

# **Data Archiving in KEK**

**EPICS collaboration meeting**

**Nov. 2000, Knoxville, Tennessee**

**Noboru Yamamoto**

**KEK, JAPAN**

# Data Archiving Tools in KEK

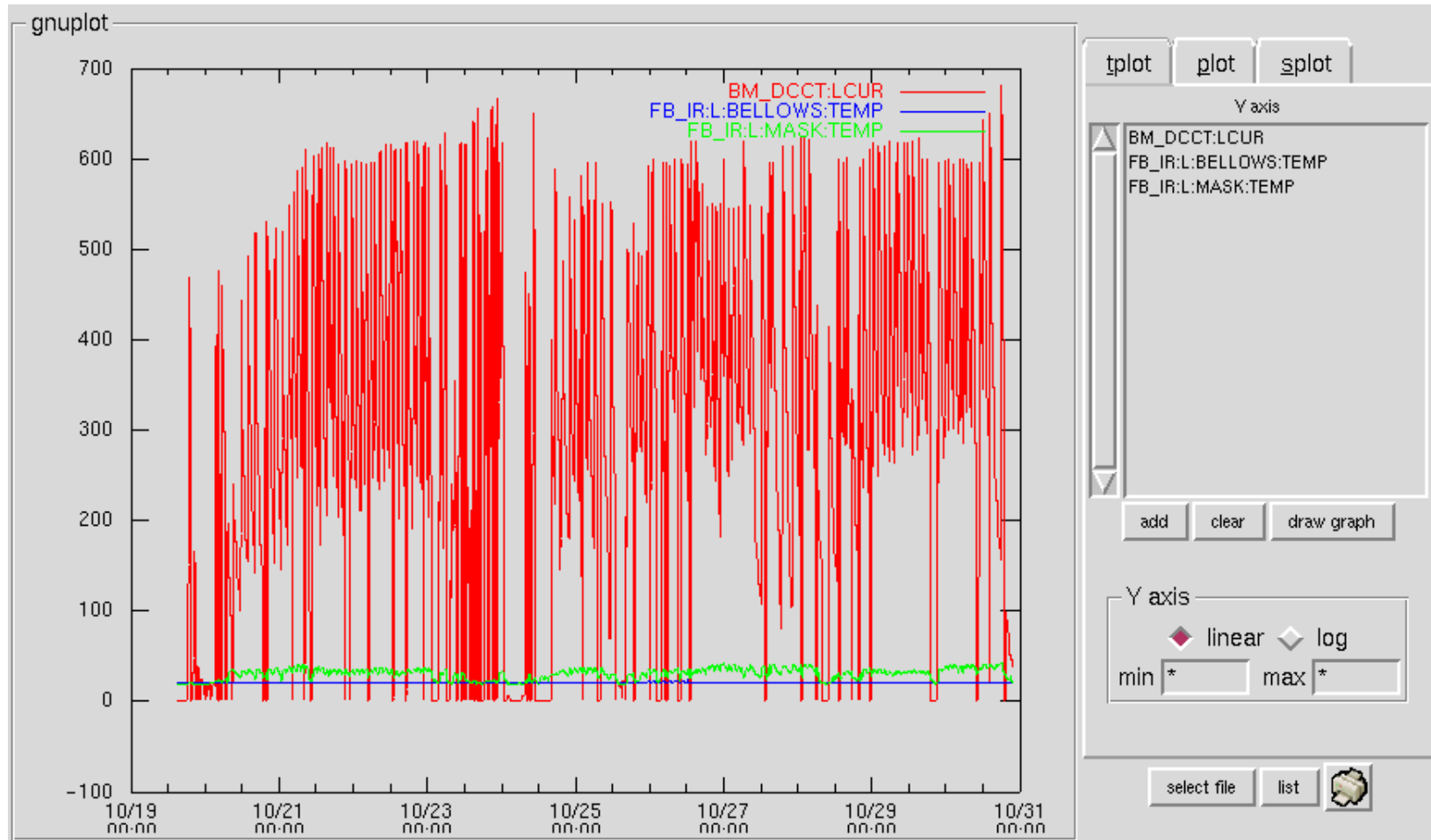
- ▶ **OLD EPICS Archiver(AR\_cmd)**
- ▶ **Akasaka's Archiving Tool (also known as KEKB logger)**
- ▶ **Channel Archiver (under evaluation on HP-UX)**

