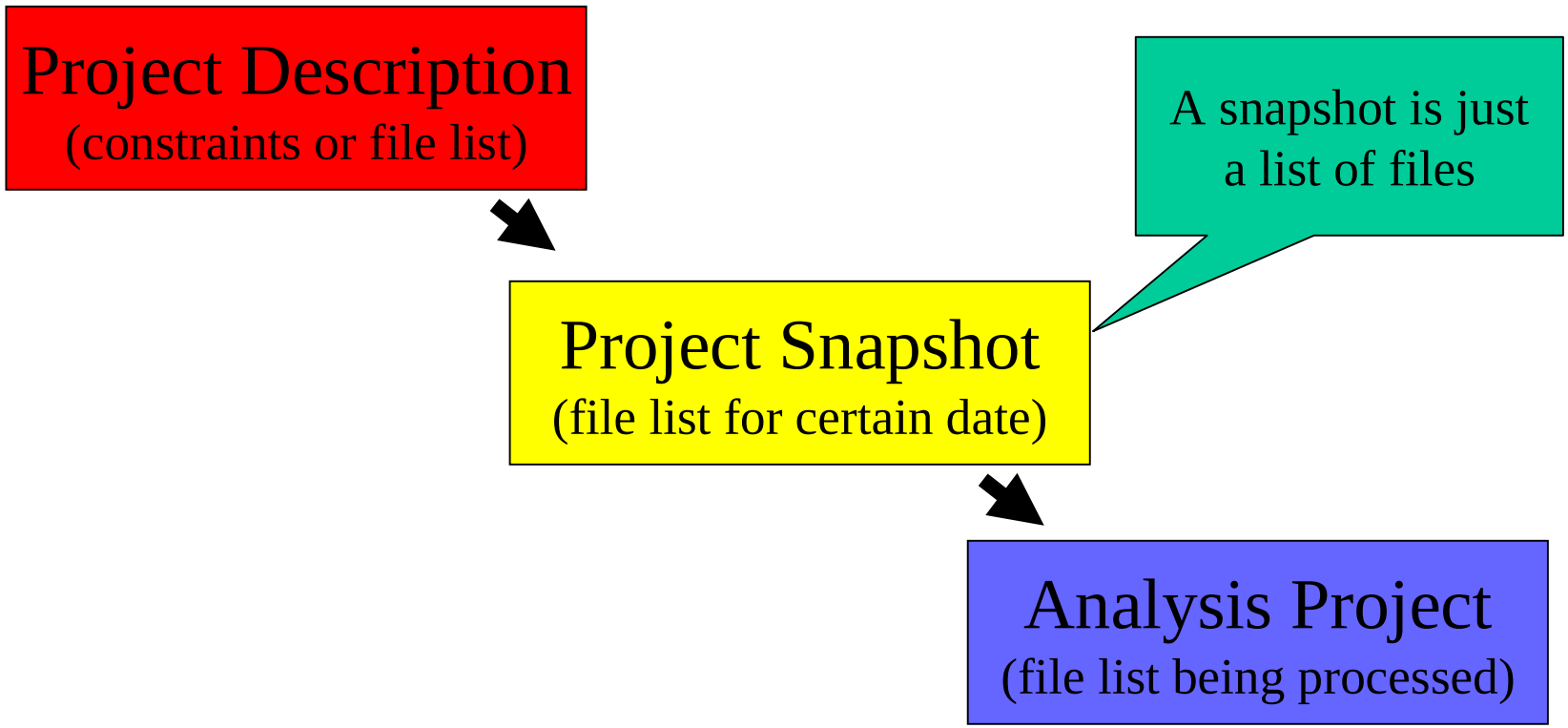

SAM Data Browser - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit
Address http://d0ora2.fnal.gov/sam/data_file_query.html
Links
SAM Data Files Query
Clear Form
Instructions
Run Type:
Run Number:
Data File Name:
%qcd%
Family:
Data Tier:
Created Before: (dd-mon-yy:)Ex:06-JUN-1999
Created After: (dd-mon-yy:)Ex:06-JUN-1999
Fetch Matching Data Files:
Run
Hint: Use percent key (%) for a wildcard. Names are case-sensitive
Query Tips
Construct your own SQL query:
Build
Edit
For help contact sam_support@fnal.gov
Home
Data Files Query
Internet zone


SAM Catalog Query Interface - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit
Address http://d0ora2.fnal.gov/misweb/cgi/misweb.pl
Links
SAM Catalog Web Query Interface
Data Files
FileName
DataTier
CreateDate
RunNumber
qcd_02_05_10k_events.isa
qcd_02_05_10k_events.isa
generated
21-MAY-99
125
qcd_05_10_10k_events.isa
qcd_05_10_10k_events.isa
generated
21-MAY-99
136
qcd_10_20_10k_events.isa
qcd_10_20_10k_events.isa
generated
21-MAY-99
140
qcd_10_20_10k_events_2.isa
qcd_10_20_10k_events_2.isa
generated
21-MAY-99
141
qcd_160_320_10k_events.isa
qcd_160_320_10k_events.isa
generated
21-MAY-99
146
qcd_20_40_10k_events.isa
qcd_20_40_10k_events.isa
generated
21-MAY-99
142
qcd_20_40_10k_events_2.isa
qcd_20_40_10k_events_2.isa
generated
21-MAY-99
143
Internet zone

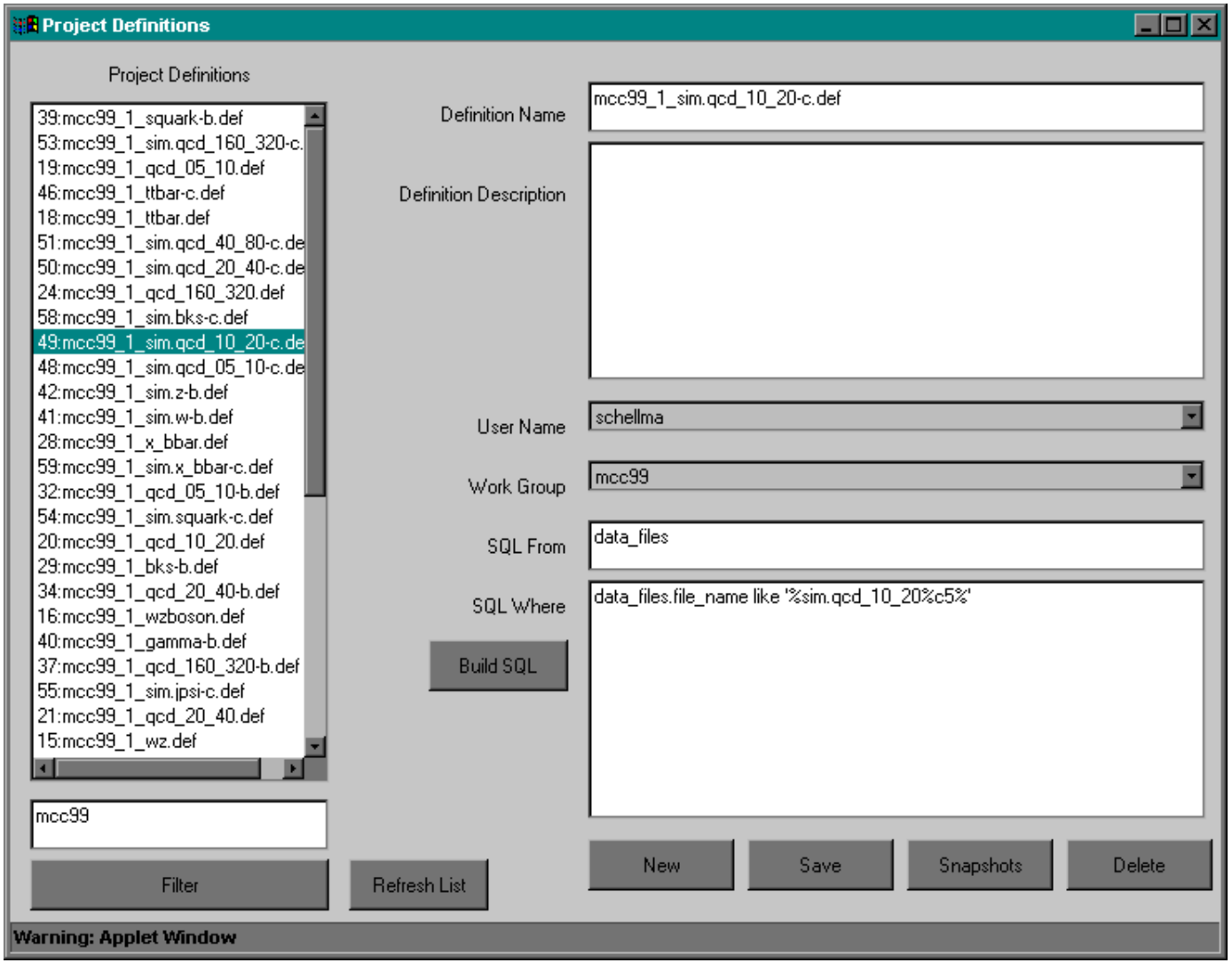


CreateDate	RunNumber	RunType	Application	Events	Status	FileLocation
21-MAY-99	125	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	136	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	140	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	141	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	146	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	142	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator
21-MAY-99	143	monte carlo	generator 5.718	1-10000	store	d02ka:/prj_root/740/mc_4/production/generator

