FOR ELCAP INTERNAL USE

FROM LOGGER TO ARCHIVE

Mark Halverson
Scott Crowder
Sharon Bailey

May 1987

Prepared for
the Bonneville Power Administration
and the U.S. Department of Energy
under Contract DE-AC06-76RLO

Pacific Northwest Laboratory
Richland, Washington 99352

FROM LOGGER TO ARCHIVE

This document contains a brief explanation of the collection,processing, and verification of the ELCAP load data. In any large scale data collection project, there is an enormous amount of data processing and manipulation of data. ELCAP is no exception to this. Every piece of data that comes into the lab has been passed through 4 different computer systems starting with the data logger and ending with the data archive on the AVAX. Every piece of data has been processed through at least 6 programs or command files. Some data has gone through even more by the time it reaches the data archive. As a data analyst you should be aware of everything that has been done to the data before it is made available to you. The assumptions made in these programs can affect not only the appearance of the data but also the availability and quality of the data. Some of the assumptions may actually bias certain types of analyis, while other assumptions may be totally unimportant to your analysis.

1.0 THE DATA LOGGGER

Electrical power is the product of the voltage and the current. The ELCAP logger explicitly measures both of these quantities and multiplies them together for each monitored channel. The value reported represents the true power consumed by the device or devices being monitored. The current is measured by a device called a current transformer (CT), which produces a signal proportional to the current flowing through the wires being monitored.

The voltage is measured with a sample voltage transformer, which produces a signal proportional to the line voltage. The current is measured separately for each sensor channel. The voltage is measured only once for each phase of the electrical service for the entire logger.

Since the logger can report only 256 different values, or counts, and we are interested in circuits from a single light bulb to the main power feed for a large commercial building, it is important that we be able to configure the instrumentation in such a way that each count will represent a different number of watts, depending on the types of circuits that are being monitored.

There are three different ways to change the sensitivity of the individual channels in the ELCAP logger. The first is by choosing a CT to match the rated capacity of the circuits being monitored. The second is to change the scaling resistor used to condition the signal produced by the CT. The third is to change the transformer used to sample the voltage of the building electrical service.

An example may be helpful in understanding this concept. Suppose there is a hot water heater attached to a two-phase, 50-amp circuit breaker. The ELCAP instrumentation would be installed using a 100-amp CT for each phase. The CTs would be attached to two channels, and both channels would be scaled so that they read full scale for 50 amps. The channel resolution for the above configuration would be 27 watts per count for hourly data. The resolution could be made finer by scaling the channels for a full-scale reading at 30 or 40 amps instead of 50 amps, because most circuit breakers are oversized. However, ELCAP practice has been to scale to the maximum circuit breaker rating, even though the resulting channel resolutions are larger than may be absolutely necessary.

One other important aspect of the logger that affects the quality of the engineering data is the offset that is applied to each channel within the logger. The function of the offset is to ensure that a positive voltage exists in the logger channel at all times, even when the signal from the current transformer is zero. The offset is individually calibrated for each channel before the logger is sent to the field. We know that the offset can drift from laboratory calibration, and we will discuss the effect of this below.

Once the sensors have been installed, the logger will record a reading once every second and convert this reading into the proper number of counts. These 1-second readings are then added together for the duration of the integration period; for instance, 3600 readings are added together for an hourly reading. At the end of the integration period, the binary sum is truncated so that it is again less than eight digits long. This is the same as dividing by 2 to however many digits have been truncated. In the case of hourly records, 12 digits are truncated or the total of the 3600 readings is divided by 4096.

The meteorological data is also recorded as a digital signal within the logger and is accumulated and truncated to form the data record in the same way as the energy data. The item of most concern to the analysis is how the temperature data has been treated.

It is this truncated number that is stored in the logger memory and subsequently transmitted to the laboratory. We then convert this digital count into engineering units by multiplying by a calibration factor that is the number of watts per count. It is very important to any analysis that the transmitted count is correctly converted into an engineering unit. The resolution of each channel depends on the size of the CT, scaling resistor, and sample voltage transformer, and on the ratio of the truncation to the number of seconds in the integration period. Figure 1 is a schematic of a typical sensor channel.

