

Channel Access Archiver Update

Thomas Birke

Toolkit Structure
Toolkit Components

I/O Library

ArchiveEngine

ArchiveManager

ArchiveExport / CGIExport

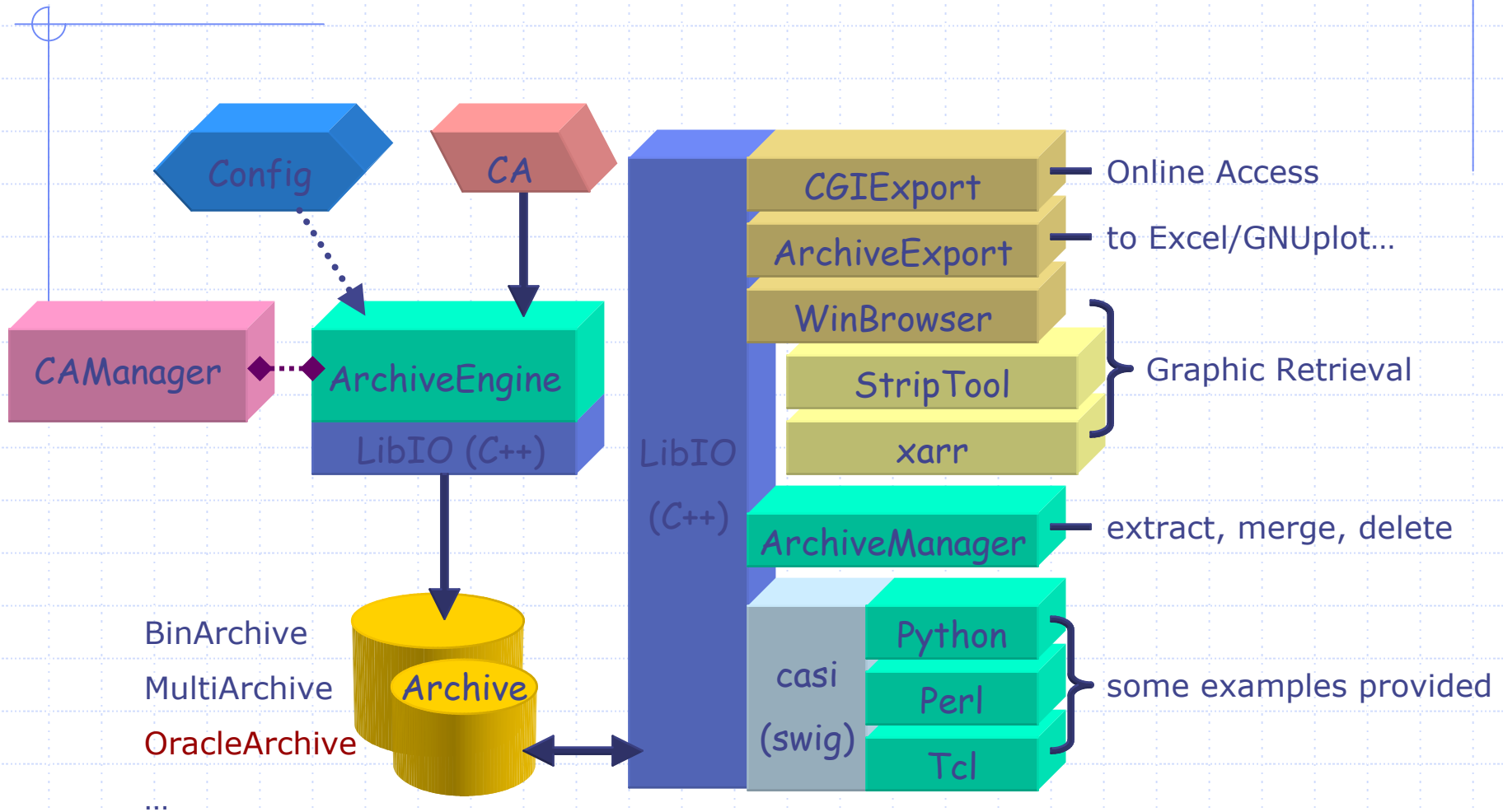
CAManager / CAbgManager

casi

Work in Progress

EPICS Meeting • November 2002 • Jefferson Lab

Toolkit Structure



Toolkit Components

- ◆ I/O Library
read/write from/to archives of any flavor
- ◆ ArchiveEngine
collect data from CA and write these into an archive
- ◆ ArchiveManager
maintain archives
- ◆ ArchiveExport, CGIExport
export data in various formats
- ◆ casi
access the I/O Library from scripting languages
- ◆ CAManager
create/manage archivers and keep them running