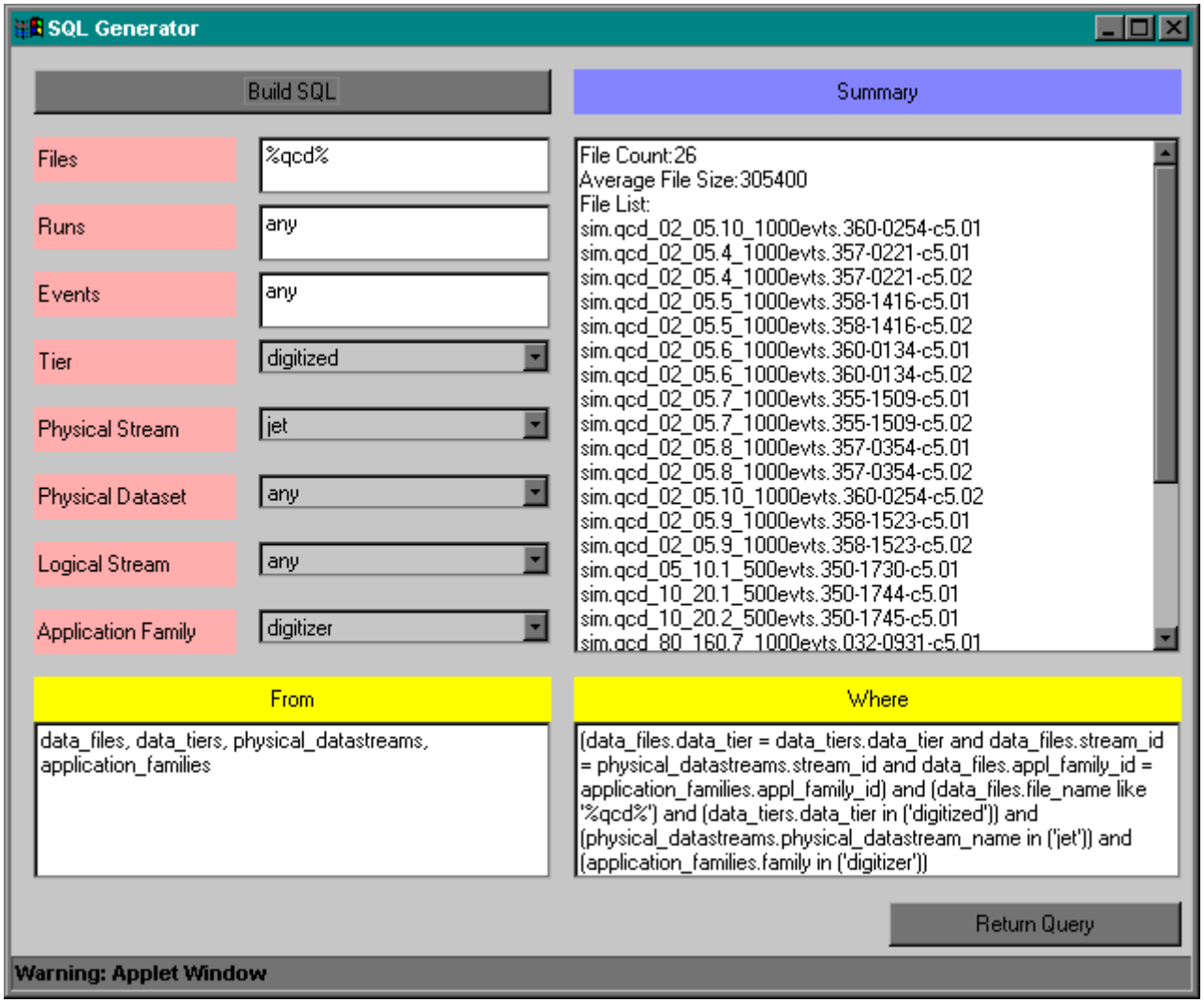
The Life of a Project



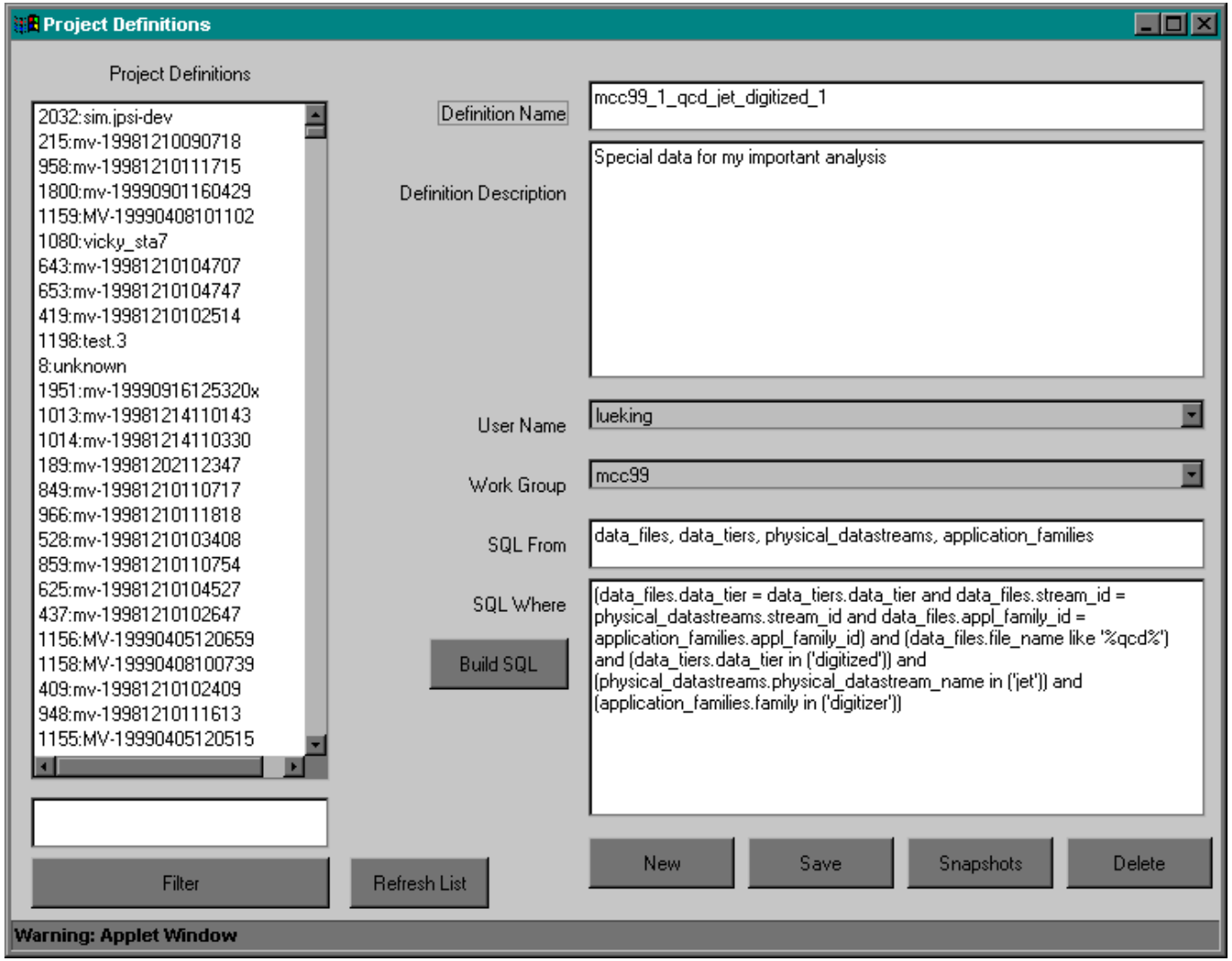
Defining a Project



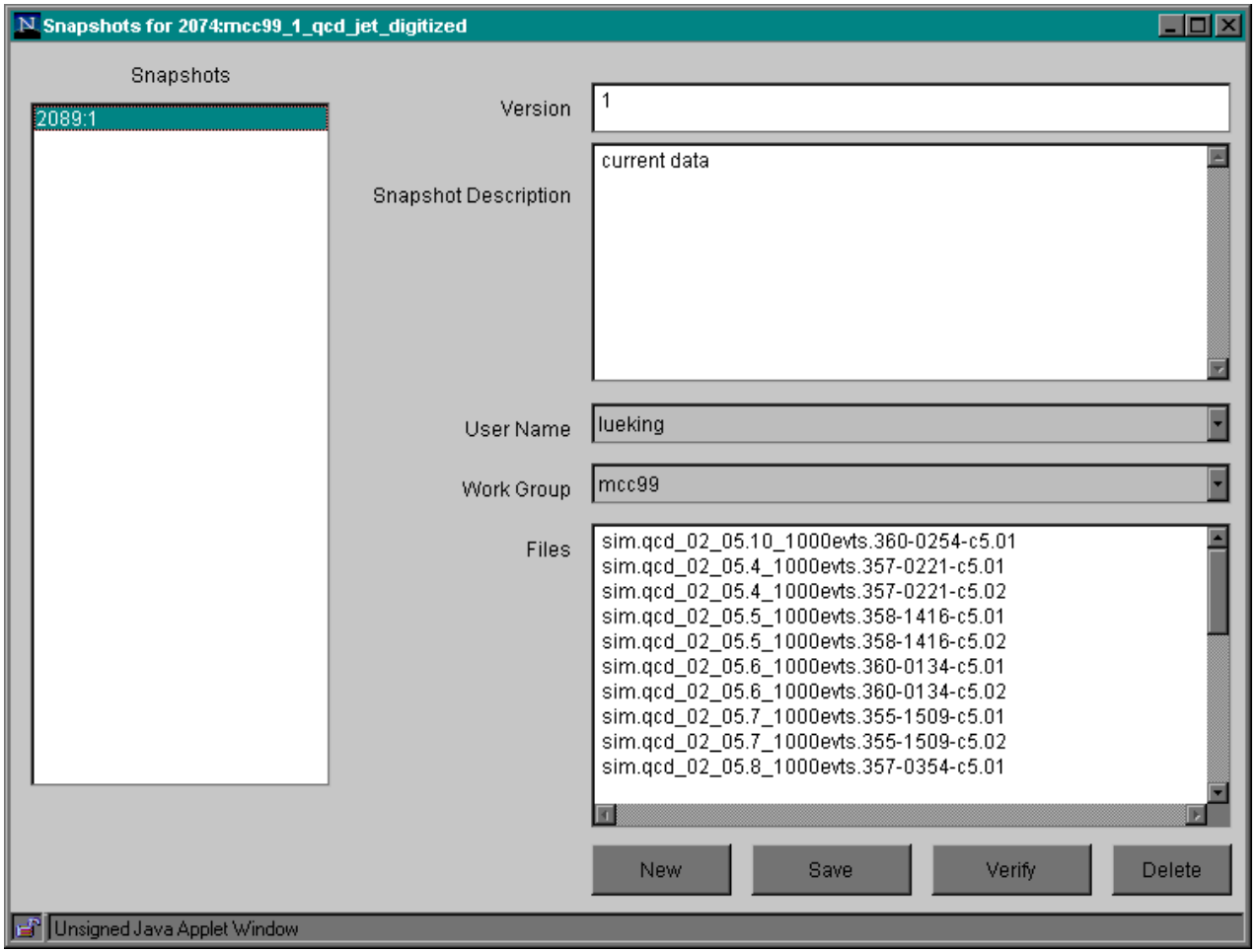
Constraint to SQL Helper



Creating New Definition



Making a Snapshot



Project Using Command Line

- Projects can be created using command line interface
  - **sam translate constraints**
  - **sam resolve query**
  - **sam define project**
  - **sam create snapshot**
  - **sam create analysis project**
  - **sam verify project**

