Submitted for Presentation at the ECNE Engineering Conference
March 23rd, 2005; Portsmouth, NH

"Smart Data Concentrator (SDC) for Serially Accessed Devices in the Substation"

By: Amir & Maria Makki, George Semati (SoftStuf) & Tony Giuliante (ATG)

Abstract: the paper addresses the functional issues that are essential for automatic collection and for analysis of power system information from various types of serially accessible devices in the substation. Scalability issues with managing large numbers of devices across large numbers of substations are also discussed. The need for having and the advantages of using this type of automation are highlighted. Sources of critical power system information are identified, and references to the helpfulness of the IEEE-PES-PSRC file naming convention in managing the critical information are frequently made. The current limitations and potential applications are described, and examples of past projects involving varying types of devices and vintages are provided.

Keywords: time sequence data, TSD, smart data concentrator, SDC, serially accessed device, smart device, digital fault recorder, DFR, digital relay, sequence of events recorder, SER, master station, network, topology, protocol, format, driver

Definitions

Time Sequence Data (TSD) File: is any type of electronic data file where each data item in the file corresponds to an instant of time that is identified by an explicit or implicit time tag, such as, transient data files, sequence of events files and periodic data files.

Smart Device: is a microprocessor based serially accessible device that is used to protect the power system and to measure critical system information. Examples of smart devices include, but are not limited to, digital fault recorders (DFR), digital relays, sequence of events recorders (SER), and transformer monitors. In general, smart devices produce extremely large quantities of TSD files.

Smart Data Concentrator (SDC): is a program for automatic collection and analysis of the large quantities of TSD files from the various types of smart devices in real time. The program notifies the proper personnel immediately upon detection of any abnormal conditions, such as, targets, fault location, overload, imbalance, and inefficiency.

Introduction

Utilities are benefiting from the introduction of smart devices into the substation environment. Utilities have replaced the light beam oscillograph with DFR's that are now capable of capturing, analyzing, and archiving faults more rapidly, accurately, and efficiently. Electro-mechanical relays have given way to digital relays that can clear faults faster, employ fault location algorithms, and provide oscillography records. Digital meters and monitors have power quality functions, provide trending abilities, and can be viewed and acted upon remotely. Data from these smart devices is used to check and update system models, replay suspicious relay operations, monitor evolving loads, and develop better restoration and contingency planning procedures.

Unfortunately, developers of smart devices have created too many distinct protocols using different operating platforms and software packages. Due to this rapid growth and expansion of digital technology within the substation new operating deficiencies have developed. The diverse and incompatible nature of each manufacturer that was once considered of minor consequence is now of major concern. Rising communication costs, limited manpower resources and expanded computer knowledge and acceptance in the substation environment have generated genuine interest in a comprehensive integration solution.

The dilemma or deficiencies that have developed are basically associated with the integration, communication, storage, and analysis of TSD files. To benefit from the strength of these new smart devices, it is essential to integrate and network them onto a common platform. The successful integration will enable users to access various types of data using a common interface, and will provide the infrastructure needed to support future automation projects. The technology to integrate and automatically manage a large number of smart devices is currently available, and is cost effective. Today's computers and multiport networks offer very high performance at low prices. It is time to explore the potential of this technology for improving overall substation operations.

Integration & Communication

State of the art integration and communication technologies offer real time capabilities with point-to-point connectivity. You can simultaneously gather and analyze information from multiple devices in real time. Old dreams of integrated operations are now being realized. A host of new applications such as remote energy management, advanced SCADA functions, and integrated protection and remedial action systems are currently being explored. These are exciting times for designers and operators of smart devices. A broad outline of the various types of integration and communication schemes for smart devices is presented bellow:

Networks: integration of smart devices is accomplished by organizing them into groups of local area networks (LANs), one LAN in each substation. The LAN could be a pure RS232 type network, or an RS485 type network, or it could be a pure Ethernet network or it could be a complex hybrid combination with multiple networking schemes.

Topologies: smart devices could be connected using a multi-drop or using a star type topology, or they can be connected in a hybrid way. In the multi-drop case the devices are placed on the same communication line, only one device can talk at any given time.

In the star topology case all of the devices can talk at the same time because they are placed on separate communication lines. During a storm, many devices may begin to report their fault information at the same time.

