

EPG Synchrophasor Data File Format (.CSV)

Version 3, 2012

1. Introduction

Synchrophasor data is used both in real-time from direct transmission and after the fact as recorded data. The SYNCHROPHASOR standard, C37.118, describes a real-time data transmission format but does not define a format for recorded data storage or exchange. Several formats have been used for synchrophasor data, most notably the Phasor File (also known as ‘dst’) data format pioneered in the WECC and supported by Bonneville Power Administration users.

The IEEE COMTRADE standard is a file format designed for time series data that is established worldwide and is supported by standards making bodies. It has a significant number of recording parameters that have been adapted for phasor data. COMTRADE file format is getting adopted quickly.

Phasor File and COMTRADE formats are binary files. It is not human friendly and hard to manipulate their content. It’s desirable to have a human readable and easy to manipulate file format for data exchange. The CSV (comma separated values) file format is a widely accepted format to store data in ASCII files in which data is stored in rows and columns. The data in rows are and delimited by comma (.). These files are human readable, and can be easily recognized by spreadsheet applications and other applications to perform data analysis and manipulation. This document proposes a format to store Synchrophasor data in CSV files, which is adopted by Electric Power Group’s synchrophasor applications, including PGDA, RTDMS, ePDC, and PMU Simulator. All EPG applications shall support importing/exporting synchrophasor data in this format.

Section 2 describes the CSV file naming convention. It is highly recommended for event data, though not required in many cases. Section 3 describes the header and data arrangement in the CSV file.

2. CSV File Naming Convention

The file name follows the convention defined in **IEEE PC37.232** standard. This file name convention defines a readable, comma delimited filename format. The “required fields” for the filename shall be as follows and in order as shown here:

Start Date, Start Time, Time Code, Station ID, Device ID, Company Name,

Additional fields may be added as needed and are called “user fields” are shown here:

User 1, User 2, User3, and so on .Extension

2.1 Start Date Field Format

This field is a fixed numeric format of (yymmdd). The start date is the date at which the first sample was recorded or the date at which the first trigger occurred. This is determined in the Time Code field.

2.2 Start Time Field Format

This field is a non fixed numeric format and can be specified to the required precision (hhMMssmmuuu). The start time is the time at which the first sample was recorded or the time at which the first trigger occurred, also determined in the Time Code field.

2.3 Time Code Field Format

This field is limited to a maximum of seven characters. The first character is the sign and is followed by up to five characters indicating the time difference between the time system used for the Start Date and Start Time fields (collectively, the Time Tag) and Universal Coordinated Time (UTC) without offset. The format for the five characters is as follows and in order: up to two digits for the hours followed by the letter “h” followed by two digits for the minutes. The last three characters are required only when fractional hours are in use. In addition, one final character, the letter “t”, should be concatenated at the end to the Time Code Field if the filename’s Time Tag is referencing the date and time of the first trigger. The calculation for the time difference should also consider whether standard time or daylight time was in affect at the time of the recording. The following examples provide a number of valid descriptions for the Time Code Field:

-4 *Time Tag is 4 hours behind UTC*
+5t *Time Tag is 5 hours ahead of UTC and references trigger time*
-7h15 *Time Tag is 7 hours and 15 minutes behind UTC*
0t *Time Tag is UTC and references trigger time*

2.4 Station Identifier, Device Identifier, and Company Name Fields Formats

Users may formulate their own codes for these fields.

The following fields are optionally recommended for packing meaningful information in the filename:

Duration Field Format:

The duration is equal to the time difference between the first and last samples in the file. The fields should be in the same format as the Start Date and Start Time Fields. For example a file with 2 years, 1 day, and 99 seconds of data will have a duration code of “020001,000139”.

Type Field Format:

This field may be used to describe the type of originating event and users may formulate their own codes. For example, the code “AG” may be used for a record created from a phase A to ground fault event.

Geographic Position Coordinates Field Format:

Two user fields are required in order to support position information. The first field is an expression of latitude and the second field is an expression of longitude.

2.5 Examples of File Names

- 120831,175215183,-4t,sta80,ben717,epg.csv

Note: The Start Time field is given in millisecond resolution and refers to when the first trigger in the record occurred with a COMTRADE configuration file extension.

- 120831,175215183,-4t,sta80,ben717,epg,000000,001359,uf,critical-frequency.csv

Note: This is a same example as above but with two user fields added and file duration fields added. “uf” added to represent that the file was created by a under frequency trigger and “critical-frequency” trigger also occurred in the record.

3. CSV file column header information

The first row contains column header information. The data block starts from the second row and continues to end of file.

3.1 The column header information format

The first row is the header and must start with Date Time and contain at least one signal column and ended with carriage return and line feed character:

DateTime(ISO Format),SignalName1,SignalName2,SignalName3,... <CR/LF>

3.2 Signal Name format

A signal name is comprised of PMU Name, Channel Name, and signal type which are separated by dots (.).