Analyzing Projects: sam run project

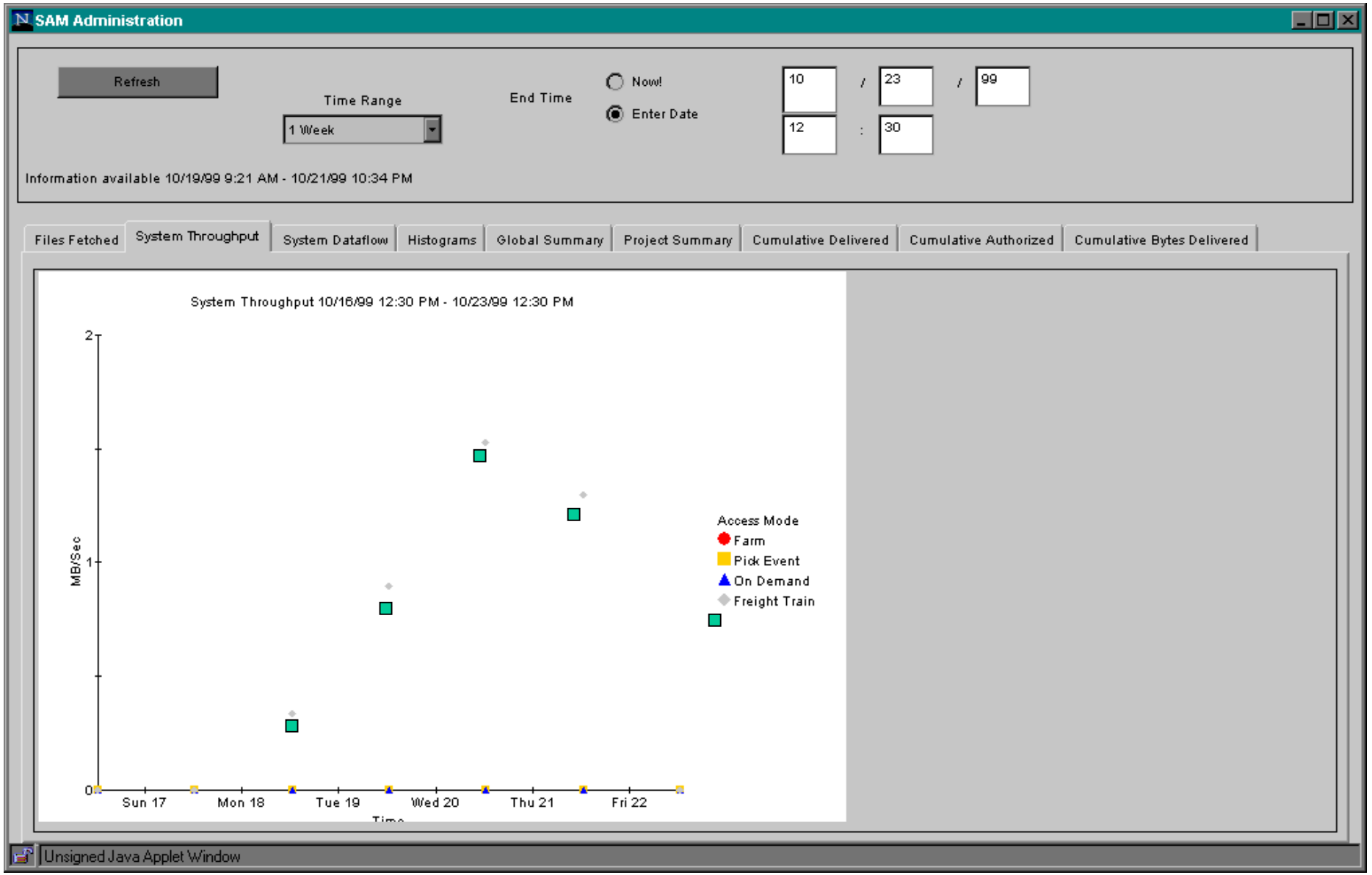
New slide

Analyzing Projects

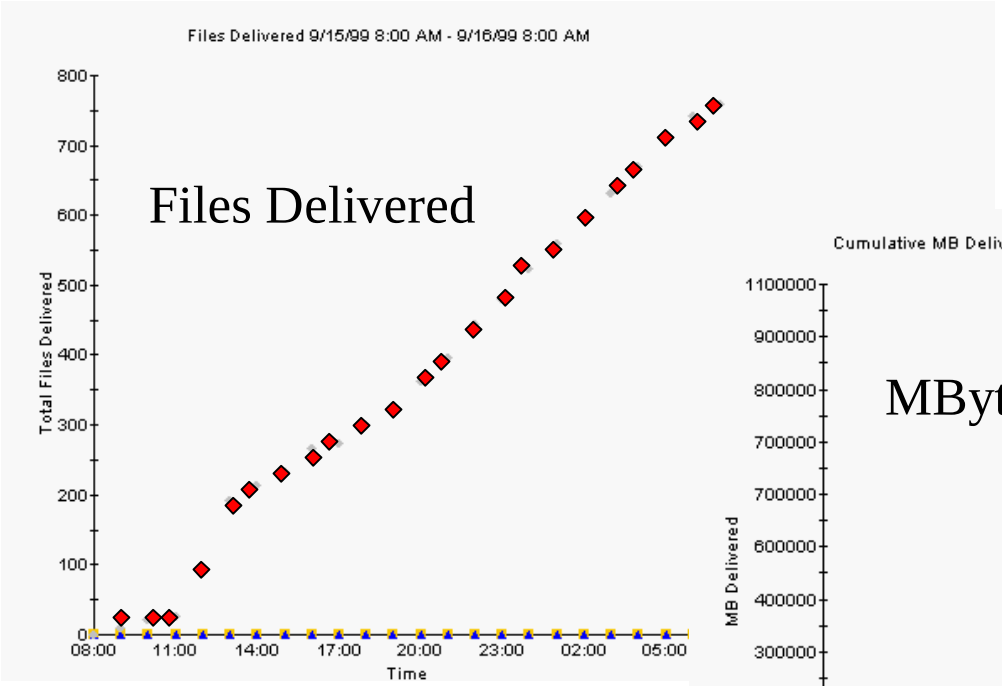
- Method 1: With Python script
  - Create configuration file project_configuration.py
  - Sam run project project_configuration.py
- Method 2: Using D0 Framework
  - In RCP setup specify InputFile=SAMInput:
  - Short script which starts project, and runs framework app.

