IF Beam Database

Igor Mandrichenko

9/15/2011

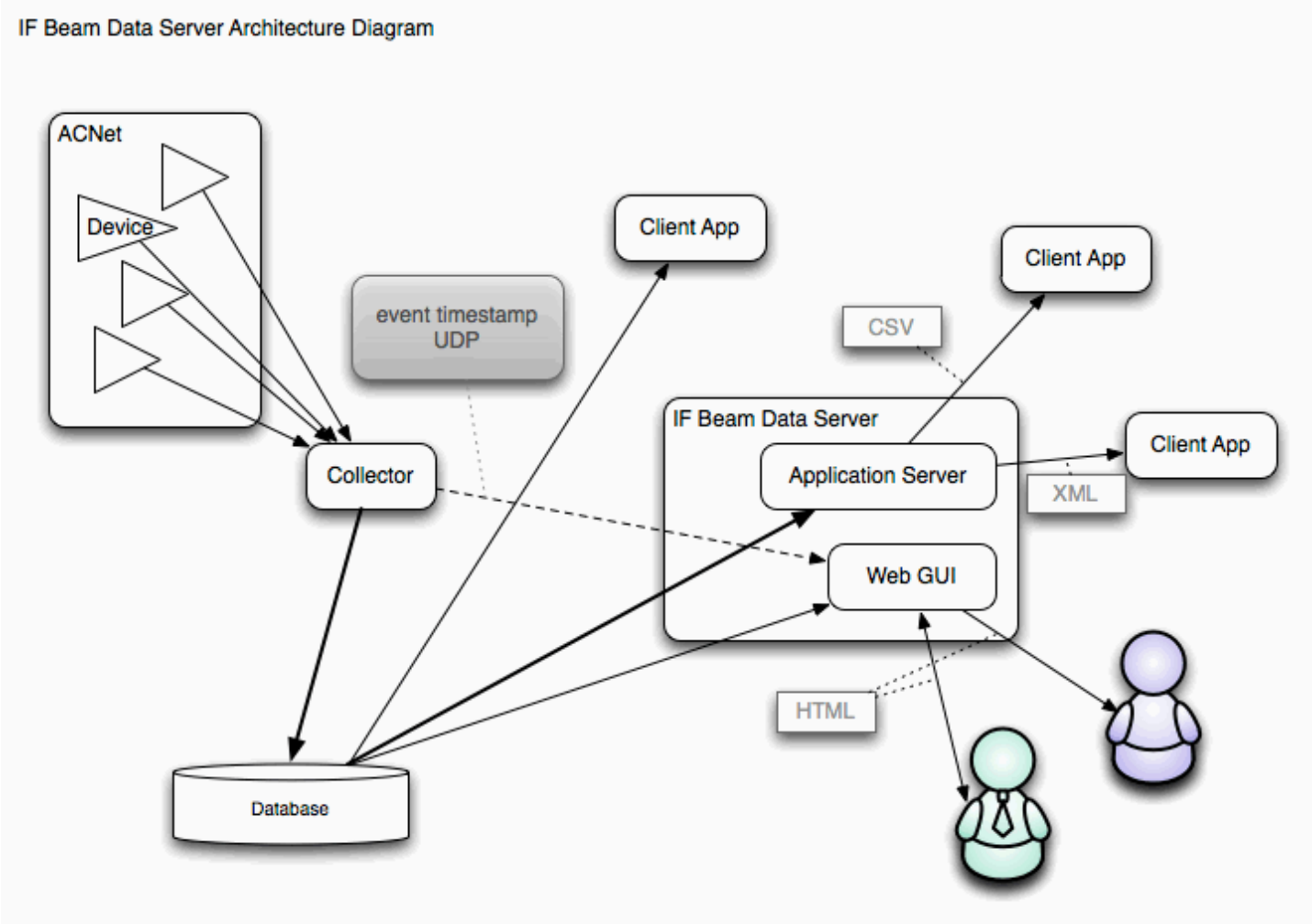
Progress since last meeting

- Added most recent event DB tables, populated in semi-real time, latency ~2 seconds
  - Historic tables have latency ~1 minute
- Added data browser GUI
- Added last event timestamp monitoring bypassing the database
  - UDP message sent directly from collector(s) to the web server process
- Added 6 physics devices charts to the dashboard
- Moved the database from rexdb02 to minervadbprod
  - Disk controller on rexdb02 failed
  - Data since ~ 8/22
- Moved collectors to GPCF and dbweb3
- NuMI_Physics bundle defines set of “important” variables
  - We will start purging the rest, keep them for 1 day initially

