Data_Inspection Long-Range Plotter

July 21, 2000

Introduction

The Data_Inspection Long-Range Plotter was developed to allow one to view trends in event/trigger rates and mean nhit values over many subruns at a time. As well, the ability to view several trend plots at once allows one to look for correlations among them.

The tool was developed in two parts. The first part, `get_stats_*.kumac`, scans data_inspection `HBK` files on `SURF`, gathering mean values from the relevant histograms and storing the results in an `ASCII` data file. The second part, `j_plot.kumac`, reads in the data file and allows the user to interactively generate plots from the data.

`get_stats_*.kumac`

The Code

The code for `get_stats_*.kumac` resides on `SURF` at

```
/users/sas/production/app_sri/scripts/get_stats_batch.kumac
/users/sas/production/app_sri/scripts/get_stats_interactive.kumac
/users/sas/production/app_sri/scripts/j_copy.f
```

All three of these files are under CVS version control. Please contact Steve Brice before making any changes to these files. `get_stats_batch.kumac` is run by `SAS` in batch mode (cmd: `paw -b get_stats_batch.kumac`) every time a new data_inspection `HBK` file arrives on `SURF`. The output data is written to

```
/bigdisk/sas/data_inspection/j_plot/j_plot.dat
```

and date/time-stamped log files are written to

```
/bigdisk/sas/data_inspection/j_plot/j_plot/log/
```

for diagnostic purposes.

`get_stats_interactive.kumac` functions exactly as `get_stats_batch.kumac` with regard to the gathering of relevant statistics from the data_inspection `HBK` files. However, it is designed to be used by an expert user to fix or reconstruct the data file in the event that it is corrupted. The main additional feature in this kumac is the ability to specify, through command-line arguments, the subrun at which to start/stop. As well, data output can be directed to a file other than the default, `j_plot.dat`. Please see the description of the available command-line arguments which is included in the first few lines of `get_stats_interactive.kumac`. `j_copy.f` is an included `FORTRAN` file which is necessary to the operation of either kumac.

Summary of Operation

`get_stats_*.kumac` first reads in the data file to determine which subrun was last analyzed. If the data file is empty or does not exist, then analysis will attempt to begin at run 10000 subrun 0. Otherwise, the kumac analyzes the data_inspection `HBK` directory (defined in the first few lines of the kumac) to determine the availability of `HBK` files for subruns which follow the last subrun in the data file.

If, for a given run, the kumac determines that an `HBK` file should exist but does not, or if the All Triggers Event Rate histogram (ID# 1018) is empty, then zeros are padded into the data file for all values for that subrun. It should be noted that the method by which the kumac determines whether or not an `HBK` file exists for a given subrun is not completely foolproof — for example, the method will fail if the `HBK` file for the last subrun of a run is missing. As well, the kumac has no ability to handle *gzip*'ed files, and as such each data_inspection file should not be *gzip*'ed until its relevant data has been captured by `get_stats_*.kumac`.

Otherwise, the kumac happily proceeds to read in histograms from `HBK` subrun files to collect the various average values in which it is interested. `get_stats_batch.kumac` begins at the next detected subrun and analyzes all subsequent `HBK` files until no more can be found. `get_stats_interactive.kumac` will do the same, unless starting and/or ending run numbers have been given as command-line arguments.

There are a few things to note concerning the method used to calculate average rates related to the way in which the original histograms are binned. The rate histograms in the data_inspection `HBK` files are binned by two hundred second intervals, yet subruns rarely start or end exactly on one of these divisions. Thus, in an effort to increase accuracy, the last bin of the histogram is not used in the average rate calculation. In the case of a run which spanned multiple subruns, nor is the first bin of the histogram used for the non-zeroth subruns. This method is effective for subruns which last longer than four hundred seconds and thus have at least three bins. However, the rate calculated for subruns which last less than four hundred seconds, while included in the data file, should not be trusted, and it is recommended that when viewing rate plots a cut be made on subrun length at four hundred and one seconds (see the section in this document on `j_plot.kumac`).

Finally, there are no sorting capabilities at the read-back stage, `j_plot.kumac`, and thus it is very important that `get_stats_*.kumac` analyze the data_inspection `HBK` files in the proper order. In interest of this, multiple instances of `get_stats_*.kumac` should never be run if using the same output data file, especially if that file is the default `j_plot.dat` file. In general, neither of the two `get_stats_*.kumac`'s should be run by the user unless s/he is certain that s/he knows what s/he is doing! Access restrictions on both the data and kumac files should prevent the average user from damaging the default data file, but if a user knows the `SAS` login password then all bets are off.

j_plot.kumac

The Code

j_plot.kumac allows the user to interactively generate plots from the data compiled by get_stats_*.kumac. Operation of j_plot is menu-driven and should be self-explanatory, but a brief overview of some of its features will be given here.

The code for j_plot resides on SURF at:

```
/bigdisk/sas/data_inspection/j_plot/j_plot.kumac
/bigdisk/sas/data_inspection/j_plot/jplot.script1
/bigdisk/sas/data_inspection/j_plot/jplot.script2
/users/sas/production/app_sri/scripts/j_copy.f
```

The two jplot.script* files are used within j_plot.kumac to merge postscript files — this feature is described in more detail in the "Summary of Operation" section below. Again, the file j_copy.f is an included FORTRAN file required by the kumac.

