GSSC LAT Data Server Overview

Tom Stephens
GSSC Database Lead

Outline

- ***Definitions***
- ***Requirements***
- ***Design Goals***
- ***Overall System Architecture***
- ***Implementation Details***
- ***Benchmarks***

Database Definitions

- ***Photon Database (D1ph) – Database that holds all LAT events considered photons and that were used to construct the IRFs. This is the primary science database.***
- ***Event Database (D1ev) – Database that holds (possibly) all reconstructed LAT Events, both photons and particles.***
- ***Pointing and Livetime History Database (D2) – Database that holds spacecraft attitude, position and instrument status information in 30 sec intervals.***

D1 Search Definitions

- ***"Standard" Search***
  - *15° radius circle or 30° x 30° box on the sky for a time period of one year (LESDR 5.2.3.1.1)*
  - *For photon database this is 50-600 MBytes of data depending on sky position*

- ***"Large" Search***
  - *Photon database: Search that would return more than 2GBytes of data per year of observation (LESDR 5.2.3.1.4)*
  - *Event database: Search that would return more than 20Gbytes of data (LESDR 5.2.4.1.4)*

D1 Database Design Requirements

- ***Search Parameters***
  - *Search on values that are real or integer numbers, Booleans, dates and times. (LESDR 5.2.1.1.1)*
  - *Times searchable to microsecond precision (LESDR 5.2.1.1.2)*
  - *2-D positions on sphere (LESDR 5.2.1.1.3)*
  - *Data quality (LESDR 5.2.1.1.4)*
- ***The database must be remotely accessible. (LESDR 5.2.1.3)***
- ***Portability – must not be tied to a single architecture or software system. (LESDR 5.2.1.6)***
- ***HEASARC Compatibility***
  - *Database will be turned over to HEASARC at the end of mission (LESDR 5.2.2.1)*
  - *Must not require excessive effort (>1 FTE) to maintain. (LESDR 5.2.2.1.1)*

Photon Database Performance Requirements

- ***Derived from statistics of current satellite data archives***
- ***Search Speeds***
  - *Standard Search – Data returned within 30 minutes per year of data searched. (LESDR 5.2.3.1.2)*
  - *Standard Search with additional sub-selections – All data returned within 45 minutes per year of data searched. (LESDR 5.2.3.1.3)*
  - *Large Search – All data returned within 3 days. Allows for processing during off peak hours. (LESDR 5.2.3.1.4)*
- ***Number of Requests***
  - *Must perform up to 60 standard searches a day. (LESDR 5.2.3.2.1)*
- ***Data Ingest***
  - *Ingest of new data must be complete within 10 minutes for a 5 hour observation data set (LESDR 5.2.3.3)*
  - *Ingest of reprocessed data may interrupt database access for no more than 60 minutes for a 5 hour observation data set. (LESDR 5.2.3.4)*
- ***Database Restoration***
  - *Must be able to restore database after a crash in <3 days per year of data (LESDR 5.2.3.4.3)*

Event Database Performance Requirements

- ***Search Speeds***
  - *Standard Search – All data returned within 10 hours per year of data searched. (LESDR 5.2.4.1.2)*
  - *Standard Search with additional sub-selections – All data returned within 15 hours per year of data searched. (LESDR 5.2.4.1.3)*
  - *Large Search – All data returned within 7 days. (LESDR 5.2.4.1.4)*
- ***Number of Requests***
  - *Must be able to perform up to 1 standard search a day. (LESDR 5.2.4.2)*
- ***Data Ingest***
  - *Ingest of new data must be complete within 100 minutes for a 5 hour observation data set. (LESDR 5.2.4.3)*
  - *Ingest of reprocessed data may interrupt database access for no more than 10 hours for a 5 hour observation data set. (LESDR 5.2.4.4)*
- ***Database Restoration***
  - *Must be able to restore database after a crash in <1 week per year of data (LESDR 5.2.4.4.3)*
- ***Requirements are generous and design goals provide better performance***