I/O Library

- ◆ Read/write from/to any archive
- ◆ Class library to represent an archive
 - Archive, Channel, Value, ControlInfo, Iterators...
- ◆ Basic I/O-functionality
 - No filtering, preprocessing, data-dependant access...
 - Access via channel and point in time
- ◆ Store/retrieve data from external storage
 - BinArchive, OracleArchive, SDDSArchive...
- ◆ Internal layout currently taylorred to BinArchive

ArchiveEngine

- ◆ Collect data via ChannelAccess
 - Frequency based or monitored
 - Writes into archive periodically
 - Archiving of a group may be switched (disabled/enabled) with a PV
- ◆ Simple Web-interface to view status and add more channels

Archive Engine

Engine Info	
Name	ArchiveEngine
Version	1.9.2 (TB), built Mar 18 2002
Description	Machine Control System
Started	05/13/2002 02:00:57.040227000
Archive	2002/20/directory
Channels	2973
Last write check	05/18/2002 13:45:40.553305000
Currently writing	Yes
Default Period	30.000000 sec
Write Period	30.000000 sec
Get Threshold	5.000000 sec
Web Clients (total)	47384
Web Clients (current)	1

[-Main-](#) [-Groups-](#) [-Config-](#)
 (Use *Reload* from the Browser's menu for updates)

Groups

Name	ID	Enabled	Channels	Connected
all	0	Yes	0	0
Beamshutter	1	Yes	24	24
BimScaler	2	Yes	79	79
HF	3	Yes	73	73
ID-corr	4	Yes	280	280
ID-gaps	5	Yes	178	178
ID-Temperature	6	Yes	16	16
LFB	7	Yes	18	18
Landau	8	Yes	36	36

Archive Engine Config.

Groups

- [List groups](#)
- Group Name:
- Parse Group File:

Channels

- [List channels](#)
- Group Name:
- Channel Name:
- Period:
- Monitored:
- Disabling:
- Channel:

[-Main-](#) [-Groups-](#) [-Config-](#)
 (Use *Reload* from the Browser's menu for updates)

ArchiveManager

Low-level management functions

- ◆ Show information (# of channels, min/max timestamp...)
- ◆ Test integrity
- ◆ Dump values for a channel
- ◆ Export data into another archive
- ◆ Read ASCII-file of data into archive
- ◆ Compare archives
- ◆ Rename a channel
- ◆ Delete a channel

ArchiveExport

- ◆ Export data in various target formats
 - GNUplot
 - Excel (CSV)
 - Matlab
- ◆ Output is ASCII
- ◆ Additional exporters to any format can be added

CGIExport

◆ Same Functionality as ArchiveExport

- GNUplot
- Spreadsheet
- Matlab

◆ Gives access to archived data over the web

◆ No data reduction (except for preview)

◆ GNUplot output "optimized"

◆ Limits for y-axis may be specified

◆ Stateless queries

◆ Temporary files have strong impact on performance

Channel Archive CGI Interface

Pattern:
 file glob

Names:

Start: Day (m/d/y) Time (h:m:s)
End: Day (m/d/y) Time (h:m:s)

Plot, y limits All Data: (default: reduced to plot size)
 Spreadsheet Status: (show channel status)
 Matlab-Spreadsheet Fill: (step-func. interpolation)
 Matlab Interpolate: secs (linear)
Log Scale:

Command Explanation:

- **List:** List all channels that match *pattern*
- **Info:** Show info on channels that match *pattern* or are in *names* list
- **Get:** Get values for given channels that are within time range
 - **Plot:** Get simple online plot.
 - **Spreadsheet:** Get a spreadsheet-type text file which you can save to your disk for further investigation.
 - **Matlab-Spreadsheet:** Same, but with slightly different formatting for Matlab. The result requires some typing in Matlab.
 - **Matlab:** Get a Matlab command file. Easier to import into Matlab but usually requires more memory.
 - **Status:** Export the channel status (disconnected, ...). Not supported for all formats.

