Grand Challenge data management system for MDC1

D. Olson

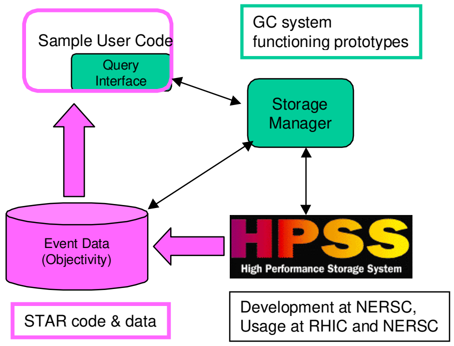
26 July 1998

STAR Collaboration Meeting

Outline

- Description & Status of GC software
- MDC1 - How do you use it?
- What needs to be done?
- Beyond MDC1

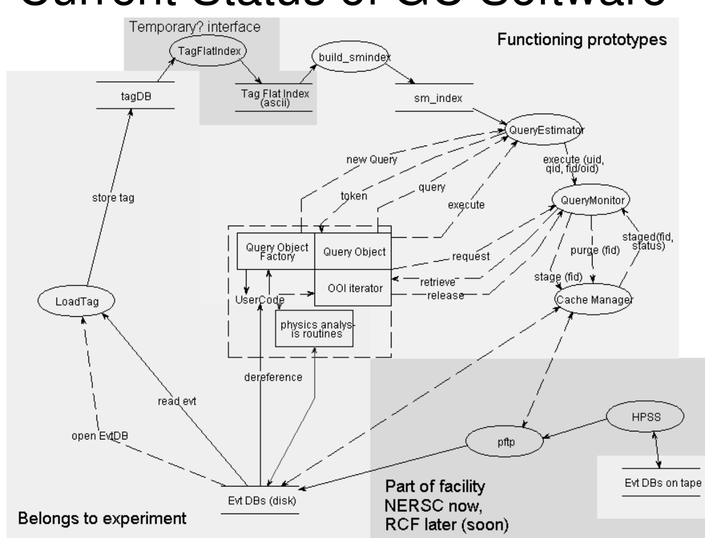
Data access s/w (simple view)



- key developers
  - Henrik Nordberg (NERSC)
    query estimator
  - Alex Sim (NERSC)
    query monitor
  - Luis Bernardo (NERSC)
    cache manager
  - Jeff Porter (LBL-STAR)
    query object
  - Dave Malon (ANL)
    order-optimized iterator & gcaResources API
  - Dave Zimmerman, (LBL-STAR)
    Mark Pollack (BNL-PHENIX)
    tagDB
  - Jie Yang (UCLA,LBL,PRC)
    testing

Machinery to query tagDB & retrieve objects (in files) from HPSS (via pftp).

Current Status of GC Software



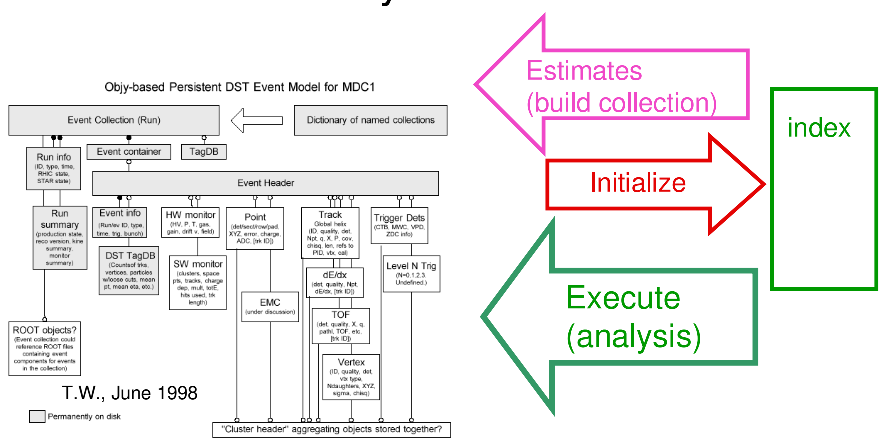
Nuclear Science Division
LBNL

How do you use it?

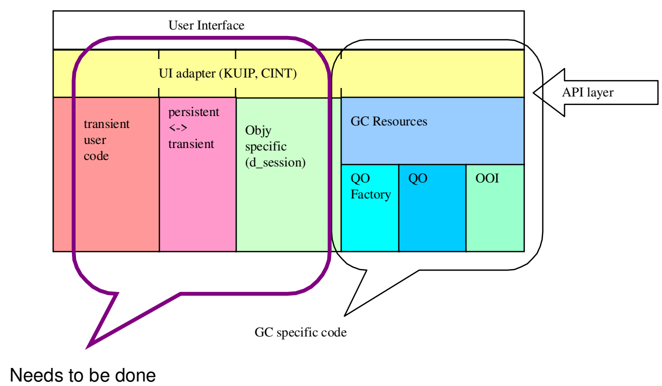
- Initialize system
  - Extract attributes (base & derived) from tagDB
  - Build index for storage manager
- Do query estimates.
  - Number files
  - Time
  - Build OID list
- Do query execution (run analysis)
  - order optimization
  - pre-fetch files
  - iterate as files are returned

Nuclear Science Division
LBNL

How you use it.



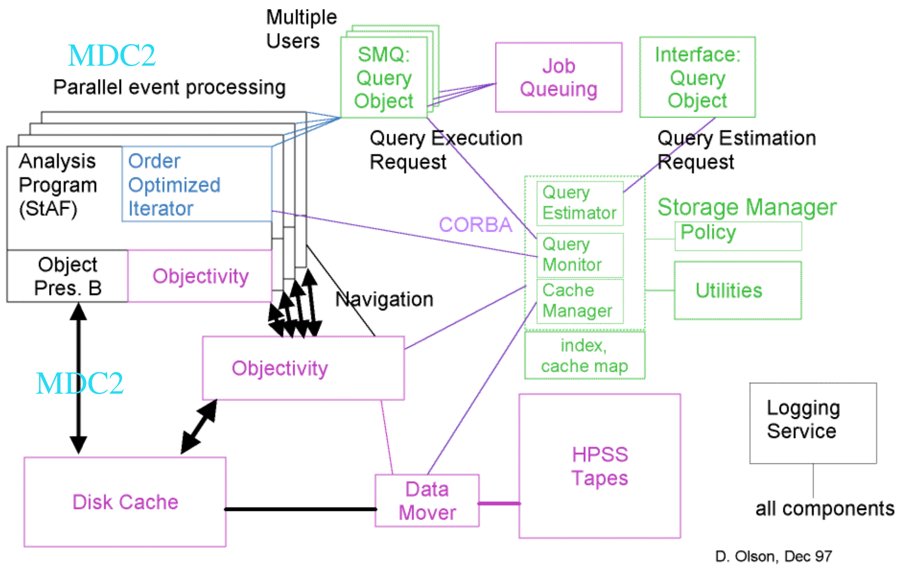
Interfacing to analysis code



In Progress

- Robustness testing, hardening
- Add logging, monitoring of GC components
- Connection to real analysis code (connect Objectivity to analysis code)
- Define MDC scenarios for measurements
- Data mining scenarios of experiments

RHIC Analysis Architecture



Questions that came up during talk

- How do you/we write data?
- How do we use tags?
- We need a clear description during this week how people will use this.
- (others I don't recall)