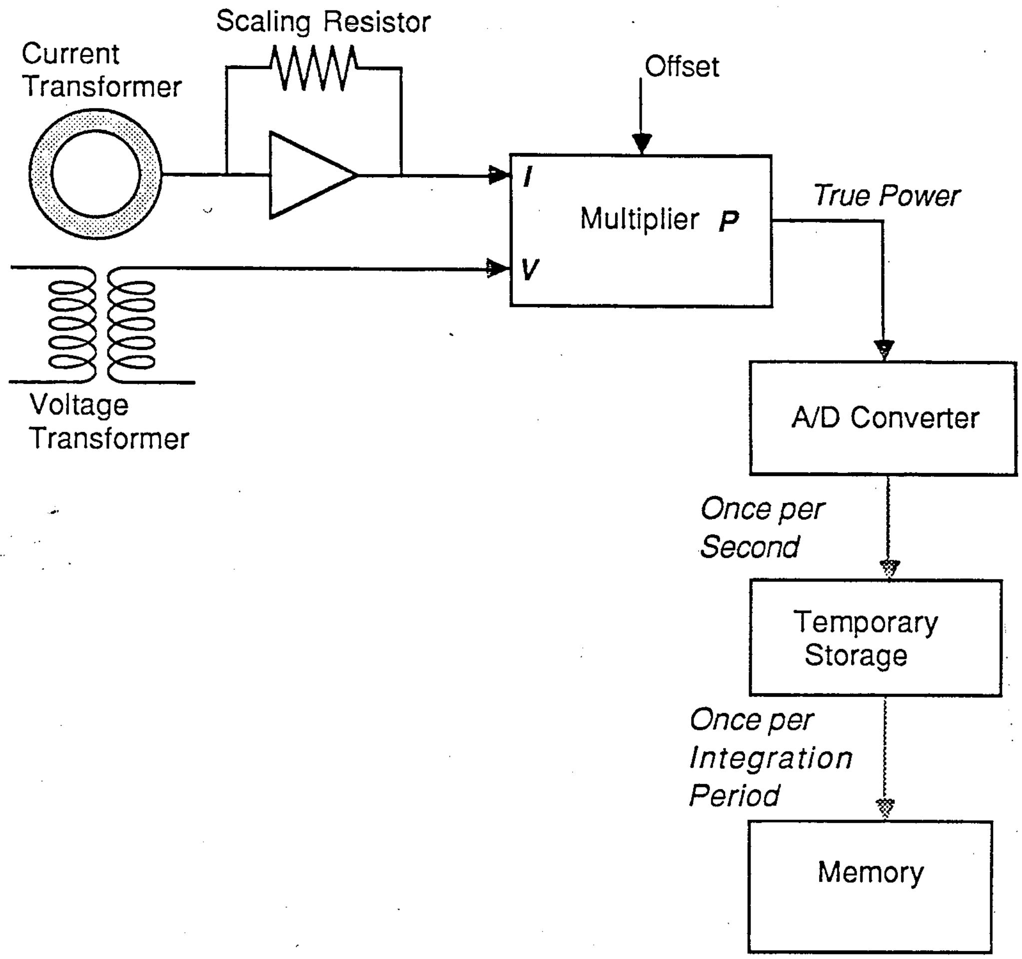

FIGURE 1. Schematic of a Typical Sensor Channel

2.0 DATA PROCESSING

Once the data has been collected and processed on the data logger, the process of bringing that data back to the lab can begin. The first data processing program that the record encounters at PNL is known as "DI". This is the data interrogation program that collects the data from the data loggers in the field and stores it on the Hewlett-Packard (HP) system. DI collects thousands of records each day from many loggers and creates a large file of all the binary records collected on a particular day. Each record contains the DI number, the time (both what the logger thinks the time is and what the HP system thinks the time is), several control parameters, and the raw data.