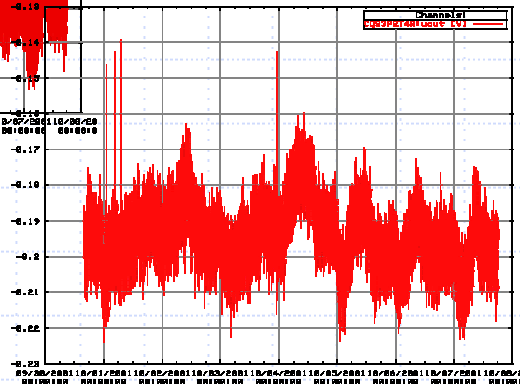
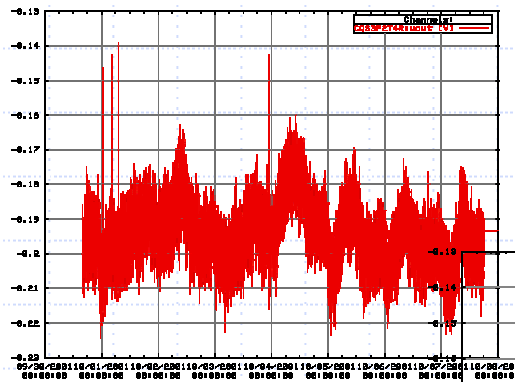
Hints concerning...

- [Regular Expressions for the pattern field](#)
- [Filling, Interpolation](#)
- [Spreadsheet programs](#)



CGIExport

- ◆ *BucketingValueIterator* reduces data on retrieval in a way that at most four values (*first, min, max* and *last*) per “bucket” (a timeslice) are returned



<i>Example</i>	<i>w/o</i>	<i>w/</i>
ArchiveExport	~30s	~5s
resulting file-size	~20MB	~80kB
GNUplot	~11s	~0.5s
CGIExport	~45s	~7s

casi - Channel Archiver Scripting Interface

- ◆ Interface to Perl, Python, Tcl
- ◆ Access to all classes/functions of I-/O library
 - Archive – open, close, getChannel..., addChannel, ...
 - Channel – next, prev, getValue..., addValue, ...
 - Value – next, prev, [gs]etValue, [gs]etCtrlInfo, ...
 - ControlInfo – [gs]et(Status, Severity, ...), ...
- ◆ Read/write data to archives
- ◆ Create archives

casi – a Python example

```
1: # USAGE: xample.py <source archive> <target archive> <pattern> <delta>
2: # copy channels in archive, but reduce values using simple deadband-algorithm
3:
4: import sys
5: from casi import *
6:
7: cmd, source, target, pattern, delta = sys.argv      # get cmdline-args
8:
9: s_arc = archive(); t_arc = archive()
10: s_chan = channel(); t_chan = channel()
11: s_val = value(); t_val = value()
12:
13: s_arc.open(source)                                # read-only access
14: t_arc.write(target, 24)                            # write access, 24 hours per file
15:
16: s_arc.findChannelByPattern(pattern, s_chan)
17: while s_chan.valid():
18:     t_arc.addChannel(s_chan.name(), t_chan)        # create channel in target
19:     t_arc.newValue(DBR_TIME_FLOAT, 1, t_val)       # create value in target
20:     s_chan.getFirstValue(s_val)
21:     firstval = 1
22:     while s_val.valid():
23:         t_val.clone(s_val)                          # copy all info about s_val
24:         if (s_val.isInfo()):                         # info (ARCHIVE_OFF, DISCONNECT...)
25:             t_chan.addValue(t_val)                  # copied
26:             firstval = 1
27:         elif ((s_val.ntype() == DBR_TIME_FLOAT) and (s_val.count() == 1)): # process only floats
28:             if (firstval or (abs(lastval - s_val.get()) >= float(delta))):
29:                 t_chan.addValue(t_val)
30:                 lastval = t_val.get()
31:             firstval = 0
32:         s_val.next()
33:     s_chan.next()
```

