

# CAS (Central Analysis Server) and GC planning

D. Olson

ANL GC Mtg

27 April 1999

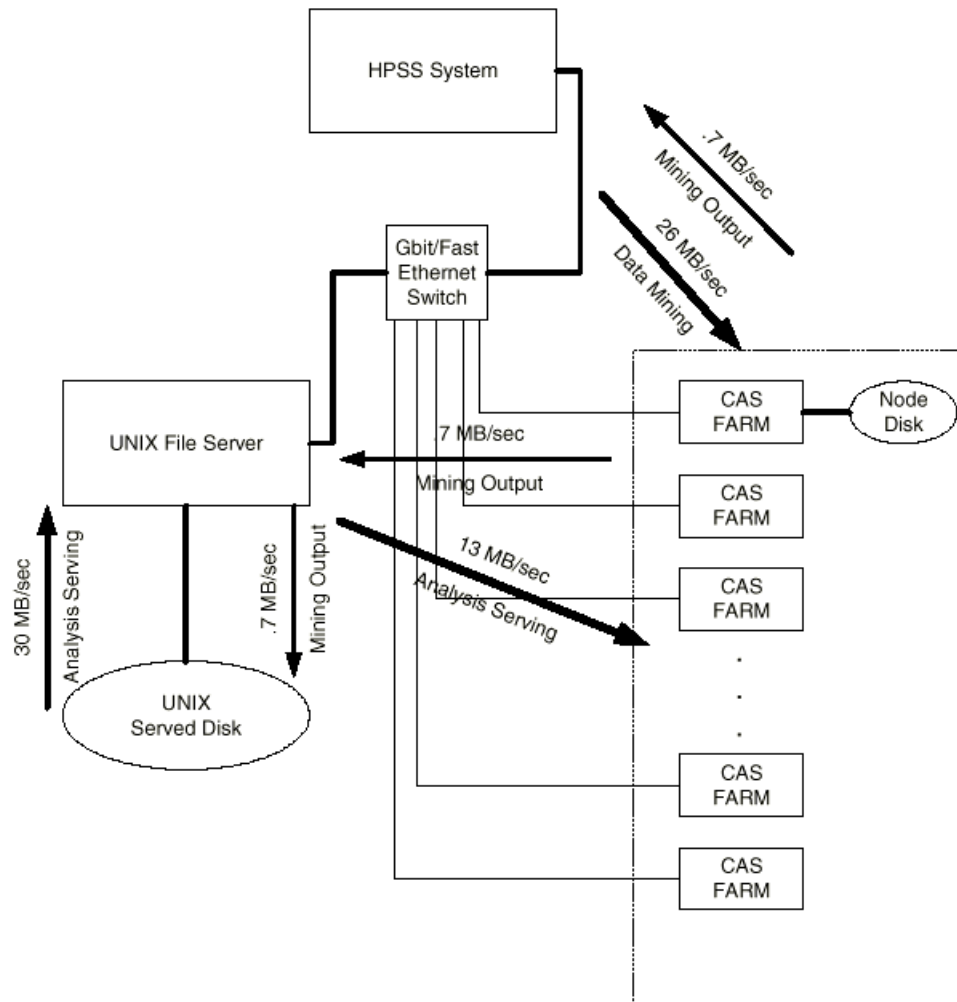
# Outline

- RHIC CAS Task Force
- Doug's presentation to task force
- Discussion