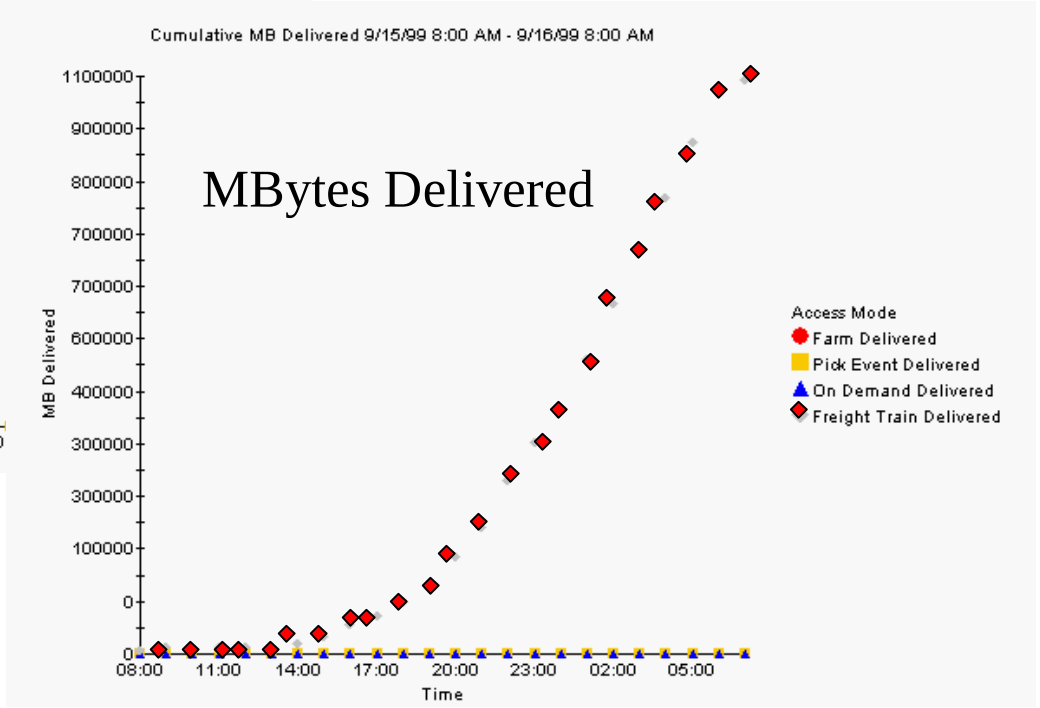
Monitoring and Diagnostics

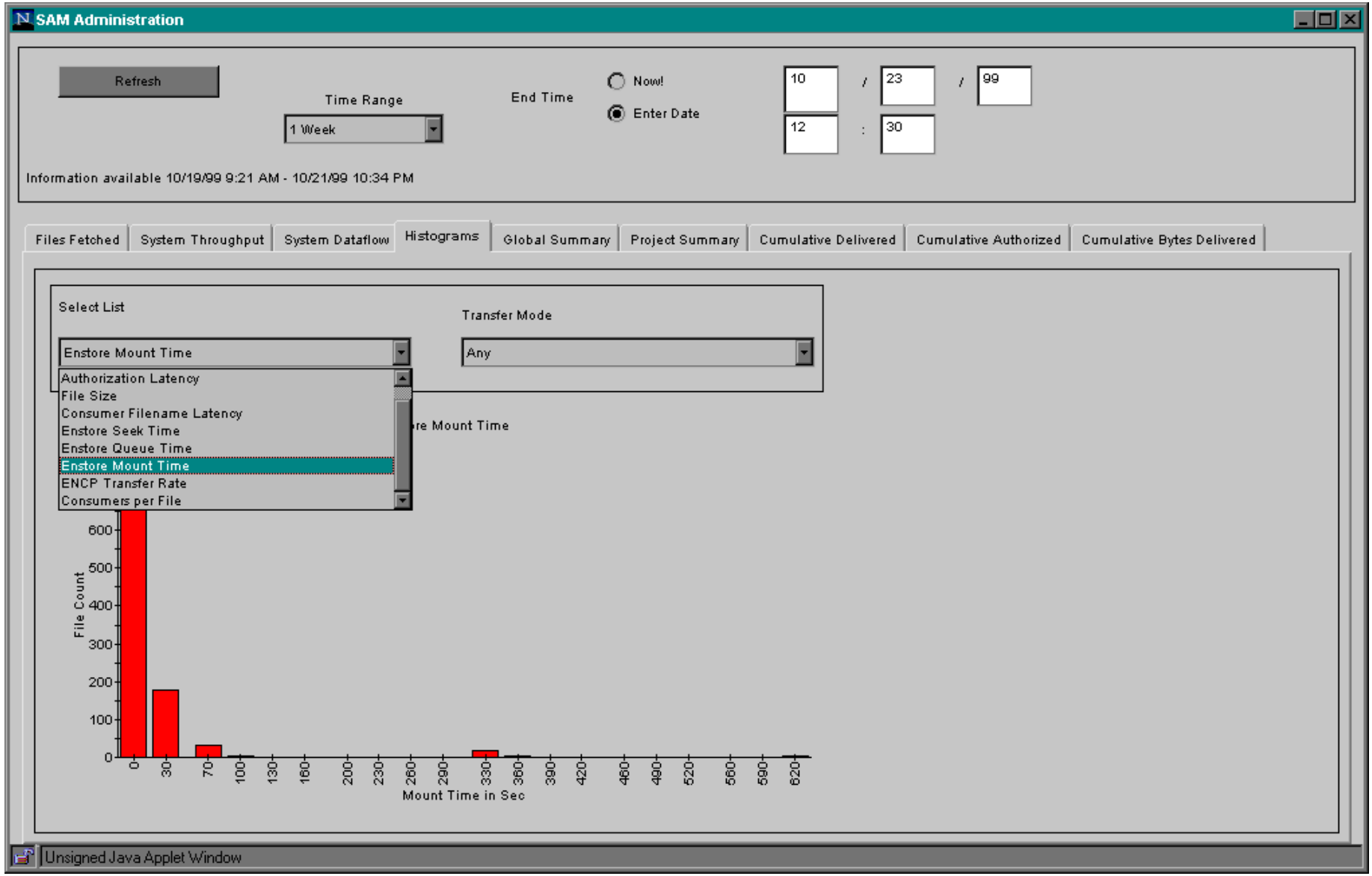
- Sam admin console
- Diagnostic page


SAM Administration
Refresh
Time Range
1 Week
End Time
Now!
Enter Date
10
/
23
/
99
12
:
30
Information available 10/19/99 9:21 AM - 10/21/99 10:34 PM
Files Fetched
System Throughput
System Dataflow
Histograms
Global Summary
Project Summary
Cumulative Delivered
Cumulative Authorized
Cumulative Bytes Delivered
System Throughput 10/16/99 12:30 PM - 10/23/99 12:30 PM
MB/Sec
2
1
0
Sun 17
Mon 18
Tue 19
Wed 20
Thu 21
Fri 22
Access Mode
Farm
Pick Event
On Demand
Freight Train
Unsigned Java Applet Window