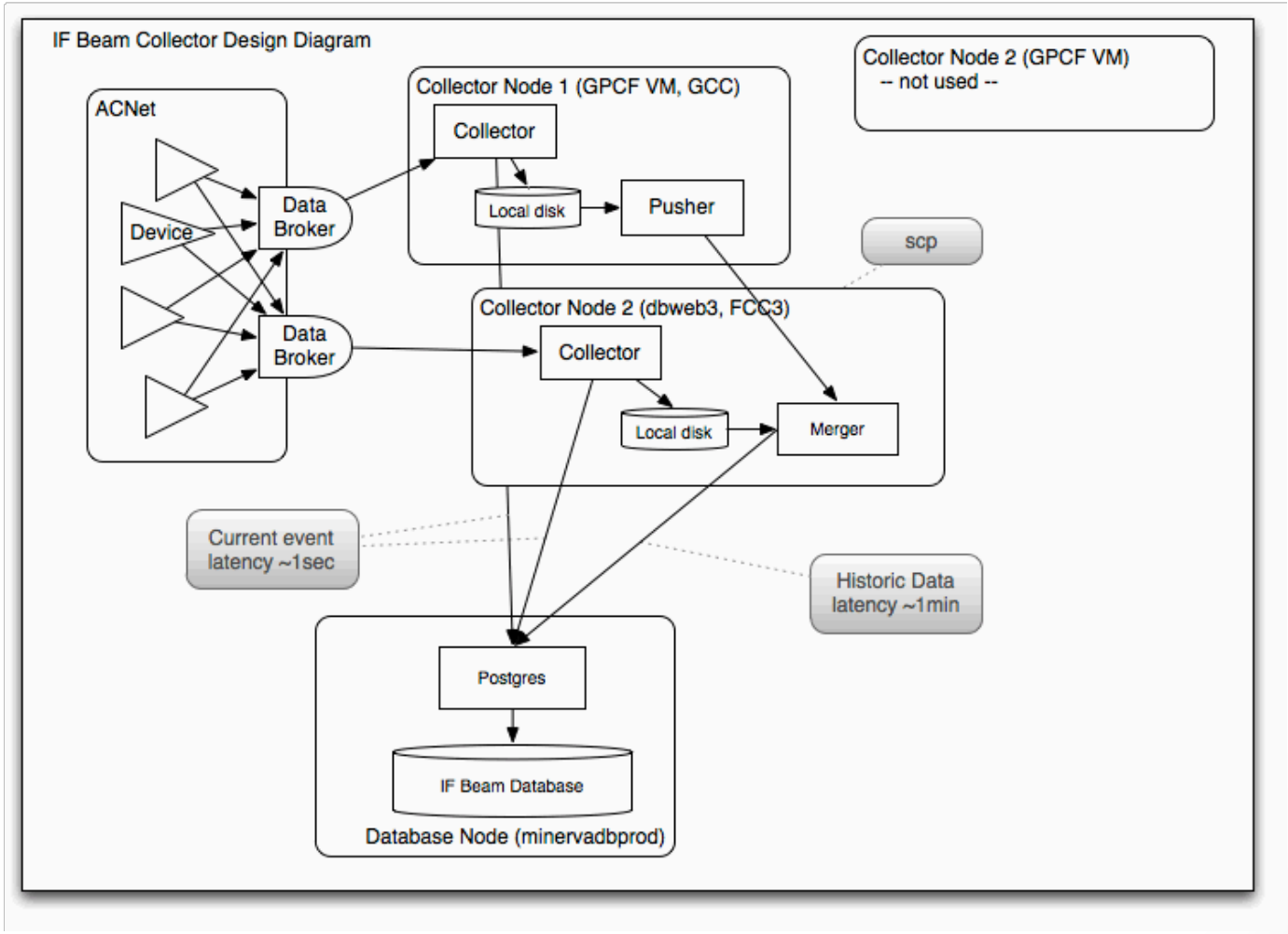
Current Database Size

tablename	rows	pages
public.data	7.10838e+08	26336206
public.events	1.76307e+06	23321
public.bundles	3	1
public.bundle_variables	686	9
public.last_event_data	1178	7860
public.last_event	5	216

