Event Database Requirements

What the level 1 event databases have to be able to do.

R. Schaefer

Introduction

- **Needed to define requirements for level 1 databases, especially the event databases.**
- **Database committee set out to define those requirements**
  - Seth “TBD filler” Digel, Pat Nolan, Richard duBois, Dave Davis, and Bob Schaefer
- **Created level 1 database requirements document.**
- **Summary presented here**
  - General requirements on searching, HEASARC, etc.
  - Some database numbers
  - Performance requirements and goals
  - Conclusions

Many Databases Needed

But the heavy lifting will be done by:

Event Database Photon Database

Search Requirements

- Searchable on fields that are integers, reals, dates, and times (to 1 microsecond precision)
- Must be searchable by 2-dimensional location on a sphere
- Must be able to include data quality as a search criterion (may be by bits in an integer field)

Miscellaneous Requirements

- API which allows access using a common GLAST LAT programming language.
- A credible upgrade path (does not require special features to perform adequately)
- Must use an SQL close enough to ANSI SQL99 so as to conform to above requirement.
- Must be mirrorable.

HEASARC Compatibility

- **Easy to maintain and operate**
  - ≤ 1 FTE
  - No excessive annual license and maintenance fees.(<$100k/year)

Assumptions About Event Data

- Estimated event rate of ~ 30 Hz (10% photons, 90% particle events)
- Size of event summary ~ 200 bytes/event.
- Assume 5 downlinks from spacecraft to ground per day

Derived Database Fun Facts

Time Period	Photons	File Sizes	Events	File Sizes
1 dwnlnk	0.05 M	0.01 Gb	0.5 M	0.1 Gb
1 day	0.3 M	0.05 Gb	3 M	0.5 Gb
1 year	100 M	20 Gb	1000 M	200 Gb
10 years	1000 M	200 Gb	10000 M	2000 Gb

Standard Searches

- **retrieve all events coming from within a circle of 15 degrees radius (or a 15 by 15 degree box)**
- **15 degree radius circle contains 1.7 % of the sky**
  - ~0.3 Gb photon data/year
  - ~3 Gb event data/year

Photon Database Performance

Action	Photons	Data size	Rate (Req.)	Rate (Goal)
Number of standard searches/ day			60/day	6000/day
Standard Searches	Š 0.15 M/year	Š 0.3 Gb /year	< 30 min/year data	3 sec/year data
Standard Search + filtering on all parameters	Š 0.15 M/year	Š 0.3 Gb /year	< 45 min/year data	5 sec/year data
Large Data sets (> 10% of DB)	> 10 M/year	> 2 Gb/year	< 3 days/ year data	2 hours /year data
# of concurrent searches			2	20
Concurrent slowdown time			Same time as sum of serial searches	1.5 * serial search sum
Ingest/pass	0.05 M	10 M	Š10 min.	Š 3 min.
Ingest data Reprocessed	0.05 M	10 M	<60 min.	10 min.
Rebuild from scratch			<3-days/year data	< 7 hours / year data

Event Database Performance

Action	Events	Data size	Rate (Req.)	Rate (Goal)
Number of standard searches/ day			1/day	100/day
Standard Searches	² 1.5 M /year	² 3 Gb /year	< 10 hours/ year data	1 min /year data
Standard Search + filtering on all parameters	² 1.5 M /year	² 3 Gb /year	< 15 hours/ year data	1.5 min/year data
Large Data sets (> 10% of DB)	> 100 M/year	> 20 Gb/year	< 7 days/ year data	1 day/year data
# of concurrent searches			2	5
Concurrent slowdown time			Same time as sum of serial searches	1.5 * serial search sum
Ingest/pass	0.5 M	100 M	²30 min.	²10 min.
Ingest data Reprocessed	0.5 M	100 M	<10 hours.	1 hour
Rebuild from scratch			<7-days/year data	< 3 days/ year data

Conclusions

- **We have identified the requirements for the database**
- **We have defined performance minimum requirements and goals for the database**
  - The more we can exceed the minimum performance the greater the options for analysis software.
- **Trade studies ongoing to evaluate possible database architectures have begun (DBMS - Oracle, PostgresSQL, Beowulf, etc.)**
- **Come back during the database breakout session to hear about them…**