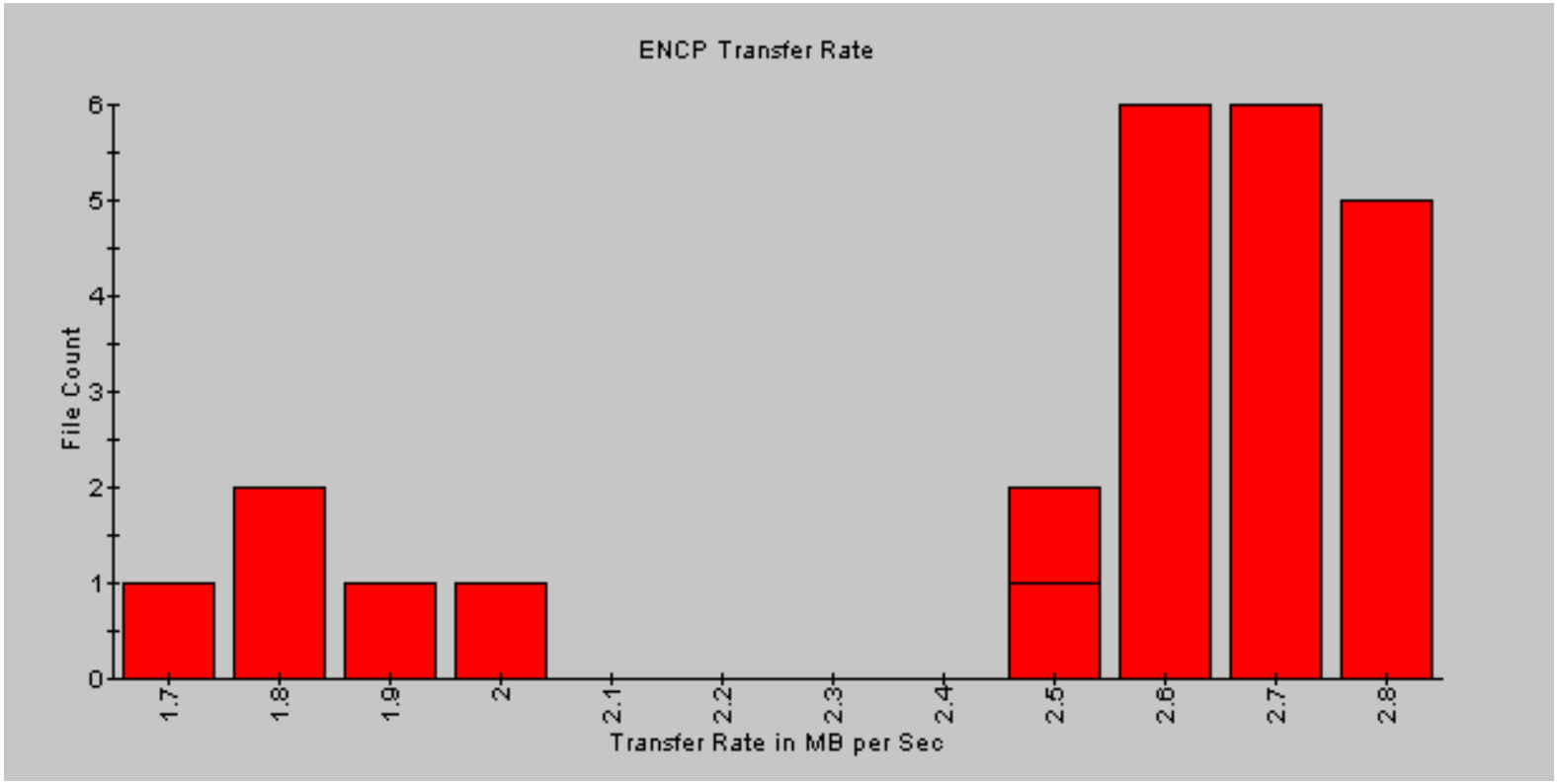
SAM Integration Tests





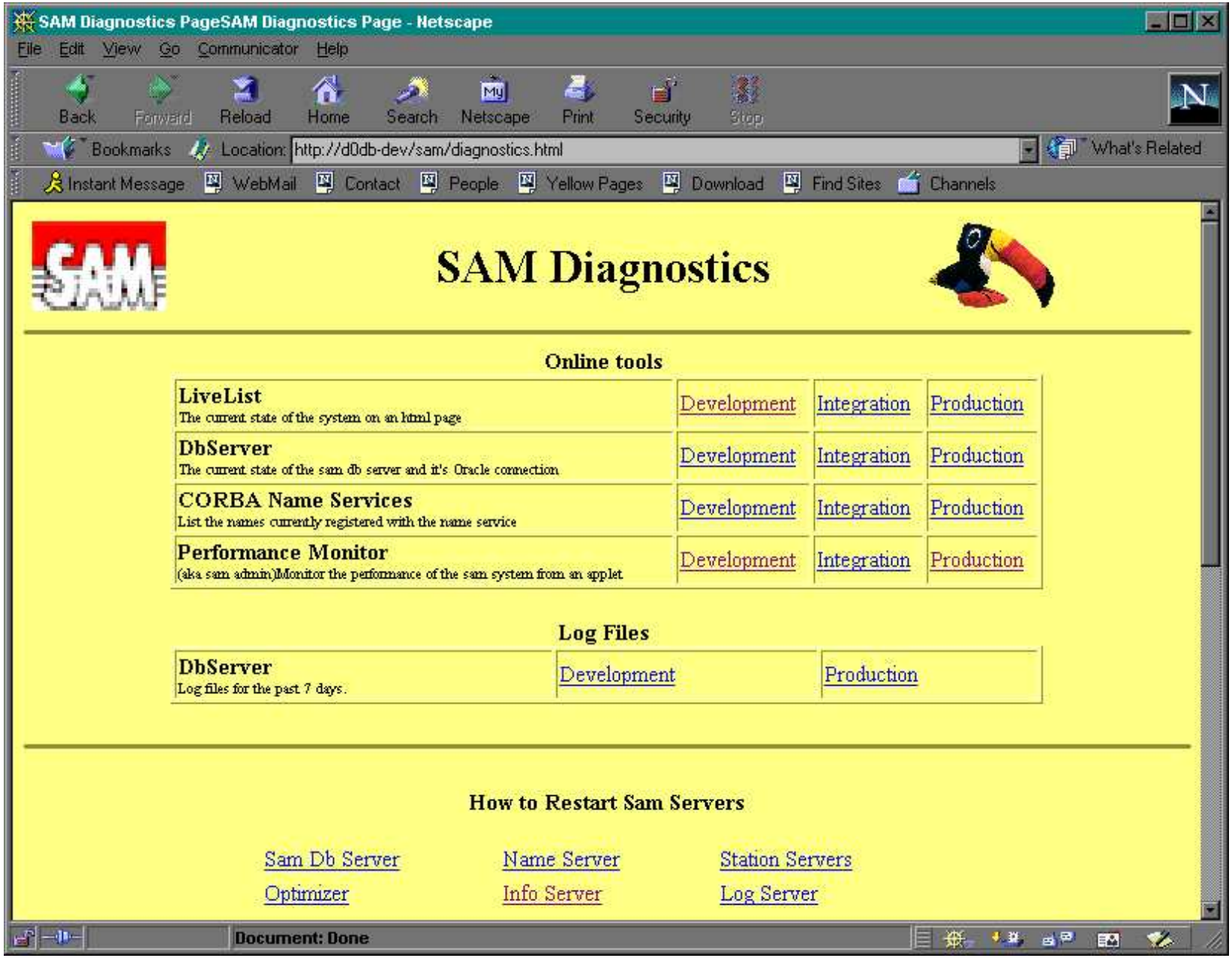

SAM Administration
Refresh
Time Range
1 Week
End Time
Now!
Enter Date
10
23
99
12
30
Information available 10/19/99 9:21 AM - 10/21/99 10:34 PM
Files Fetched
System Throughput
System Dataflow
Histograms
Global Summary
Project Summary
Cumulative Delivered
Cumulative Authorized
Cumulative Bytes Delivered
Select List
Transfer Mode
Enstore Mount Time
Any
Authorization Latency
File Size
Consumer Filename Latency
Enstore Seek Time
Enstore Queue Time
Enstore Mount Time
ENCP Transfer Rate
Consumers per File
File Count
Mount Time in Sec
Unsigned Java Applet Window

SAM- Integration Tests



SAM- Integration Tests

Station Name	Project Name	Cnsmr Counts	Mode	Succ. File Xfers	Avg/ StdDev ENCP Latency	Avg/ StdDev File Size	Avg System Thruput	Inst. Thruput MB/Sec
d02ka	simple-test-1-9_99-22.53.15	0	Freight Train	1	711.49/1006.21	396.0/0.0	0.110	2.631
d02ka	simple-test-1-9_99-04.12.33	0	Freight Train	1	1035.67/1464.66	347.0/0.0	0.097	2.696
d02ka	simple-test-1-9_99-00.56.26	0	Freight Train	2	291.48/546.2	382.0/56.57	0.212	2.281
d02ka	simple-test-1-9_99-23.38.48	0	Freight Train	3	945.46/475.68	391.33/41.3	0.327	2.728
d02ka	simple-test-1-9_99-06.30.11	0	Freight Train	1	956.83/1353.16	344.0/0.0	0.095	2.62
d02ka	simple-test-1-9_99-09.32.19	0	Freight Train	2	372.33/644.89	205.0/289.91	0.114	2.809
d02ka	simple-test-1-9_99-02.31.24	0	Freight Train	2	376.89/78.23	358.0/55.15	0.199	2.694
d02ka	simple-test-3-9_99-04.38.59	0	Freight Train	3	488.74/576.12	568.33/126.06	0.473	2.736
d02ka	simple-test-1-9_99-01.16.52	0	Freight Train	1	275.55/389.69	410.0/0.0	0.114	2.035
d02ka	simple-test-1-9_99-06.10.35	0	Freight Train	1	1725.57/0.0	397.0/0.0	0.110	2.644
d02ka	simple-test-4-9_99-09.32.22	0	Freight Train	1	653.73/0.0	692.0/0.0	0.192	1.86