Every day, the output from the previous day's DI run is fed to another program called CONVERSION. This program also resides on the HP system and its' sole function is to convert the binary data records to records containing real engineering units. To do this, data about the scaling resistor and current transformer sizes used at the installation must have previously been entered. There is always a potential for data entry errors in this step, but these are almost always caught in a series of checks later on in the process.

After CONVERSION has run, the large data file now containing real engineering records is transferred to the ELCAP3 Microvax for further processing. The first step is to take the large file of all the data collected on a particular day and split it up into data for individual sites. This program is called the REFORMATTER and it is one of most important programs in data processing. Among other things, the REFORMATTER separates out data for individual sites into separate files, checks for bad reference voltages in the logger, checks for incomplete data records in the data, checks for any periods of power outages in the data, checks for clock drift in the data, reconstructs the timestamps on power outage data where it can, and reconstructs clock drift data when it identifies it. As you might guess, the REFORMATTER has a large number of assumptions about what is "good data" built into it. The REFORMATTER has been improved several times over the years so there is data that has passed through old versions of the REFORMATTER in the archive with data from the new version. All data should appear the same to the analyst who looks at the data now. If you are interested in the assumptions used by the REFORMATTER to check and correct the data, please contact the data processing manager.

At this point in the process, we have individual data files of real engineering records on the ELCAP3 Microvax. The files are of random length depending on how long it had been before we collected data from a particular site. The next step is to get the data to the data archive on the AVAX in a form that is accessible to the analyst. The current storage form on the AVAX is in week long files of channel level data for individual sites. A series of 3-step processes known as APPEND is used to move the correct data to the AVAX and append it into week long files. It is at this stage that each data record is checked for reversed current transformers and is check summed to comfirm that no changes or failures have occured at the site. A data quality flag is added to each data record which allows the analyst to pick and chose the actual quality of data desired for the analysis at hand. Data is totally unavailable to analysts until these steps are complete.

Before analysis can begin, a program designed to update EASE files must also be run. This is typically run in conjunction with the LINEGRAPH reports (see the section on "Getting Started" for more information on this.)

All this sounds like an incredible amount of data handling. It is. Keep in mind that the description above is only for data we believe is of relatively good quality to begin with. If we have any suspicions that the data might be bad, a whole additional series of steps are performed to check and correct the data. This is discussed below.

3.0 DATA VERIFICATION

We apply a process known as verification anytime anything is fixed or changed at a site. This includes the original installation and any other site modifications made at a later date. The object of verification is to ensure that the installation or modifications were made correctly and that no problems were missed or created. This is really one of the chief distinquishing features that separates ELCAP data from most other metering projects.

We have the ability to tell if our data is bad long before analysis begins and while there is still time to fix it. The whole verification process is based on redundant monitoring. We monitor both the mains coming into a site and the feeders out of the breaker panel or switch gear. The energy entering the site should equal the energy leaving the breakers at any one time. This is a very straightforward application of the First Law of Thermodynamics. Of course, given the processing that goes on in the data logger, and the limitations of the monitoring equipment, the energy entering and leaving the logger never quite add up to zero. We have built up quite a bit of experience as to how close to zero this difference should be and use this as the basis of our data quality checks. Small changes to the offset values noted above can have a great effect on the sum-checking procedure described here. An additional check of the load shapes of the various enduses is made to insure that the channel marked "refrigerator" is really a refrigerator (or at least looks like one) and that the "furnace" is really a "furnace" and not an "hvac" system with cooling capabilities.

A typical site will be set up to collect short integration period data after any major change. This period is 5 minutes for a residential site and 15 minutes for a commercial site. This allow us to verify the cycling of those devices that we expect to cycle and to check the usage patterns of those devices. The data logger is then switched to a collection period of one hour and the verification checks repeated. Only after the verfication staff is convinced that the data is correct and processing is complete before the analyst allowed to access it.