# OLD EPICS archiver

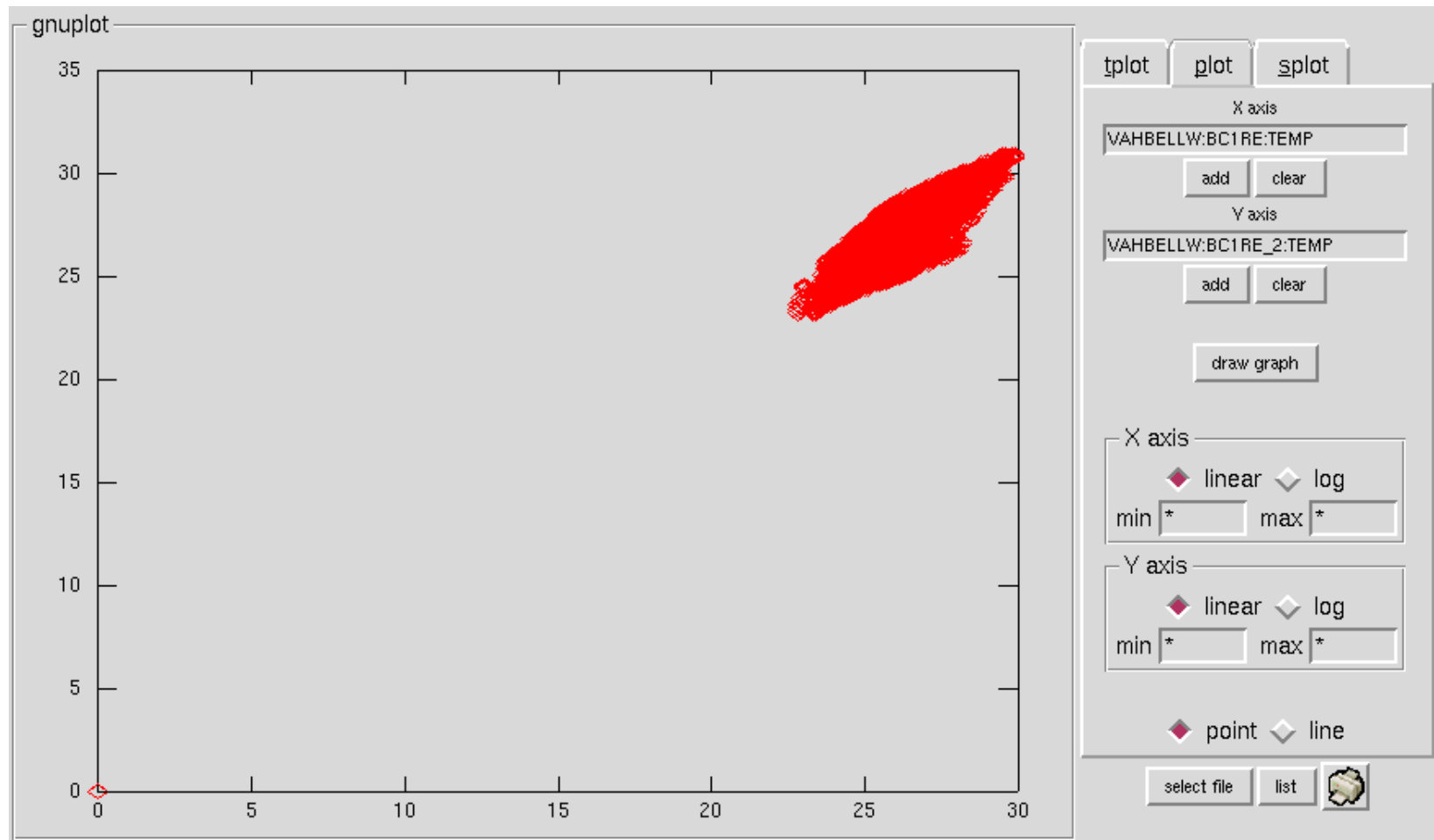
- ▶ Used for some several application.
- ▶ AR\_cmd in both standalone/server modes.
- ▶ ar2cwn command converts archived data into ASCII and cwn\* binary format.
- ▶ Python program , arr\_plot , is developed to draw graph from archived data.