<PMU Name>.<Signal Channel Name>.<Signal Type>

Examples:

```
Bull Run.Frequency.FR
Bull Run.Frequency.DF
Bull Run.500 kV Bus +SV.VM
Bull Run.500 kV Bus +SV.VA
Bull Run.Watts Bar +SI.IM
Bull Run.Watts Bar +SI.IA
Bull
Run.B16;B15;B14;B13;B12;B11;B10;B09;B08;B07;B06;B05;B04;B03;B02;B01.DG
```

Note: A Digital signal has 16 bits in C37.118 standard. There are 16 channel names for a Digital signal to present every bit. These 16 channel names will be separated by semi-colon where the first channel represents the highest bit (bit 15) and last channel represents the lowest bit (bit 0). For example:

```
Bull Run.B16;B15;B14;B13;B12;B11;B10;B09;B08;B07;B06;B05;B04;B03;B02;B01.DG
```

Note: There is no name for Status flags in C37.118. Status should be used as the *Signal Channel Name*. For example:

```
Callaway.Status.ST
```

Note: There is no name for frequency in C37.118. Frequency should be used as the *Signal Channel Name*. For example:

```
Callaway.Frequency.FR
```

3.3 Signal Type format

The following code describes each signal type along with its unit:

Label	Description	Engineering Unit
<i>FR</i>	Frequency	Hz
<i>DF</i>	DF/DT	Hz/s
<i>VM</i>	Voltage magnitude (line to line voltage)	kV

VA	Voltage angle wrapped to (-180 180] degrees	Degree
IM	Current magnitude (line current)	A
IA	Current angle wrapped to (-180 180] degrees	Degree
PP	Active power	MW
PQ	Reactive power	MVar
DG	Digital	N/A
ST	Status flag, a 16-bit unsigned integer mapped to C37.118 STAT word.	N/A
AN	Analog	Depending on the signal
ZZ	Generic type other than specified above, such as for derived/calculated values	depending on the signal
UN	Unknown or missing information	depending on the signal

3.4 Column Header Row Example

```
Date Time, Callaway.Status.ST, Callaway.Frequency.FR, Collinsville.500 kV
Line.VA, Cordova.Line1.IM
```

4. Data information

4.1 Date Time Data in ISO 8601 Format

The first column is the Date Time information in human readable format following ISO 8601 standard. Exactly the components shown here must be present, with exactly this punctuation. Note that the "T" appears literally in the string, to indicate the beginning of the time element.

```
Year:
  YYYY (eg 1997)
Year and month:
  YYYY-MM (eg 1997-07)
Complete date:
  YYYY-MM-DD (eg 1997-07-16)
Complete date plus hours and minutes:
  YYYY-MM-DDThh:mmTZD (eg 1997-07-16T19:20+01:00)
Complete date plus hours, minutes and seconds:
  YYYY-MM-DDThh:mm:ssTZD (eg 1997-07-16T19:20:30+01:00)
Complete date plus hours, minutes, seconds and a decimal fraction of a
```

second

YYYY-MM-DDThh:mm:ss.sTZD (eg 1997-07-16T19:20:30.45+01:00)

where:

YYYY = four-digit year
MM = two-digit month (01=January, etc.)
DD = two-digit day of month (01 through 31)
hh = two digits of hour (00 through 23) (am/pm NOT allowed)
mm = two digits of minute (00 through 59)
ss = two digits of second (00 through 59)
s = three or four digits representing a decimal fraction of a second
TZD = time zone designator (Z or +hh:mm or -hh:mm)

There are two ways to handle time zone offsets:

1. Times are expressed in UTC (Coordinated Universal Time), with or without a special UTC designator ("Z").
2. Times are expressed in local time, together with a time zone offset in hours and minutes. A time zone offset of "+hh:mm" indicates that the date/time uses a local time zone which is "hh" hours and "mm" minutes ahead of UTC. A time zone offset of "-hh:mm" indicates that the date/time uses a local time zone which is "hh" hours and "mm" minutes behind UTC.

Note: to preserve high enough resolution, three or four digits are used to represent the decimal fraction of a second.

4.2 Signal Data

Signal data starts at the second column and continues till (n+1) column where n is the number of signals. For Status and Digital signals, data should be saved as 16 bits unsigned integer. For phasors, analogs, frequency, and MW/MVAR signals, values are in floating point. The signal data is required to be evenly spaced ordered by increasing date/time. If the data is unavailable for any signal for a particular time, that value should be set to -9999.

Example of data for frequencies:

```
2010-10-08T04:50:00.0000,59.994,59.994,59.994
2010-10-08T04:50:00.0333,59.995,59.992,60.003
2010-10-08T04:50:00.0667,59.994,59.995,59.991
2010-10-08T04:50:00.1000,59.993,59.992,59.975
2010-10-08T04:50:00.1333,59.995,59.992,60.006
2010-10-08T04:50:00.1667,59.994,59.991,59.993
```