# CAS Task Force

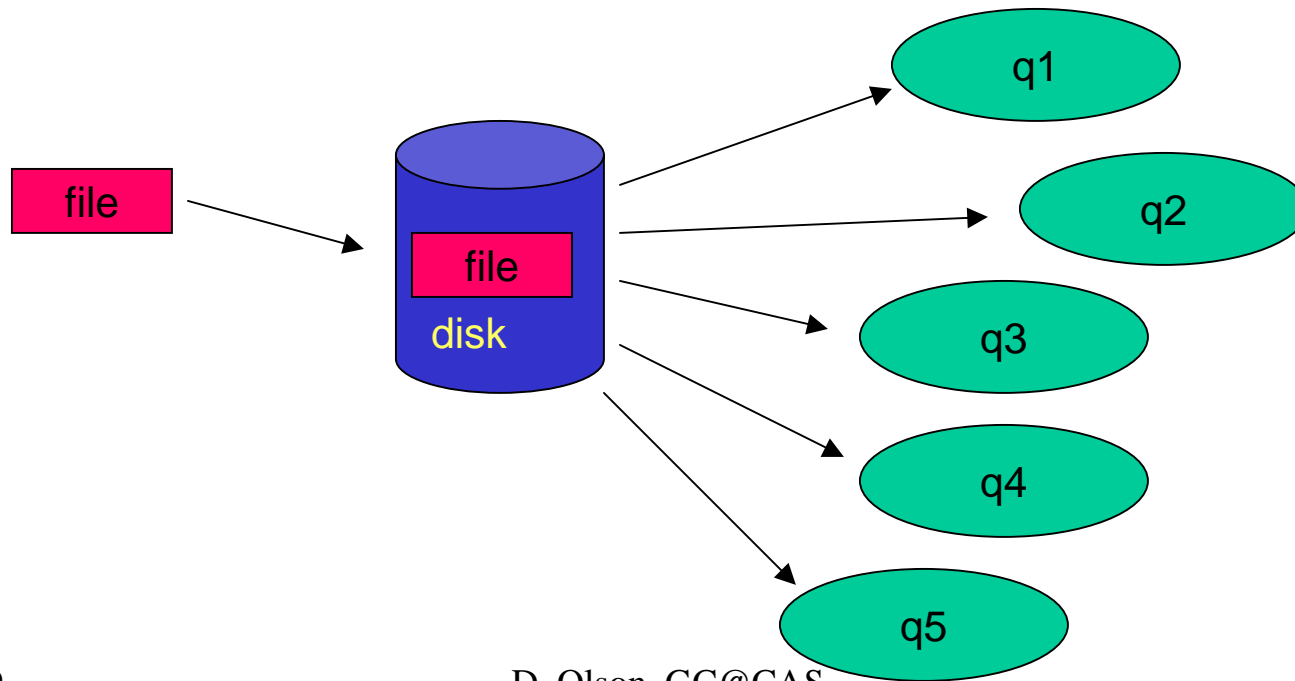
- A committee with reps from RCF, RHIC experiments and GC are discussing the setup for the Central Analysis Server.
- April - May 1999 time frame.
- Goal is to define configuration for FY00 (RHIC expt. startup)