A pure star topology is therefore the most suitable topology for integrating smart devices. In general, an RS232 network is a star topology network, an RS485 network is a multi-drop network, and an Ethernet network can be either star or multi-drop depending on the type of addressing scheme (if each device has a unique internet protocol “IP” address, then the LAN is a star topology).

Integration: in order to integrate the LANs from multiple substations, companies have historically chosen the phone system for their remote communications. A modem is placed in each LAN and then a bank of host modems is made available in the office. The host modems are usually placed as standalone peripherals on some unmanned computer in the office called the master station. The master station services the host modems on a 24/7 basis and acts as a “firewall” between the host modems and the office network. The host modems should not be configured as “open modems” so as to block any potential saboteurs from gaining remote access to the office network.

Some companies also add a computer to each substation LAN. Such computers are used to continuously poll the smart devices and look for trouble spots. Upon detection of trouble, the substation computers will use the LAN modems to immediately dial out and report their findings to the host modems (this feature is called “report by exception”). Alternatively, the computers can be configured to hold onto their findings until the user asks for them (this is called “report upon demand”). The computers report their findings by saving a new set of summary and supporting TSD information onto a shared folder on the office network. Such computers are called substation data concentrators, or substation communication processors, or substation gateways, or just SDCs as defined in this paper.

The master station and the SDC computers do not have to be "super-duper" workstation or server platforms. Any computer with a Gigahertz processor, 256 Megabytes of RAM, 100 Gigabytes of available hard disk memory, and a few Ethernet ports is more than adequate. A graphical depiction of the overall integration scheme starting with the smart device in the substation and ending with the users on the office network is shown in Figure-1 below.

Protocols: the order in which bytes are transmitted or received is called the protocol. In general, protocols are divided into two categories: a protocol is either standard (MOD-BUS, DNP, UCA, FTP, TCPIP) or proprietary (each manufacturer defines their own). An extensive library of protocols (software drivers) may be required to automatically communicate with multiple types devices using various types of drivers. Each driver in the library corresponds to a specific type of protocol for a specific type of smart device.

Communications: smart devices communicate at low baud rates. Typical rates are: 2400, 4800, 9600, or 19200 baud. When a large array of smart devices, say 128 devices, are considered collectively the resulting data transfer rates may exceed 100 kilobytes per second. However, today's computers can easily handled such rates with ample time left over for other functions such as processing, archiving, analysis, and reporting. For example, the data transfer rate for Industry Standard Architecture (ISA) is 2.4 Megabytes per second or 32 Megabytes per second for Enhanced ISA (EISA).