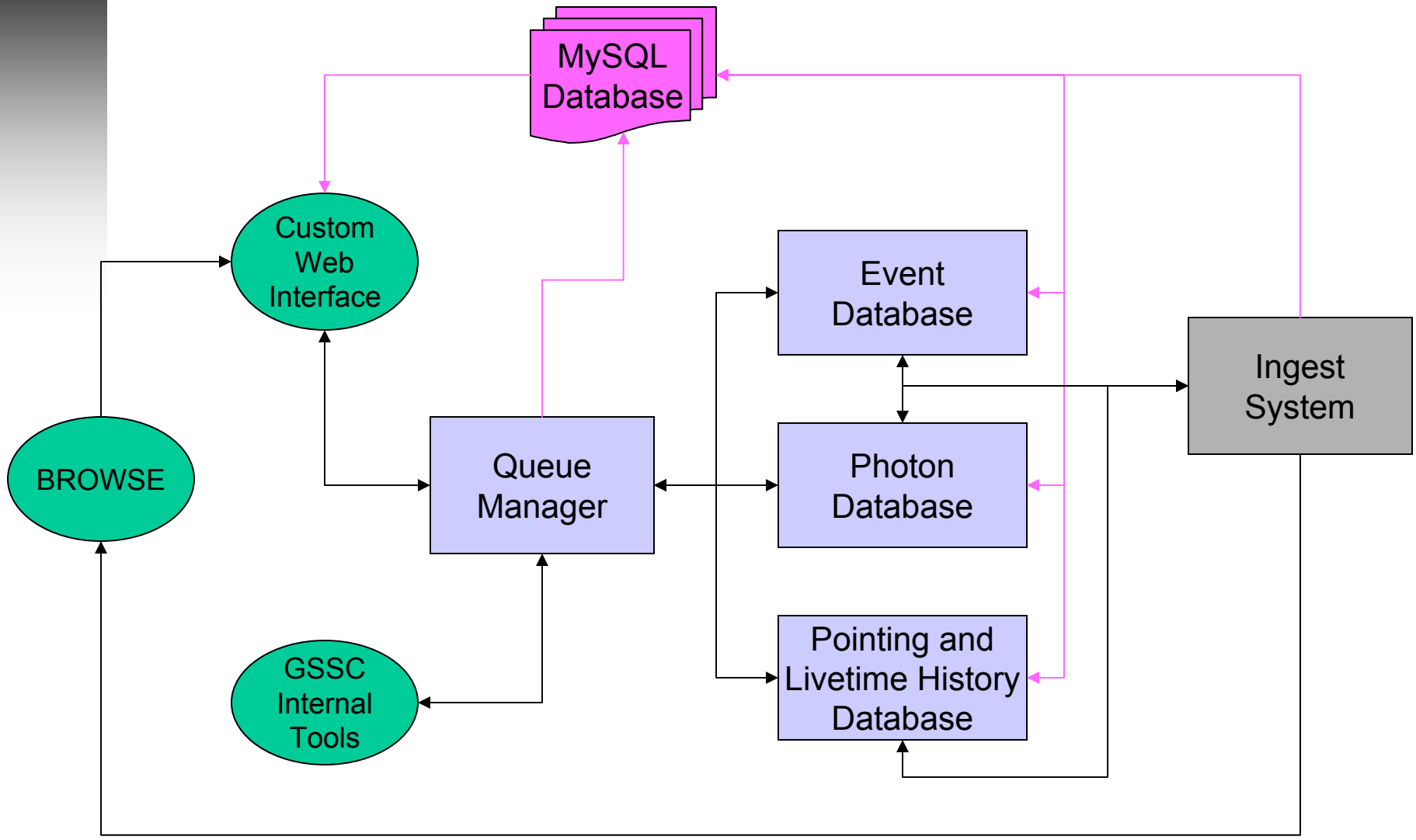
D2 Database Design Requirements

- ***Search Speed***
  - *Retrieve 6 months of consecutive data (~50 MBytes) in 1 minute (SAEDR5.4.1.5.2)*
- ***Number of searches***
  - *Must be able to handle >1500 searches a day (SAEDR 5.4.1.5.2.3)*
- ***Data Ingest***
  - *Ingest of new data (5 hours of spacecraft operation) in 1 minute (SAEDR 5.4.1.5.1)*
  - *Ingest of reprocessed data (5 hour period) in 5 minutes (SAEDR 5.4.1.5.3)*
- ***Database Restoration***
  - *Must be able to restore database after a crash in <1 day (SAEDR 5.4.1.5.4)*