CAManager

- ◆ Two processes
CAManager with GUI, CAbgManager without
- ◆ Manage archives and archivers (no retrieval!)
- ◆ Configure ArchiveEngines
host, port, files, directories, schedule...
- ◆ Starts/Stops/Restarts ArchiveEngines
- ◆ Current status via http-interface
- ◆ Runs on Windows and Unix
- ◆ Access to ArchiveManager and ArchiveExport
test/info, export data to spreadsheet, GNUplot, Matlab or
another archive, rename/delete channels from archive

The screenshot shows the Channel Archive Manager GUI. The main window has a menu bar with 'File', 'Edit', 'Tools', and 'Help'. Below the menu bar is a table with columns for 'Host', 'Port', 'Schedule', 'Last Update', and 'Status'. The table contains several rows of data for different archive engines. A 'Host' dialog box is open in the foreground, showing configuration options for a host named 'bogart'. The dialog box has fields for 'Host', 'Description', 'Configuration', 'Archive file', 'Multi-Archive file', and 'Log file'. There are 'Check', 'New', 'Cancel', and 'OK' buttons at the bottom of the dialog box.

Host	Port	Schedule	Last Update	Status
bogart	7410	every 3 minutes	since 11/06/2001 17:06:06	NO
bogart	7420	every 5 minutes	since 11/06/2001 17:08:06	NO
bogart	7421	every 5 minutes	since 11/06/2001 17:05:06	NO
bogart	7422	every 5 minutes	since 11/06/2001 17:06:06	NO
bogart	7423	every 10 minutes	since 11/06/2001 17:00:06	NO
bogart	7411	every hour	since 11/06/2001 17:00:06	NO
bogart	7416	every 3 hours	since 11/06/2001 17:00:07	NO
bogart	7417	every 6 hours	since 11/06/2001 14:26:07	NO
bogart	7414	every day	since 11/06/2001 02:00:07	NO
bogart	7417	every 2 days	since 11/05/2001 13:36:55	NO
bogart	7418	every 3 days	since 11/05/2001 13:36:55	NO
bogart	7416	every week	since 11/05/2001 13:36:55	NO
bogart	7419	no schedule	NO	NO
bogart	7411	no schedule	NO	NO