d02ka	simple-test-3-9_99-01.45.50	0	Freight Train	1	85.67/0.0	727.0/0.0	0.150	2.843
d02ka	simple-test-3-9_99-09.32.21	1	Freight Train	2	163.45/162.17	672.5/48.79	0.373	2.265
d02ka	simple-test-3-9_99-09.04.12	0	Freight Train	1	463.13/0.0	177.0/0.0	0.049	2.764
d02ka	simple-test-4-9_99-05.50.58	0	Freight Train	1	371.92/205.86	682.0/42.43	0.352	2.123
d02ka	simple-test-4-9_99-03.55.39	0	Freight Train	0	0.0/0.0	0.0/0.0	0	0
d02ka	simple-test-4-9_99-06.05.16	0	Freight Train	0	0.0/0.0	0.0/0.0	0	0
d02ka	simple-test-4-9_99-02.43.46	0	Freight Train	2	188.0/0.49	568.5/108.19	0.316	2.702
d02ka	simple-test-4-9_99-08.50.11	0	Freight Train	1	151.02/0.0	652.0/0.0	0.100	2.862
d02ka	simple-test-1-9_99-09.04.12	0	Freight Train	0	0.0/0.0	0.0/0.0	0	0
d02ka	simple-test-1-9_99-04.57.59	0	Freight Train	0	0.0/0.0	0.0/0.0	0	0

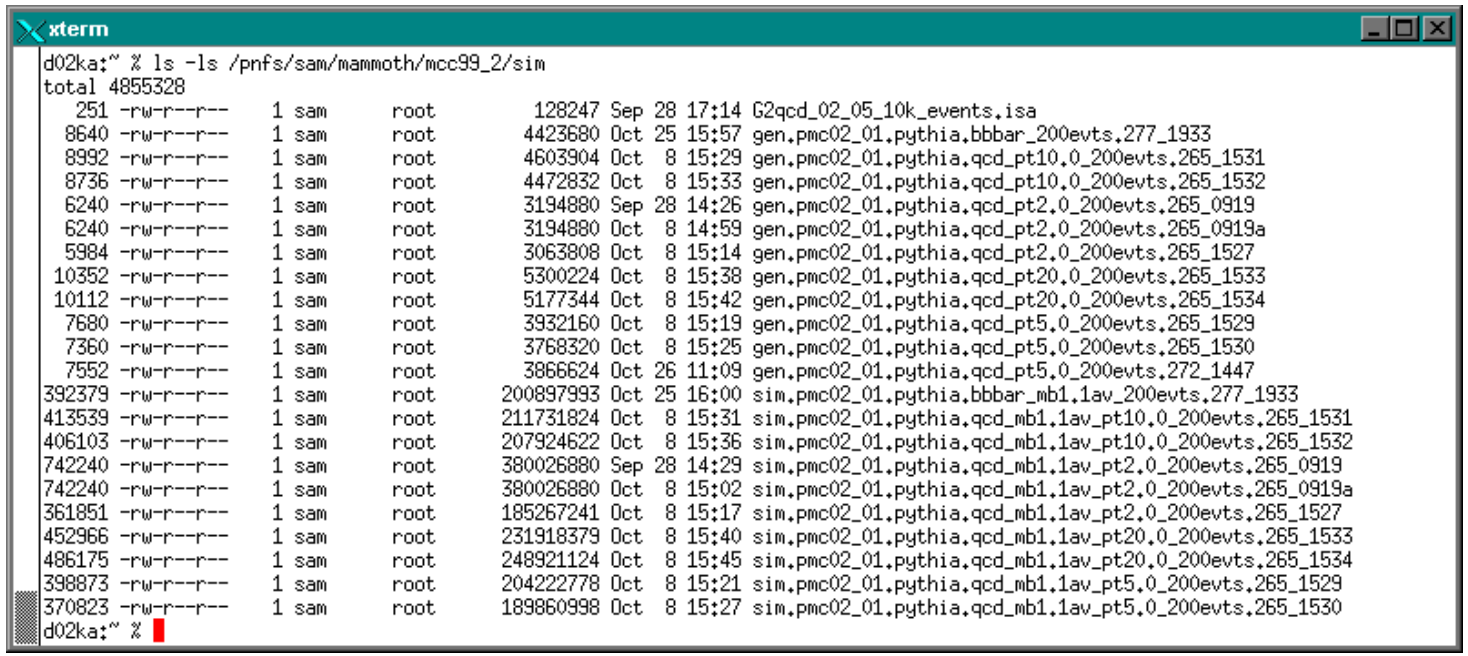

SAM Diagnostics PageSAM Diagnostics Page - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Stop
Bookmarks Location: http://d0db-dev/sam/diagnostics.html What's Related
Instant Message WebMail Contact People Yellow Pages Download Find Sites Channels
SAM
SAM Diagnostics
Online tools
LiveList
The current state of the system on an html page
Development Integration Production
DbServer
The current state of the sam db server and it's Oracle connection
Development Integration Production
CORBA Name Services
List the names currently registered with the name service
Development Integration Production
Performance Monitor
(aka sam admin)Monitor the performance of the sam system from an applet
Development Integration Production
Log Files
DbServer
Log files for the past 7 days.
Development
Production
How to Restart Sam Servers
Sam Db Server
Name Server
Station Servers
Optimizer
Info Server
Log Server
Document: Done