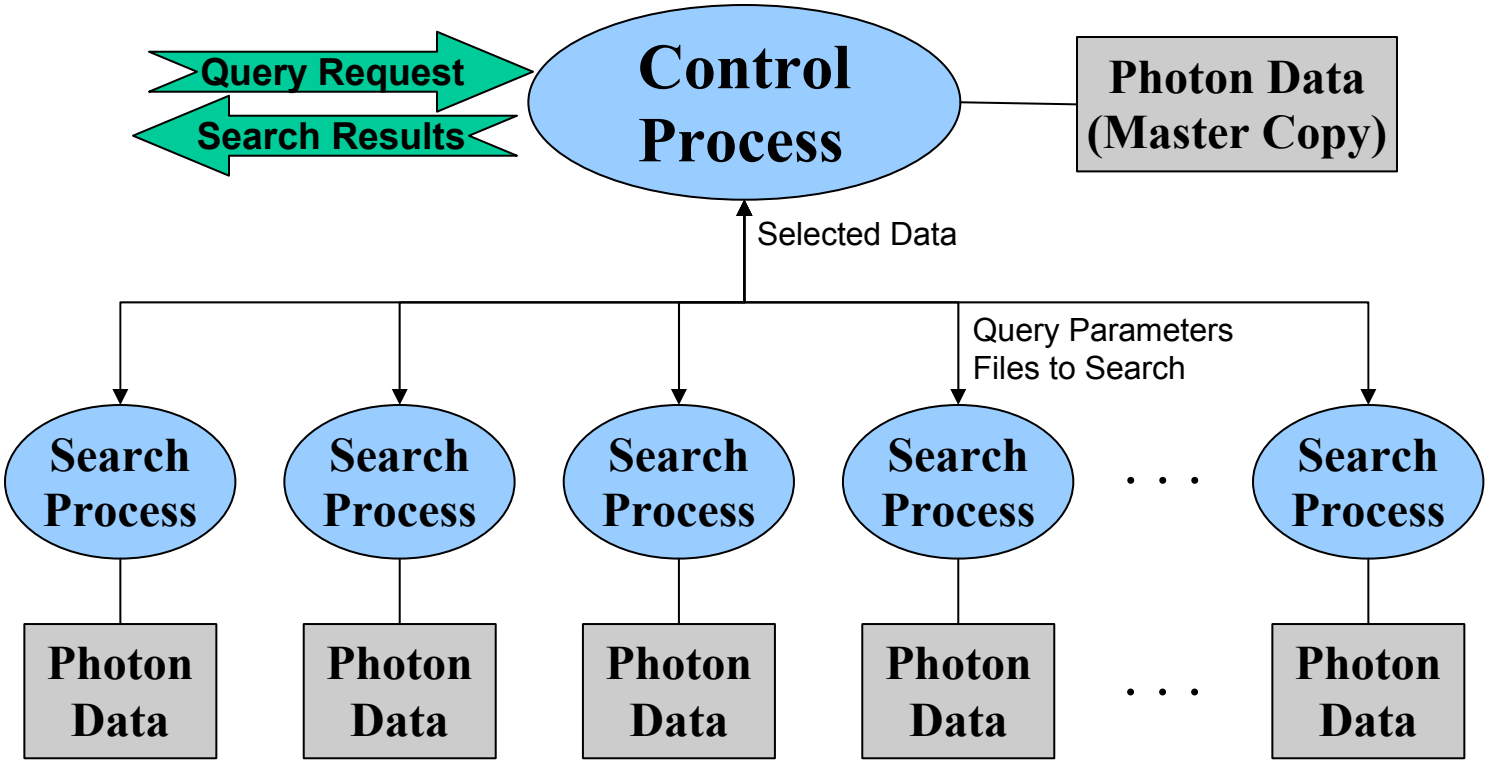
Database Design Goals

	Design Requirement	Design Goals	Current Performance
Standard D1 photon search – 1 year of data	*30 min*	*1 min*	*~40 sec*
Standard D1 event search – 1 year of data	*10 hrs*	*30 min*	*N/T*
D2 search – 6 months of data	*60 sec*	*60 sec*	*7 sec*
D1 photon ingest, new data – 5 hours of data	*10 min*	*2 min*	*0.5-5 min*
D1 event ingest, new data – 5 hours of data	*100 min*	*20 min*	*N/T*
D1 photon ingest, reprocessed data – 5 hours of data	*60 min*	*12 min*	*N/T*
D1 event ingest, reprocessed data – 5 hours of data	*10 hrs*	*2 hrs*	*N/T*
D2 Ingest, new data – 5 hours of data	*1 min*	*1 min*	*10 sec*
D2 Ingest, reprocessed data – 5 hours of data	*5 min*	*1 min*	*N/T*
D1 photon Data Restoration – year of data	*3 days*	*3 hrs*	*10 min*
D1 event Data Restoration – year of data	*7 days*	*3 days*	*N/T*
D2 Data Restoration – entire database	*1 day*	*1 hr*	*10 min*

D1/D2 Database System Design



D1 Photon Database Design



Photon Database Internal Storage

- ***All data is in HEASARC compatible FITS files***
- ***Each node (control and search) has a complete copy of the photon data.***
  - *Fast data access from internal disk*
  - *Multiple backups in case of failure of a single data disk*
- ***Data broken into sky regions and time periods in internal data files***
- ***Hierarchical Triangular Mesh (HTM) used to define regions***
  - *Developed for Sloan Digital Sky Survey at Johns Hopkins*
  - *Recursively divides sky into spherical triangles*
- ***Conducted trade study to determine optimal combination of HTM pixelization level and time binning***
  - *Best time of ~39 sec was level 3 pixelization (512 sky regions) with 2 month time bins*

File Metadata Database I

- ***Currently 3 database tables (will eventually be 7)***
- ***Ingest_data – version information for database***
  - *Database name – (Photon, Event, Spacecraft)*
  - *Start time of current data file (Mission Elapsed Time seconds)*
  - *Current file version – incremented if reprocessed data received for this file, reset to 0 upon creation of new file*
  - *Database version – incremented every time reprocessed data is received. Will allow "roll-back" to earlier version of database if necessary*
- ***Photon_file_comp – what composes the data files***
  - *Filename base – This is a stub that contains the file data start time and version number of the set of data files the input data was added to.*
  - *Input filename – The name of the data file that was ingest*
  - *Ingest date – The data the file was added to the data set.*