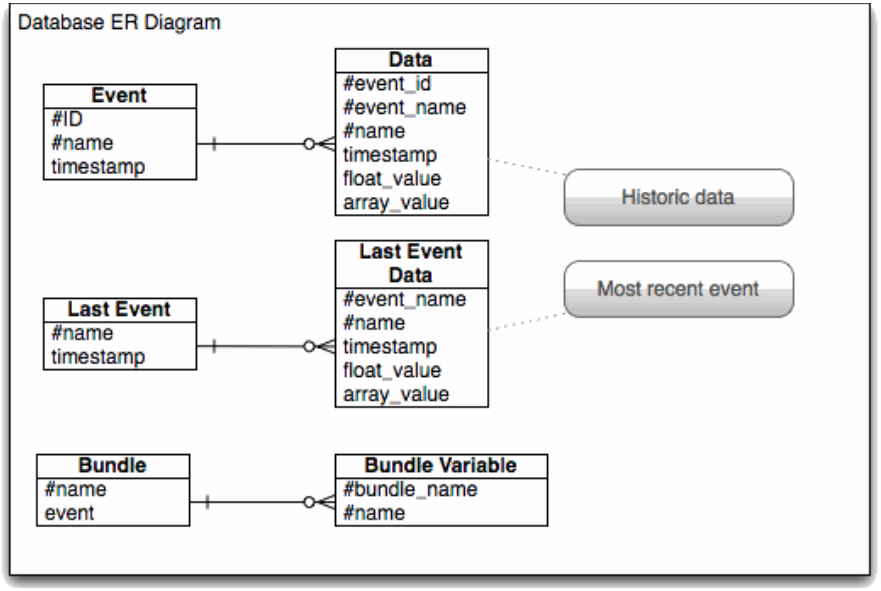
Data Server


IF Beam Data Server Architecture Diagram
ACNet
Device
event timestamp
UDP
Client App
Client App
CSV
IF Beam Data Server
Application Server
Client App
XML
Collector
Web GUI
HTML
Database

Collector


IF Beam Collector Design Diagram
ACNet
Device
Data Broker
Data Broker
Collector Node 1 (GPCF VM, GCC)
Collector
Local disk
Pusher
Collector Node 2 (GPCF VM)
-- not used --
scp
Collector Node 2 (dbweb3, FCC3)
Collector
Local disk
Merger
Current event latency ~1sec
Historic Data latency ~1min
Postgres
IF Beam Database
Database Node (minervadbprod)

Database Schema



Bundle is a combination of
- event name and
- list of devices to read

Database Size Estimate

- 100 Bytes/device/event in average
- Event rate: 1Hz
- 500 devices for monitoring (short term storage, 1 week)
  - 500*100*7*24*3600 = 1.3GB
- 50 devices for long term storage (1 year)
  - 50*100*3*10^7 = 150 GB
- Overhead:
  - Indexes – factor 2
  - Disk backup/standby instance – factor 2
  - Integration instance – factor 1.5
- Total: 1 TB

Database Issues

- Database host: minervadbprod
- Allocated size: 363 GB
  - There is no space for more than 1 backup snapshot
  - Backup takes ~5 hours
- I suspect there is unused space there
- Current event tables are highly dynamic, creates fragmentation, needs to be defragmented
- Historic data will become dynamic too
- Trying to schedule defragmentation next week
- We need more help from DBAs in performance monitoring, tuning, capacity planning
- Restore procedures need to be tested

What we need to know

- Device list
  - More than 100 devices appear to be “dead”
- Trigger events
- Data access patterns
  - By event + time
  - By time
  - All devices or single device
  - Data subsets
  - CSV, XML, what else ?
    - XML schema ?

What needs to be done

- We need input/feedback on:
  - Bundle control
  - Dashboard
  - Monitoring
  - Alerts

Deployment

- Database
  - IF DB host in purchasing. ATA – Fall 2011
  - Using minervadbprod for integration, feedback
  - Read load – minimal, but disk space is an issue
- Collector
  - We have 2 GPCF machines set up for us
  - Collector processes:
    - dbweb3
    - GPCF VM1
  - Merger process:
    - dbweb3
- Web server
  - Dbweb0 (HTTP redirector), dbweb3, dbweb4
  - HTTP redirector now can “gently” probe and know which server can actually connect to the database