When executed, j_plot.kumac reads in the data produced by get_stats_*.kumac from

/bigdisk/sas/data_inspection/j_plot/j_plot.dat

The directory

/bigdisk/sas/data_inspection/j_plot/j_plot/

contains two backups of j_plot.dat, staggered by two weeks, which are maintained by SAS. This directory also holds the log/ directory, where the get_stats_batch.kumac log files are written.

In consideration of the end-user, soft-links to the files necessary to run j_plot exist in the main data_inspection HBK directory:

/bigdisk/sas/data_inspection/hbk/

Summary of Operation

Event/trigger/quantity types for which plots are available are listed in Table 1.

Event/trigger rates, mean NHIT, or raw values (as appropriate) for the quantities listed in Table 1 may be plotted against run number, raw subrun length, or cumulative time, though not all choices of x-axis value are available for each listed quantity. In addition, several mean Analog Sum plots are available for "All Triggers".

To start j_plot, do:

```
SURF> cd /bigdisk/sas/data_inspection/hbk
SURF> paw
PAW > exec j_plot
```

Subrun Length	Pulse_GT	Ring Of Fire
All Triggers	Prescale	Orphans
NHIT100_Hi	Sharkfins	# Crates Online
NHIT100_Med	Flashers	# Cards Online
NHIT100_Lo	Wet-End Breakdown	# Relays Set
NHIT20	Neck Events	# 100ns Triggers Enabled
ESUM_Hi	HIB (EXT2)	# Sequencers On
ESUM_Lo	Muons	# Chan w/ HV On
OWLN	ITC	
OWLE_Hi	QvNHIT	

Table 1: Available Plots

`j_plot.kumac` will accept one command-line argument – a fully qualified path to an alternate data file. If no such argument is given (and under most circumstances, no such argument *should* be given) then `j_plot.dat` is used by default.

```
Choose an option:
0 - Quit
1 - Start over
2 - Go back to previous menu
3 - Linear Y-axis
4 - Log Y-axis
5 - Linear X-axis
6 - Log X-axis
7 - Add Error bars
8 - Remove Error bars
9 - Perform cuts
10 - Remove cuts
11 - Reset cut parameters to default values
12 - Zone another plot
13 - Remove a plot
14 - Capture current plot window to file in your home dir
15 - Change marker
16 - Manually set Y-scale
17 - Fit function to plot
18 - Remove fit from plot
```

Figure 1: Plot Options Menu

After loading the data file, the user will be confronted with a text menu where s/he may choose from the quantities listed in Table 1. Next, the kumac will ask to plot either Mean Rate, Mean NHIT, or Analog Sum. However, as mentioned, the Analog Sum plots are only available for "All Triggers". Finally, the user must decide which x-axis to plot against: run number, raw subrun length, or cumulative time. After this decision, the requested plot is displayed with default options, and a post-plot menu appears allowing the user to change the appearance of the plot (See Figure 1).

The default options are log scale on the y-axis, linear scale on the x-axis, no error bars, no cuts (unless the x-axis is cumulative time, in which case all zero length subruns are cut), horizontal-line marker, auto-scale on the y-axis, and no fit.

Notes on Plot Options:

- "`Log X-Axis`" option is only available for plots versus Raw Subrun Length.
- Error Bars are not available for Subrun Length plots, nor the Detector Configuration plots (# Crates Online, etc.).
- When using Option #9 ("`Perform Cuts`"), `j_plot` will ask for minimum and maximum run numbers as well as minimum and maximum subrun lengths, even if the Subrun Lengths plot is not currently zoned. Next, `j_plot` will, for every zoned plot in turn, ask if the user wishes to make changes to the minimum and maximum cut values for the raw or RMS values for the quantity of that plot. Through changing these cut values, the user is able to cut outlying points or points for which the error is too large, for example.
- When a cut is specified, it is applied to all plots currently zoned. However, this does not mean that if one specifies a maximum raw value on one plot, that same maximum value will be applied to every other zoned plot. Rather, if (for example) a particular subrun fails a maximum raw value cut on one plot, that subrun will not appear in the post-cut regeneration of any of the zoned plots, regardless of the cut parameters defined for the other plots.
- Up to four plots may be zoned at a time.
- When capturing a plot window to a postscript file, it is possible to append the postscript data to an existing postscript file in the user's home directory – when asked for a filename, the user should just enter the name of the existing postscript file, and `j_plot` will ask if it should append or overwrite. This function is only guaranteed to work for existing files that were created with `j_plot` and have the same orientation (portrait or landscape) as the desired output — using this function in other situations will yield less than desirable results.
- When fitting, it is possible to remove outliers by first cutting those subruns for which the quantity in question falls above/below a specific value.

Final Notes

- As of the writing of this document, `SAS` has been running `get_stats_batch.kumac` successfully for several weeks. As well, `j_plot.kumac` has been successfully in use for months and should not contain any major bugs. Should any minor glitches appear in the data file (for whatever reason) following my departure in September, Noel Gagnon will be notified by `SAS` and will be capable of handling the situation.

- Because of a bug introduced into `SNOMAN` during the regeneration of the data_inspection `HBK` files this past June, the data for the Detector Configuration plots (# crates online, etc.) were missing from a small number of the `HBK` files. Thus, for those subruns, `j_plot` will show a non-zero run length yet zero crates online, zero cards online, etc., which is clearly nonsense. These subruns may be removed from those plots by cutting on the raw value of the plot, specifying a minimum of 1.