File Metadata Database II

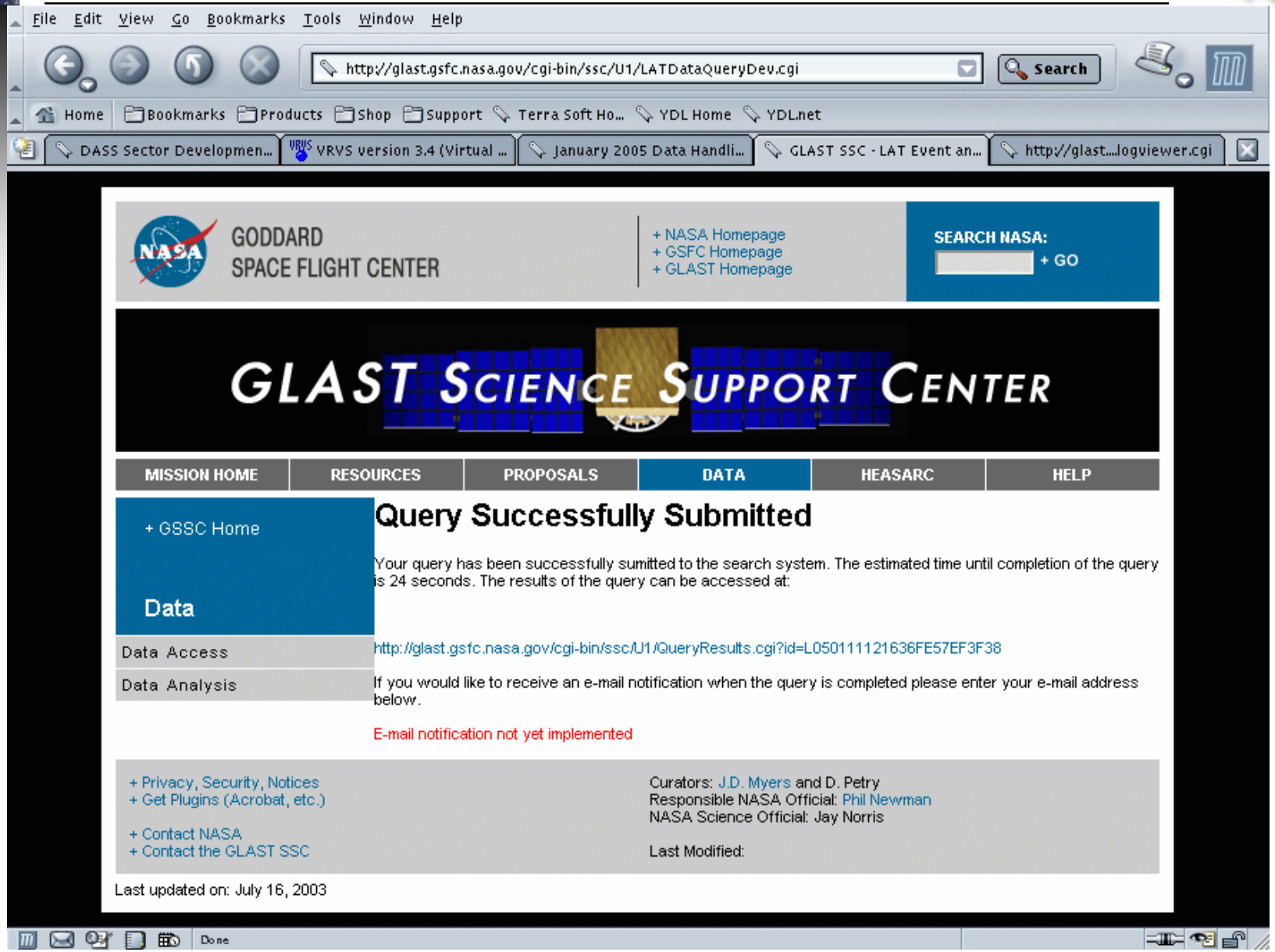
- ***Photon_file_data – What is in the actual files***
  - *Filename – The name of the internal data file*
  - *Date modified – The date and time the file was last modified*
  - *N_photons – The number of photons in the data file*
  - *startTime – The start time of the data file*
  - *stopTime – The end time of the data file*
  - *First_DB_version – The first database version the file is valid in*
  - *Last_DB_version – The last database version the file is valid in*
  - *HTMpixel – the HTM pixel the file corresponds to.*

Sample photon_file_data Entries

filename	modified	n_photons	start_time	stop_time	f_DB_ver	l_DB_ver	HTM_pix
N3321_000000000.811_V01.fits	2004-09-14 23:01:22	78352	0.810755359	5192964.39219536	1	1	N3321
N3321_005192964.422_V01.fits	2004-09-15 23:23:06	75676	5192964.42167179	10380720.8784425	1	1	N3321
N3321_010380720.970_V01.fits	2004-09-16 19:15:23	84248	10380720.9699239	15573604.9799768	1	1	N3321