D0 Integration and Users

- SAM is distributed via ups/upd to user and test platforms:
  - Reconstruction farm
  - D02KA
  - D0MINO
  - D0TEST
  - D0Online (not yet)
- Testing done to shake down the system
- Getting users involved
  - V0.0 had several early bird users.
  - V1.0 has several users starting.
  - Will start SAM training the next couple of weeks (Nov 11 General Computing and Software Meeting at D0).

D0 Integration and Users

- Data availability
  - Data: MCC 99 phase I - generated, digitized and reconstructed data
  - Data: MCC 99 phase II - being added as it is generated



User Documentation

- Documentation on the web
- Example user docs in binder
  - Methods for Creating Sam Analysis Projects
  - Project Command Line Interface Users' Guide
  - File Retrieval Using SAM: Running Projects
  - Using SAM within D0 Framework
  - Storing Data Files with SAM
  - Declaring Data File Information and Locations
  - Using the SAM Administrative Console
  - SAM Data Models (and ER diagram)

Data Exchange with non-FNAL Centers

- Small:(~10 GB/week)
  - File transfers over network connection (Amsterdam transfers >1GB/hr)
  - Staging server for remote delivery in V1.5.
- Medium: (~100 GB/week)
  - One or two tapes per week
  - By arrangement with FNAL
- Large: (more than 100 GB/week)
  - Many tapes per week
  - Tapes sent to FNAL can be placed directly into the d0robot, if properly formatted and accompanied with Enstore metadata.
  - Tapes sent from FNAL will have “Enstore/D0 standard” format, unless otherwise arranged.

Journey to Production Operations

- There is a long journey to achieve a production system.
- More robust hardware and software
  - D0Ora1 and RAID array
  - Exabyte Robot and Legato for backups
  - Robustness and recovery features in SAM
- User channels:
    - Mail list sam-users@fnal.gov,
    - Bug tracking using Remedy (like Helpdesk)
    - User training and SAM Users Group
- Data entry and maintenance
  - Keep database up to date, and "clean"
  - Add new supporting data, like versions, users, etc.
- 12/7 and eventually 24/7 support for
    - Production database
    - SAM servers
    - Enstore system

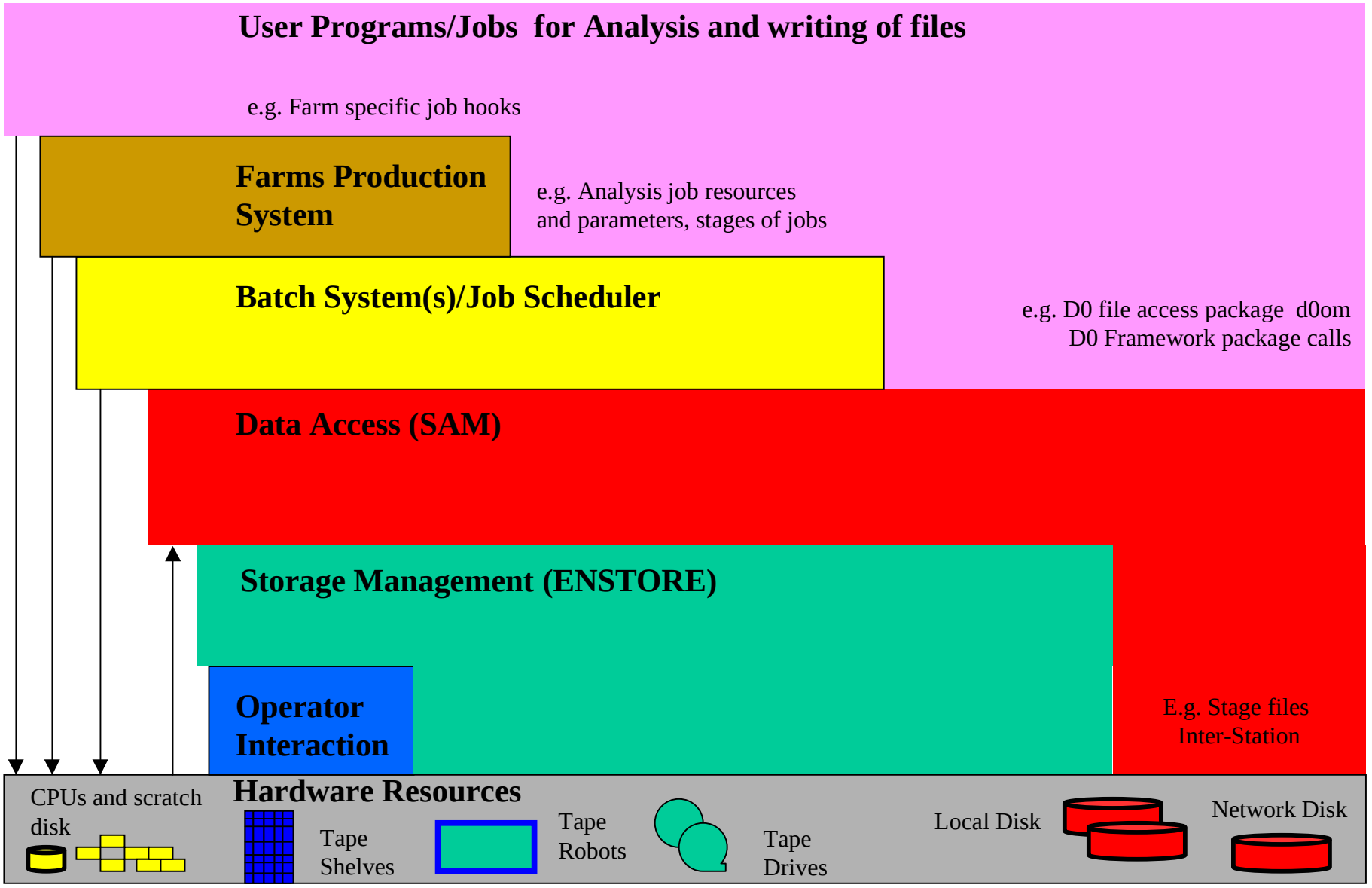
Topics

- Summary of Release Schedule
- Version 1.5 Features
- Code Packages Release
- Who is doing what
- System Functionality
- D0 Integration and Users
- User Documentation
- Moving into Production Operations