# Data Flow for "Average" Experiment (min. node disk)



# The optimization?

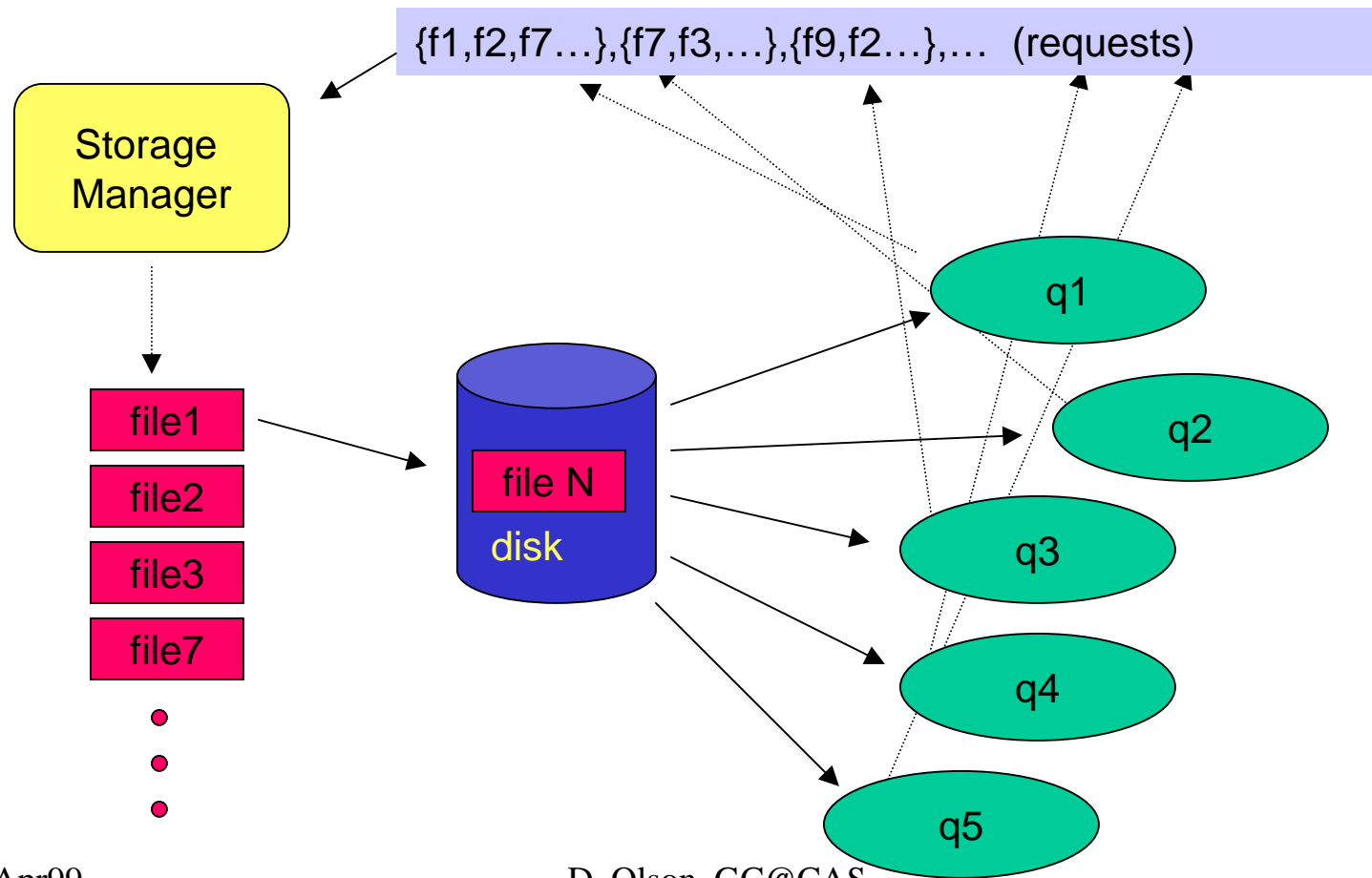
Read once from tape to disk, access by many processes from disk. The disk $\leftrightarrow$ process bandwidth is much greater than the tape $\leftrightarrow$ disk bandwidth. Also, disks can be read very many times. Tapes can be read a finite number of times due to media and head wear.



# How to optimize?

- Each process (query) should request many files.
- Storage manager optimizes order of file delivery and access by processes to minimize number of tape reads.

# How to optimize (2)



# How to optimize (3)

- Figure is simplified view:
  - request is list of interesting physics objects (events & components) and not just files
  - Storage manager maps physics objects to files and coordinates across multi-file event components



# What is relation between file on disk and query process?

- Case 1: process reads every byte in file
  - **CPU intensive analysis:**  
if read rate  $\ll$  network bandwidth  
--> put file on local disk (like CRS)
  - **I/O intensive analysis or data mining:**  
if read rate  $\geq$  network bandwidth  
--> run process on shared disk server

# What is relation between file on disk and query process?

- Case 2: process needs few bytes from file
  - **CPU intensive physics analysis**  
read rate < network bandwidth  
leave file on shared network disk  
run process on farm node
  - **I/O intensive physics analysis**  
read rate  $\geq$  network bandwidth  
run process on shared disk server