N3321_015573605.217_V01.fits	2004-09-17 08:52:05	58666	15573605.2165841	20758694.9444996	1	1	N3321
N3321_020758695.407_V01.fits	2004-09-20 14:51:22	86182	20758695.4070489	25956942.8102773	1	1	N3321
N3321_025956943.127_V01.fits	2004-09-21 05:26:23	84268	25956943.1269949	31141488.8384952	1	1	N3321
N3321_031141488.938_V01.fits	2004-09-21 05:40:23	7789	31141488.938296	31539599.4946347	1	1	N3321
S3321_000000000.811_V01.fits	2004-09-14 23:00:30	20638	0.810755359	5192964.39219536	1	1	S3321
S3321_005192964.422_V01.fits	2004-09-15 23:22:15	22436	5192964.42167179	10380720.8784425	1	1	S3321
S3321_010380720.970_V01.fits	2004-09-16 19:14:30	23566	10380720.9699239	15573604.9799768	1	1	S3321
S3321_015573605.217_V01.fits	2004-09-17 08:51:10	19569	15573605.2165841	20758694.9444996	1	1	S3321
S3321_020758695.407_V01.fits	2004-09-20 14:50:24	25107	20758695.4070489	25956942.8102773	1	1	S3321
S3321_025956943.127_V01.fits	2004-09-21 05:25:17	21376	25956943.1269949	31141488.8384952	1	1	S3321
S3321_031141488.938_V01.fits	2004-09-21 05:40:20	2239	31141488.938296	31539599.4946347	1	1	S3321