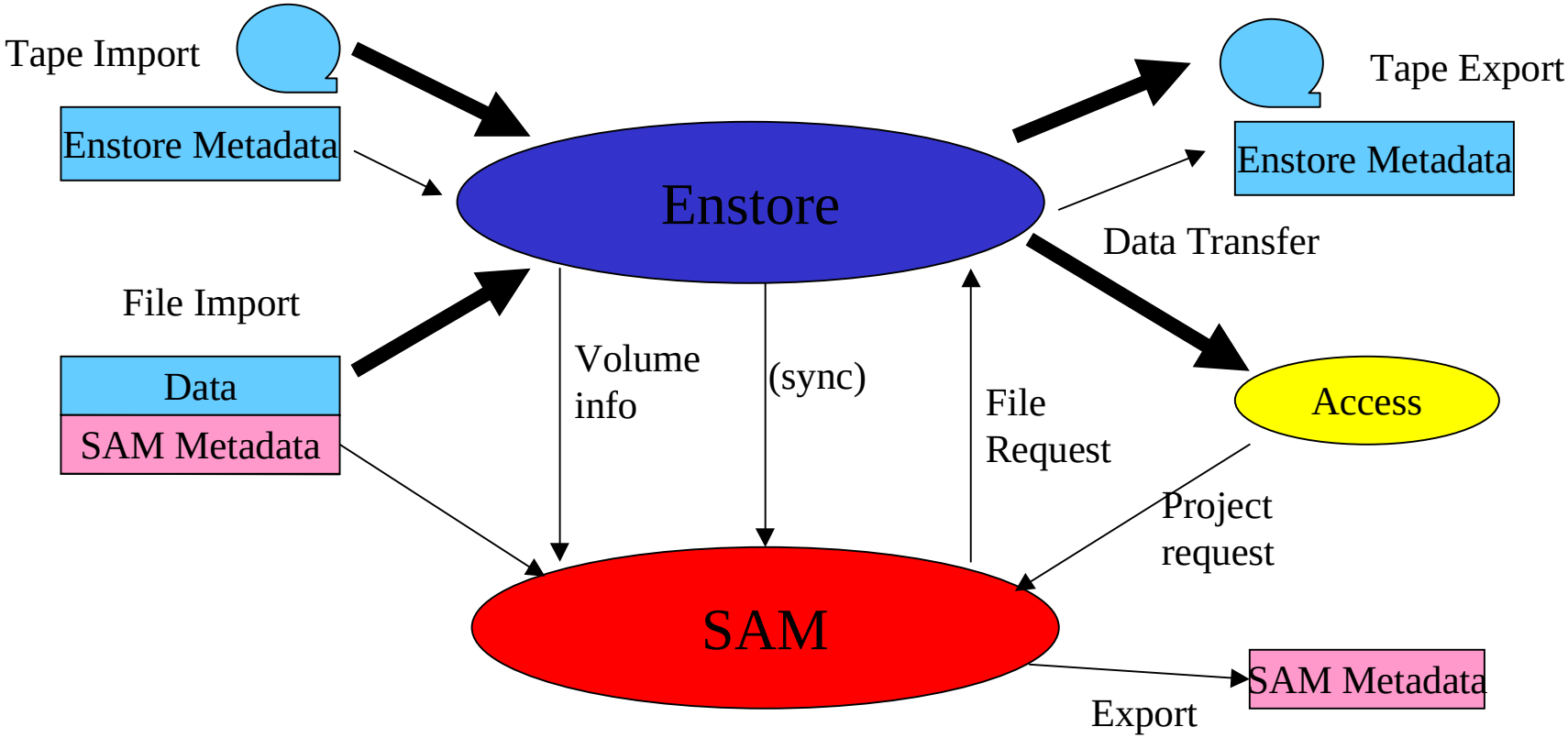
Summary

The slides after this are just for backup

Software Layers for Data Management and Access



Data Import/Access/Export



Enstore Tape Import/Export

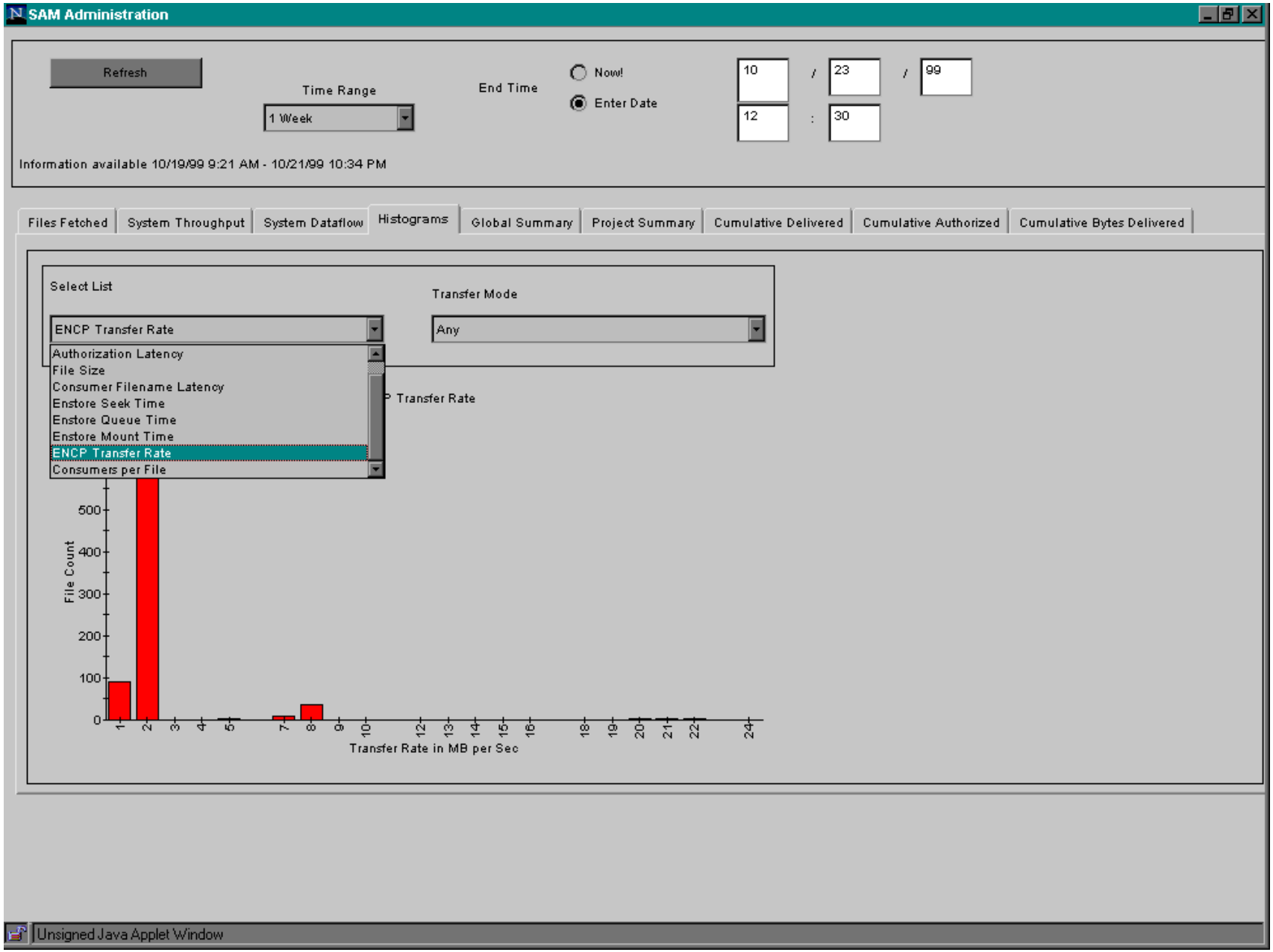
- Remote users can create cpio formatted tapes which can be added directly to the FNAL Enstore system.
- Interface

  enstore_tape init[--tape-device=devname] [--tape-db=dbdir] [--verbose] --volume-label=label

  enstore_tape write [--tape-device=devname] [--tape-db=dbdir] [--verbose] --volume-label=label file_list [file_list ...]

  enstore_tape flatten_db [dbname] | output_file

  perscription for reading tapes using generally available GNU cpio
- More details at: rip8.fnal.gov/enstore/volume_import.html


SAM Administration
Refresh
Time Range
1 Week
End Time
Now!
Enter Date
10
/
23
/
99
12
:
30
Information available 10/19/99 9:21 AM - 10/21/99 10:34 PM
Files Fetched
System Throughput
System Dataflow
Histograms
Global Summary
Project Summary
Cumulative Delivered
Cumulative Authorized
Cumulative Bytes Delivered
Select List
Transfer Mode
ENCP Transfer Rate
Any
Authorization Latency
File Size
Consumer Filename Latency
Enstore Seek Time
Enstore Queue Time
Enstore Mount Time
ENCP Transfer Rate
Consumers per File
Transfer Rate
File Count
0
100
200
300
400
500
1
2
3
4
5
7
8
9
10
12
13
14
15
16
18
19
20
21
22
24
Transfer Rate in MB per Sec
Unsigned Java Applet Window