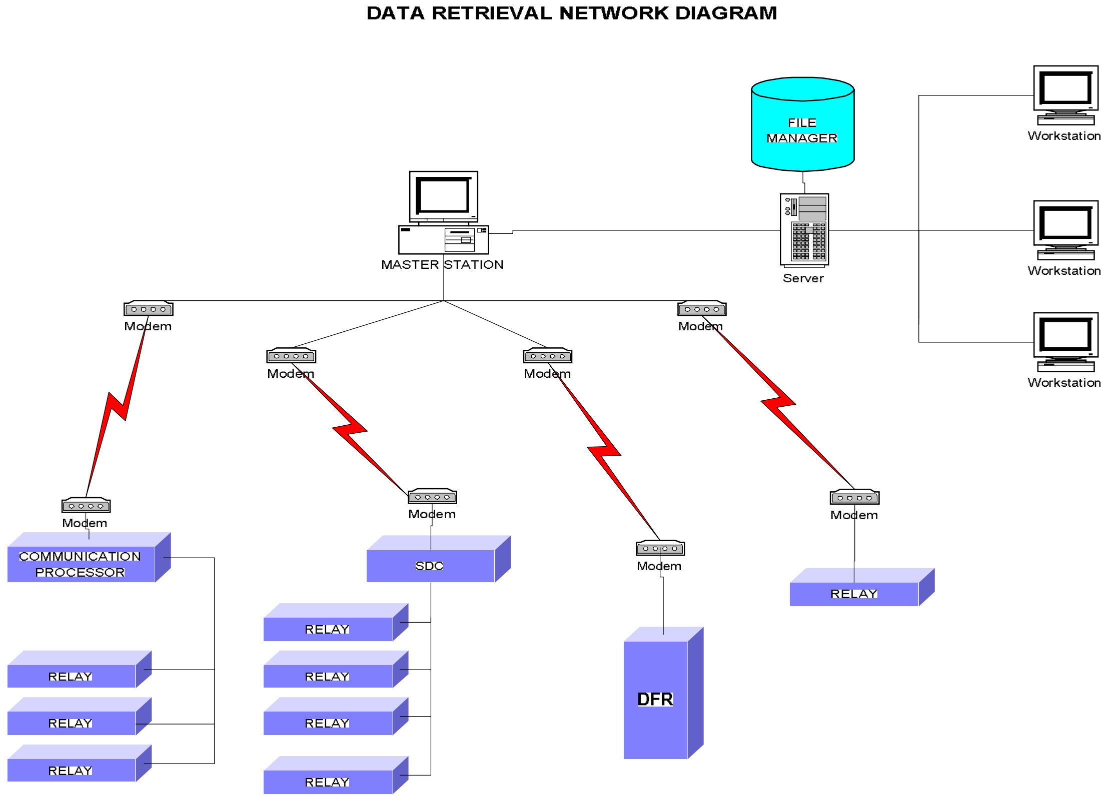

Figure-1: Integration of Smart Devices and Communications Architecture

The typical methods for communicating with smart devices are: 1) locally, by going to the device and down loading to a portable computer, or 2) remotely, by using a dial-up or direct connection to the device. As advances in communication networks continue to impact the design of smart devices, more and more electronic data about the power system is being generated. All of the critical system information that we have always wished for is actually measured or calculated and archived by today's smart devices.

However, automatically collecting large volumes of data from smart devices is not a simple task. The data overload problem is clearly a visible downside. The nature of this data is complex and requires the use of new management tools and standards in order to achieve the goal of total integration. In what follows a survey of the TSD types and a description of the latest in TSD management tools and standards is presented and a novel solution to the data overload problem is proposed:

Storage & Exchange

Once the substation devices are properly integrated a new world of automation applications emerges. DFR records can be automatically analyzed and “cross checked” with events from digital relays and SERs, upon occurrence. Decisions can then be made about the state of the power system and acted upon automatically whenever an event has occurred. The decision making capabilities of future automation systems will only be limited by the amount of knowledge (rules, methods, templates) that are built into the software. Current expert systems allow many years of expertise and experience to be programmed and organized into knowledge bases. With this type of technology true substation automation is only a few steps away, but what is the first step?

The first step is to automatically collect the critical data from the smart devices. The collection process includes and is not limited to: polling the smart devices with the appropriate protocols, extracting the critical data, and finally saving the data (or concentrating the data) onto a common platform. Theoretically, this is not a very difficult task to do because the critical data is already archived in the smart devices and the data exists in the form of TSD files. So, basically, automatic collection means copying or moving the files directly from the smart devices and saving them onto a shared folder on the office network (via the master station and the SDCs). In real life, however, there are many types of TSD files, storage formats, and naming conventions to contend with:

Oscillography Files: these are high speed, transient data records with sampling rates in the 5 kHz range. The records are used to electronically preserve snap shots of the voltage and current phases during fault and disturbance conditions. The records are primarily used for postmortem analysis just like in the case of the airplane’s “black box” recorder. The transient data records are also used to measure angles and harmonics, to study status and event sequence, and to calculate fault location, type, duration, and magnitude. An example of a transient data record is shown in Figure-2 below.

Periodic Log Files: these are low speed, steady state data logs with a commonly accepted sampling frequency of once every 15 minutes. The logs are used to preserve instantaneous measurements of the load flow and power transfer conditions and of the related environmental factors as they vary in time. Amps, votls, watts, vars, efficiencies, total power in/out, peaks, imbalances, overloads, ambient temperature and relative humidity, transform temperature and relative saturation, and vegetation overgrowth are all good examples of periodically logged measurements and factors. An example of a periodic data log is shown in Figure-3 below.

Sequence of Events Files: these are readable, ASCII text files that provide a chronological summary of the event occurrences during a particular fault or disturbance condition. An event could be a target, or an alarm, or a breaker operation, or it could be a communication signal or an error. Each event in the file has it’s own time tag and the resolution is typically accurate to the millisecond. An example of a sequence of events file is shown in Figure-4 below.