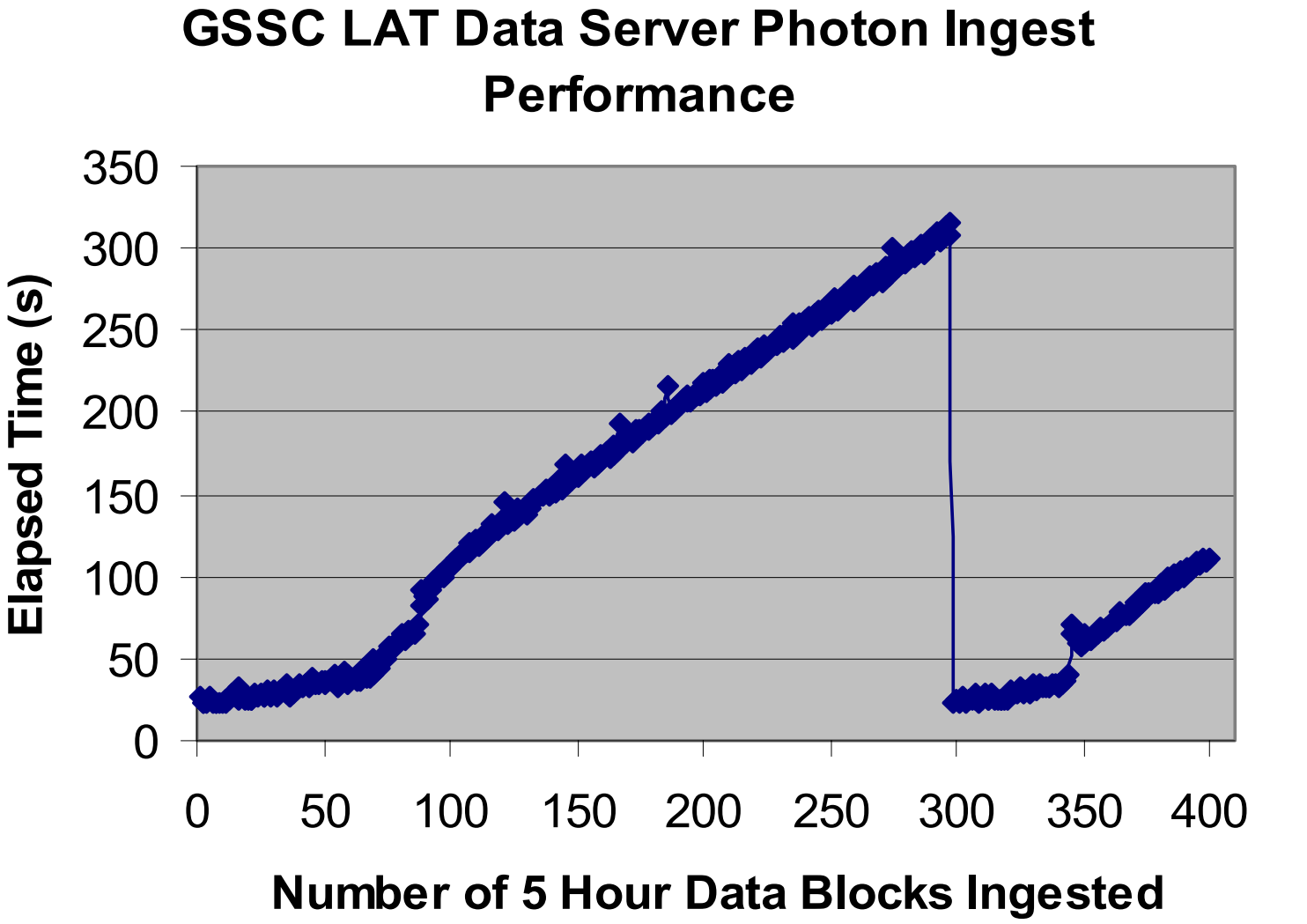
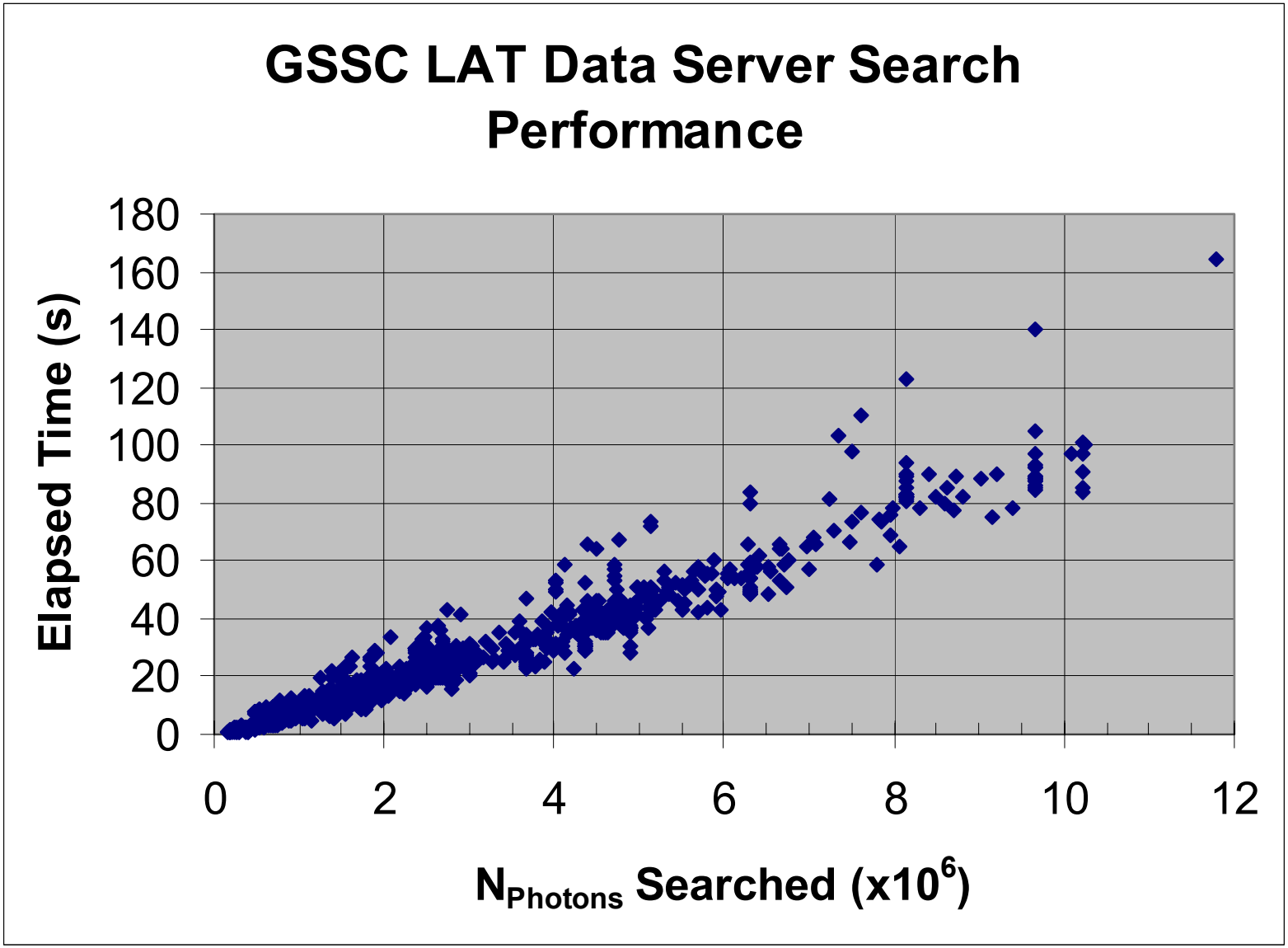
Screenshots – Search Page

Screenshots – Query Submitted



Screenshots – Results Page

Ingest Performance



Search Performance