CAManager

- ◆ Overview of configured archivers
- ◆ Start/stop archivers manually
- ◆ Block archivers from being started/stopped
- ◆ Access ArchiveManager and ArchiveExport
- ◆ Simple configuration file editor

```
home/birke/Archives/BI/PowerSupplies
# Group: PowerSupplies
# This file was auto-created by the ChannelArchiver Engine
# Instance: Machine Control System

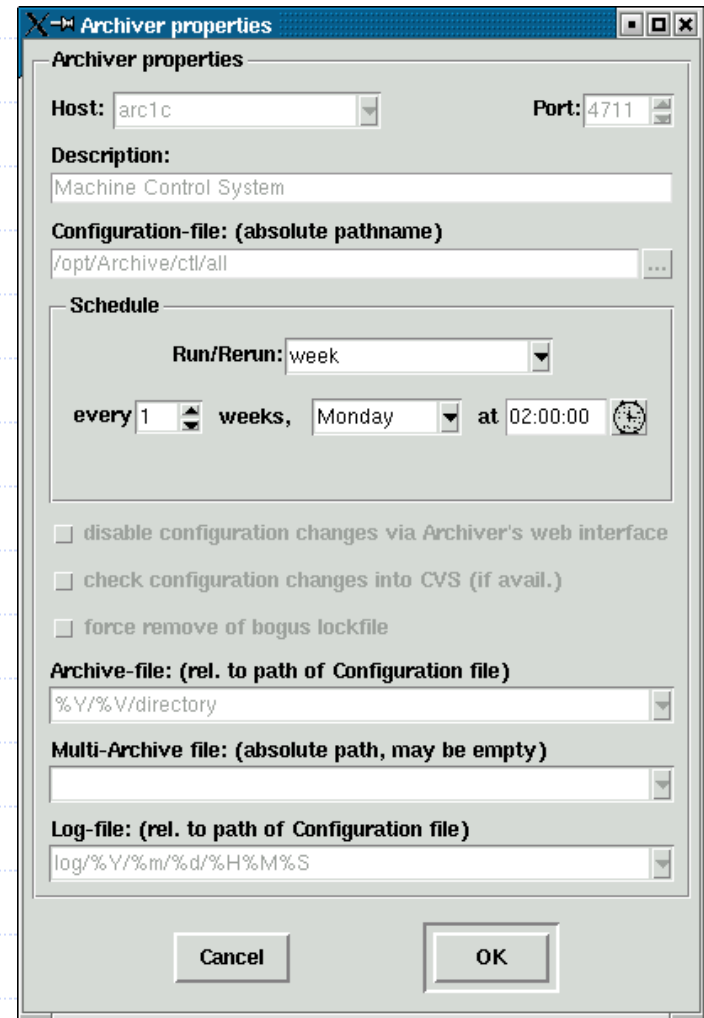
!write_period 30
!default_period 30
!get_threshold 60
!file_size 24
!ignored_future 6
!buffer_reserve 30

!group PowerSupplies-rdbk
!group PowerSupplies-set
!group PowerSupplies-alebk
!group PowerSupplies-ugut
!group PowerSupplies-offset
```

Host	Port	Description	Running?	Block
bogart	7410	every 3 minutes	since 11/07/2001 10:48:07	NO
bogart	7420	every 4 minutes	since 11/07/2001 10:48:07	NO
bogart	7421	every 5 minutes	since 11/07/2001 10:50:07	NO
bogart	7422	every 6 minutes	since 11/07/2001 10:48:08	NO
bogart	7423	every 10 minutes	since 11/07/2001 10:50:07	NO

CAManager

- ◆ Host, port, description, config-file
- ◆ Schedule
 - ◆ hour, day, week, month
- start at x , every y
 - ◆ from – to, always or no schedule at all
- ◆ en-/disable online-config
- ◆ Use cvs for config-files
- ◆ Archive- and log-file
%-substitution
- ◆ MultiArchive



The screenshot shows a window titled "Archiver properties" with the following fields and options:

- Host:** arc1c
- Port:** 4711
- Description:** Machine Control System
- Configuration-file: (absolute pathname):** /opt/Archive/ctl/all
- Schedule:**
 - Run/Rerun: week
 - every 1 weeks, Monday at 02:00:00
- disable configuration changes via Archiver's web interface
- check configuration changes into CVS (if avail.)
- force remove of bogus lockfile
- Archive-file: (rel. to path of Configuration file):** %Y/%V/directory
- Multi-Archive file: (absolute path, may be empty):** (empty)
- Log-file: (rel. to path of Configuration file):** log/%Y/%m/%d/%H%M%S

Buttons: Cancel, OK

CAbgManager

- ◆ Background process (no GUI)
- ◆ Status via web-interface
- ◆ Starts/stops/restarts archivers according to required schedule
- ◆ Prepares directories
- ◆ Checks online config-changes into CVS and uses new config for next start
- ◆ Updates all MultiArchives that are affected

Channel Archiver - bgManager

Configured ArchiveEngines for config
/home/operators/archiver/.CAManager/archivers
of user *archiver* on host *arc1c*

ArchiveEngine	Port	running?	run/rerun	command
Machine Control System	4711	since 05/13/2002 02:00:57	Mondays @ 02:00:00	STOP & BLOCK
Beamline Control System	4712	since 05/14/2002 16:07:49	Mondays @ 02:00:00	STOP & BLOCK
Control System - Surveillance	4721	since 05/13/2002 02:00:57	Mondays @ 02:00:00	STOP & BLOCK
Test System	4731	NO	daily @ 02:00:00	UNBLOCK & START