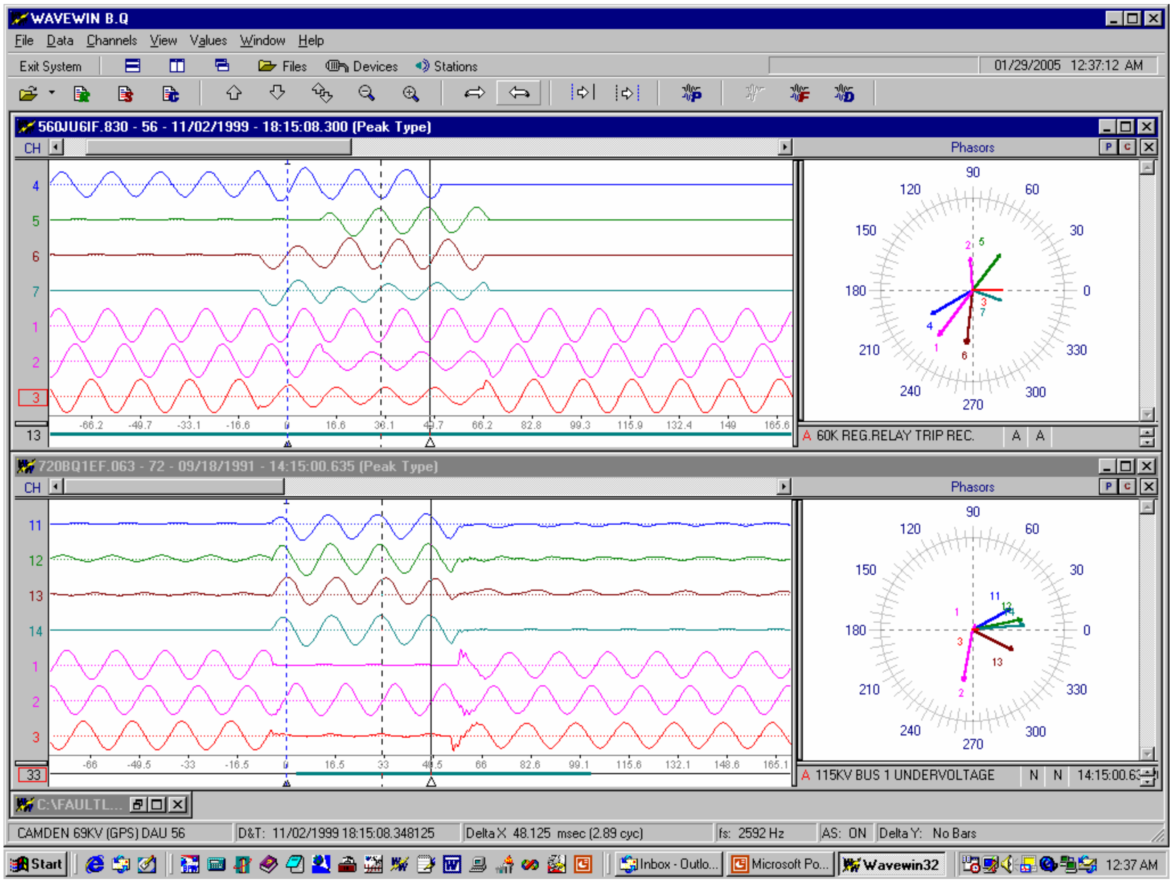

Figure-2: Oscillography Example Showing Synchronized TSD Files

Configuration Files: these are device specific files that include information about the configuration parameters (scale factors, sampling frequency, communication rate, and setting groups), or include information about testing and maintenance operations. The setting formats and the configuration procedures vary from device to device depending on the manufacturer and the application. The testing procedures and methods for replaying transient records into smart devices also vary by test-set depending on the manufacturer. Examples of such device information files are shown in Figure-5.