\*) cwn format is a binary data format for PAW(Physics Analysis Workbench) program developed at CERN.

# arr\_plot.py graph example



# 2D plot in arr\_plot.py



# KEKB Archiving Tool

N. Akasaka

1. Target
2. File format
3. Retrieve command

# 1. Target

## 1. Record as much data as possible

- +Record all the monitored value sequentially in binary format
- Long retrieve time

## 2. Simplicity and ease of use.

- +Only a file which contains the names of the records are necessary
- Lack of sophisticated triggering etc. as in ar.

## 3. Save disk space

- +Compressed off line using zlib in small chunks (~100kB).
- Time stamp resolution is 1/100 sec.

## 2. File format

- ▶ When the log starts by a command “kblog Example“, two files with names “Example\_20000517110000.log1” and “Example\_20000517110000.log2” are created. The record names for archiving is taken from the file “Example.lcf“. Typically kblog continues to run for 3~7 days.
- ▶ “.log1” file is the index file, which contains an index number, type of value and number of elements for each record. In the case of "chunked" files, the offset values to the start of chunks follow.
- ▶ “.log2” file is the data file. For each of archive events, the offset to the next record (2 bytes), index number (2 bytes), time stamp (4bytes), and value(s) are written. Changes to the connection state are also recorded. In the case of "chunked" files, latest values of all the records are written at the start of each chunk.



# 3. Retrieve command

- ▶ The archived data is read by kblogrd command:
  - ▶ kblogrd [-r <record name>[,<record name>].....
  - ▶ [-t <start time>[-[<end time>][d<time interval>]]]
  - ▶ [-f <output format>] <log file name>
- ▶ The record name can be a regular expression.
- ▶ kblogrd command convert binary data into ASCII format. The output format is one of “free”, “sad” and “kaleida”.
- ▶ The time range which is not contained in one data file is handled automatically by kblog.

## 4. Operation

- ▶ 20 kblog processes stores 2 GB/day data onto local disk.
- ▶ These processes use 15.5 % of CPU time on one of Alpha server.
- ▶ Old archived data are compressed (50% reduction in size) and send to a mass storage system in the BELLE data analysis computer system.

# kblog processes for KEKB operation

```
acsad3.kek.jp.20: ps -ae |grep kblog
  6338 ttyp6      S N+    10:30:48 ./kblog AnalogData
 22472 ttyp9        S N+    44:59.36 ./kblog BinaryData
  6039 ttypa        S N+    01:10:13 ./kblog BTMagnets
  5772 ttypb        S N+    01:25:38 ./kblog CCG
  5915 ttypc        S N+    04:10:12 ./kblog LERMagnets
  5920 ttypd        S N+    03:22:26 ./kblog HERMagnets
13211 ttype        S N+    51:14.47 ./kblog BPM
16714 ttypf        S N+    23:55.66 ./kblog -f Base
16724 ttyq0      S N+    24:57.31 ./kblog -f -m Physics
19154 ttyq1      S N+    32:53.05 ./kblog -f PEPII
  6116 ttyq2      S N+    53:20.14 ./kblog Env
  6214 ttyq4      S N+    28:49.75 ./kblog DCCT
  5851 ttyq5      S N+    16:16.69 ./kblog BCM
25647 ttyq6      S N+    01:21:28 ./kblog Misc
19067 ttyq7      S N+    35:06.98 ./kblog Tune
19069 ttyq8      S N+    45:34.21 ./kblog Abort
25421 ttyq9      S N+    01:01:04 ./kblog BTBPM
  5839 ttyqa      S N+    43:38.02 ./kblog VATemp
  6432 ttyqb      S N+    22:07.95 ./kblog OctoPos
```

CPU usage in total = 15.5 % on Alpha server