Messages (most recent first)

[Starts/Stops](#) / [Errors/Warnings](#) / [Misc.](#)

```
2002/05/14 16:07:49: start "Beamline Control System" (/opt/Archive/blc)
2002/05/14 16:07:49: Conflict copying "SMU" - /opt/Archive/blc/SMU is newer!
2002/05/14 16:07:49: Conflict copying "OMSX" - /opt/Archive/blc/OMSX is newer!
2002/05/14 16:07:49: Conflict copying "MICFOC" - /opt/Archive/blc/MICFOC is newer!
2002/05/14 16:07:49: Conflict copying "MCFG" - /opt/Archive/blc/MCFG is newer!
2002/05/14 16:07:49: Conflict copying "CHOP" - /opt/Archive/blc/CHOP is newer!
2002/05/14 16:07:48: Conflict copying "S02D21" - /opt/Archive/blc/S02D21 is newer!
2002/05/14 16:07:48: Saving config-files of "Beamline Control System"
2002/05/14 16:07:47: Lockfile for "Beamline Control System" exists - 2. try in 10 seconds
2002/05/14 16:07:47: manually starting "Beamline Control System"
2002/05/14 16:07:33: stop "Beamline Control System"
2002/05/14 16:07:33: manually stopping "Beamline Control System"
2002/05/13 02:00:57: Conflict copying "SMU" - /opt/Archive/blc/SMU is newer!
2002/05/13 02:00:57: Conflict copying "OMSX" - /opt/Archive/blc/OMSX is newer!
2002/05/13 02:00:57: Conflict copying "MICFOC" - /opt/Archive/blc/MICFOC is newer!
2002/05/13 02:00:57: Conflict copying "MCFG" - /opt/Archive/blc/MCFG is newer!
```

Work in Progress

ORACLE – an alternative to BinArchive

Motivation

- ◆ Archives hold important data not only covering measurements but also long-term surveillance of a machine
- ◆ Keep data in a reasonable resolution online forever
- ◆ BESSY e.g: 2.5 years of data -> ~150GB BinArchive
 - First approach: One big archive
 - ◆ Unmaintainable after a few months!
 - Second (and current) approach: weekly archives combined with MultiArchive-functionality
 - ◆ Better, but...
 - ... retrieval (still) requires some patience ...
And retrieval-time scales with time of archived data!!!

Work in Progress

ORACLE – an alternative to BinArchive

- ◆ DB-layer done by SLAC Bob Hall, Lee-Ann Yasukawa
I/O-library extensions by BESSY Thomas Birke
**Connected! Was planned to run in July 2002.
But is still not complete!
I'm far from where I wanted to be by now...
Retrieval basically works – Archiving doesn't!**
- ◆ **ArchiveManager, -Export and CGIExport work properly with new LibIO**
 - ◆ Tests/dry-runs are promising
 - ◆ Inserting up to ~20k values per second – scales with server size
 - ◆ Partitioning, Direct-Path inserts – indexing “older” partitions
 - ◆ Access time should be fairly constant – benefits from partitioning and indexing
 - ◆ Need “real” machine and storage-system
 - ◆ Should be able to serve multi-TB-online-archives

Work in Progress

Network API – access remote archives

Motivation

- ◆ Perform certain data-reduction on server-side
 - Averaging
 - Any other statistical analysis
 - FFT, ...
 - Whatever one may ask for (modular, pluggable...)
- ◆ Reduce amount of data transferred over the net
- ◆ Two approaches
 - DESY – “AAPI” – Albert Kagarmann et al.
 - Jlab – Corba based? – Chris Larrieu et al.

Conclusion

- ◆ An awful lot of work has already been done
 - Kay Kasemir et al. @ LANL
 - SLAC, DESY, BESSY...
- ◆ Lots of things still to do
- ◆ Trying to create a modular system where Labs may pick/improve/develop features they need and have a core system that works for everyone
- ◆ Let's work out some directions...