TSD Formats: there are too many types of proprietary TSD formats in circulation today (ASCII and binary). The formats vary depending on the manufacturer and type of originating device. Figure-6 shows a few examples of proprietary TSD file formats from various types of smart devices. As for standards, there are only a few formats in circulation today and the most popular of them all is the standard IEEE C37.111 COMTRADE format. In addition, an event may be composed of one or multiple TSD files or multiple events may be saved in the same TSD file. In other words, events and TSD files have a many to many relationship. For example, the COMTRADE standard requires three files for each fault record: the header, configuration and data files. The files share the same name but have different extensions: "HDR", "CFG" and "DAT":

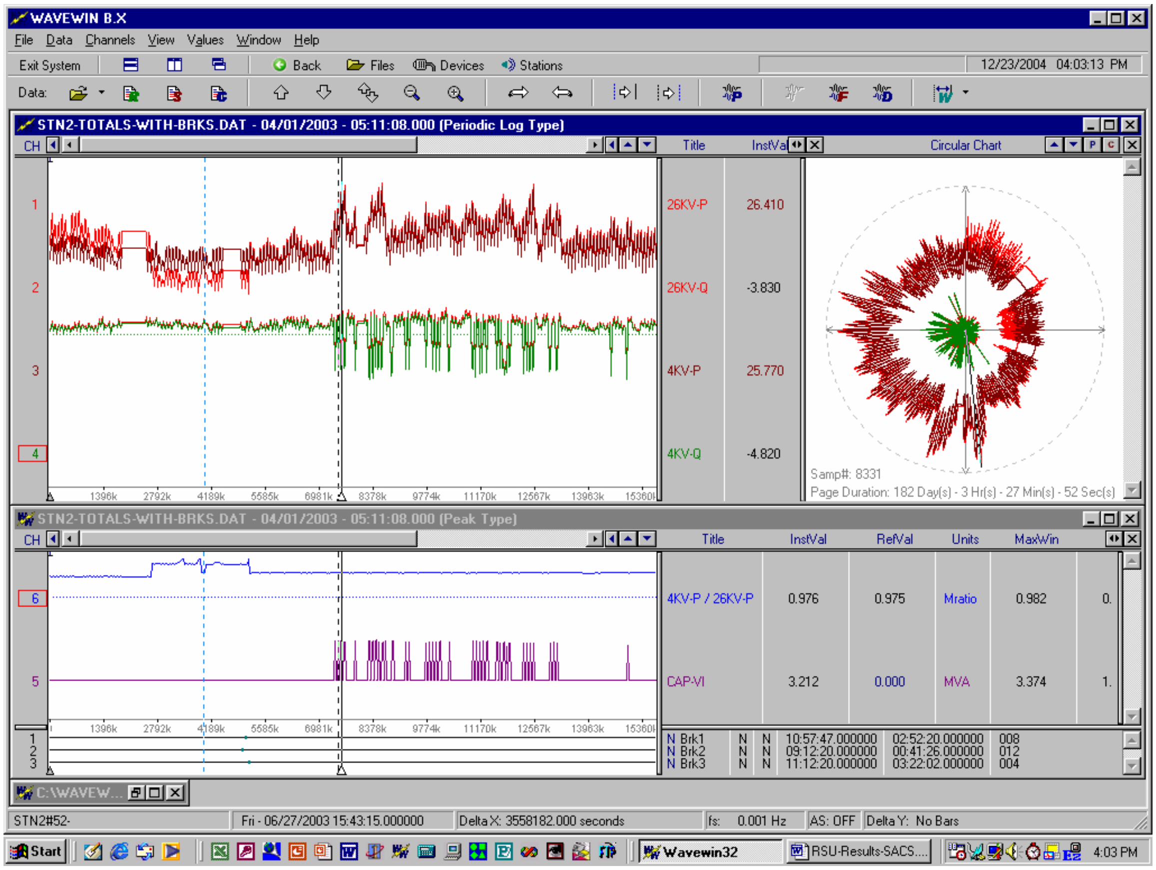

Figure-3: Periodic TSD Log File Showing Total Substation Power flow (In/Out)