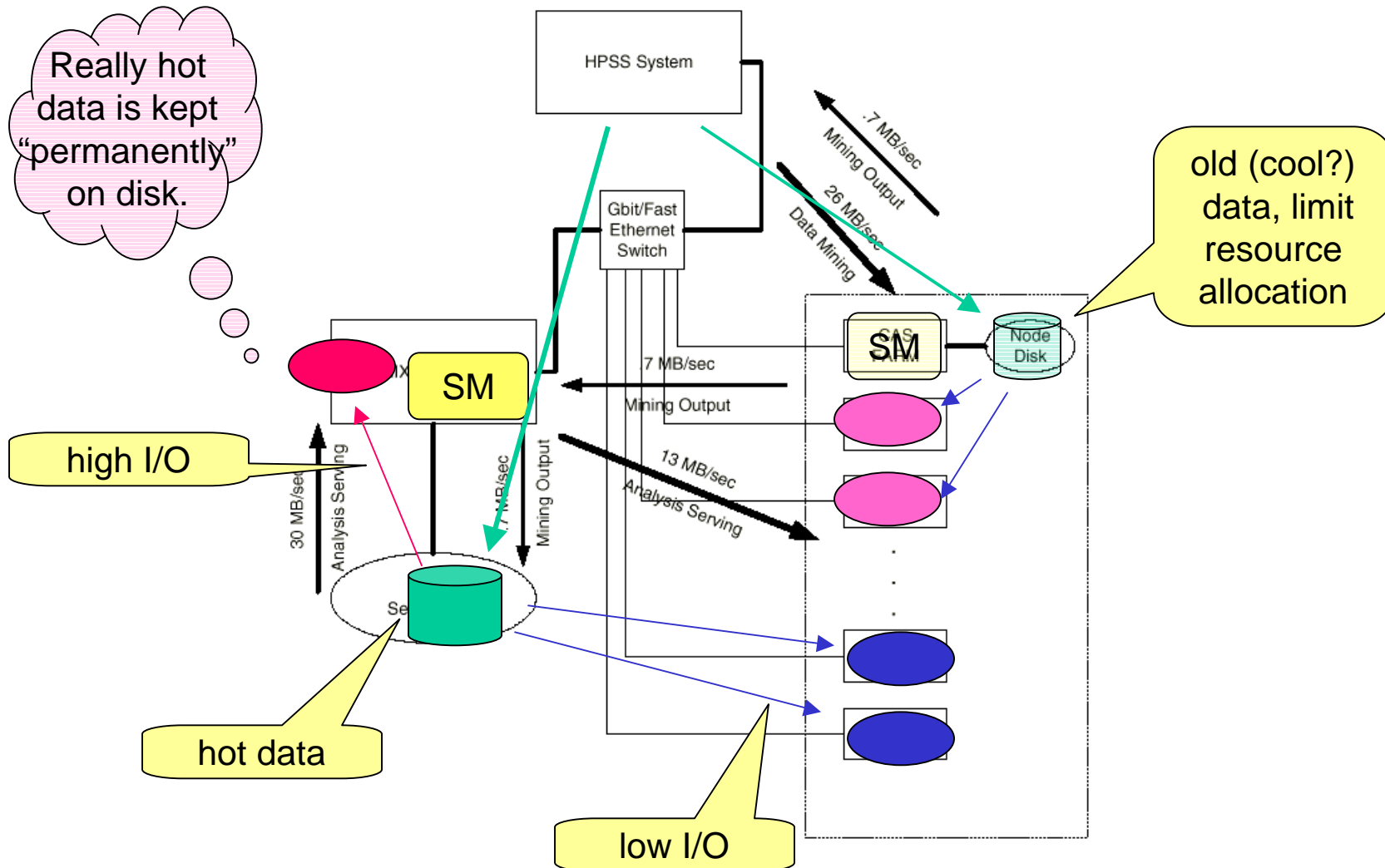
# Characteristics of the shared disk

- How much space?
  - an optimization parameter, depends upon usage
- Use HPSS disk?
  - No. It can not handle fine-grained access.
- Unix server?
  - good for high-bandwidth read
- Use single farm node shared disk?
  - works for low-bandwidth read
- I would start with Unix server disk and also use a farm node as shared disk if necessary.

# What is granularity of a Storage Manager - data system instance?

- The largest set of data likely to be accessed by overlapping queries (analysis codes) and for which the same set of attributes apply.
- Probably corresponds to 1- few macro runs per year, per experiment. I.e., same beam / detector configuration.
- The resource allocated to each instance is disk cache size, which is a configuration parameter.

# Data Flow for "Average" Experiment (min. node disk)



7 April, 1999

CAS Task Force

- B. Gibbard

8

16Apr99

D. Olson, GC@CAS

13