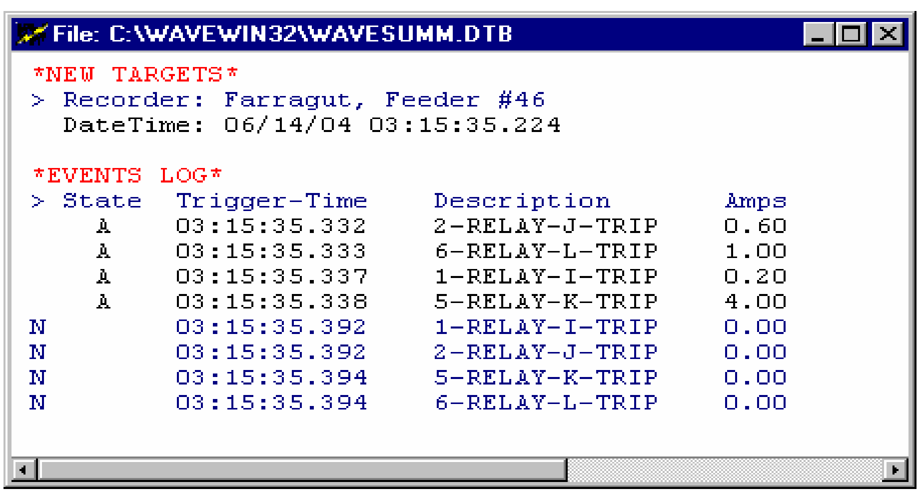

Figure-4: A TSD File Example Showing A Typical Sequence of Events Report

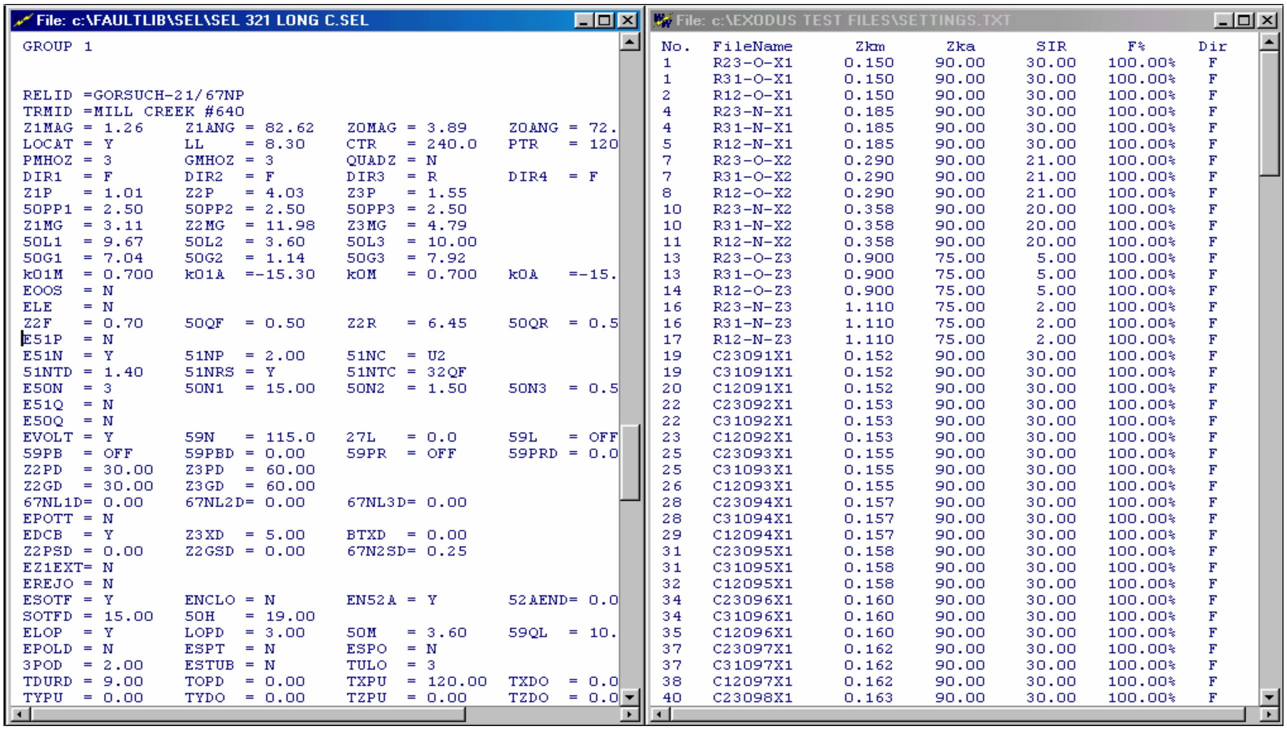

Figure-5: TSD File Examples Showing Relay Settings & Test Plan Information

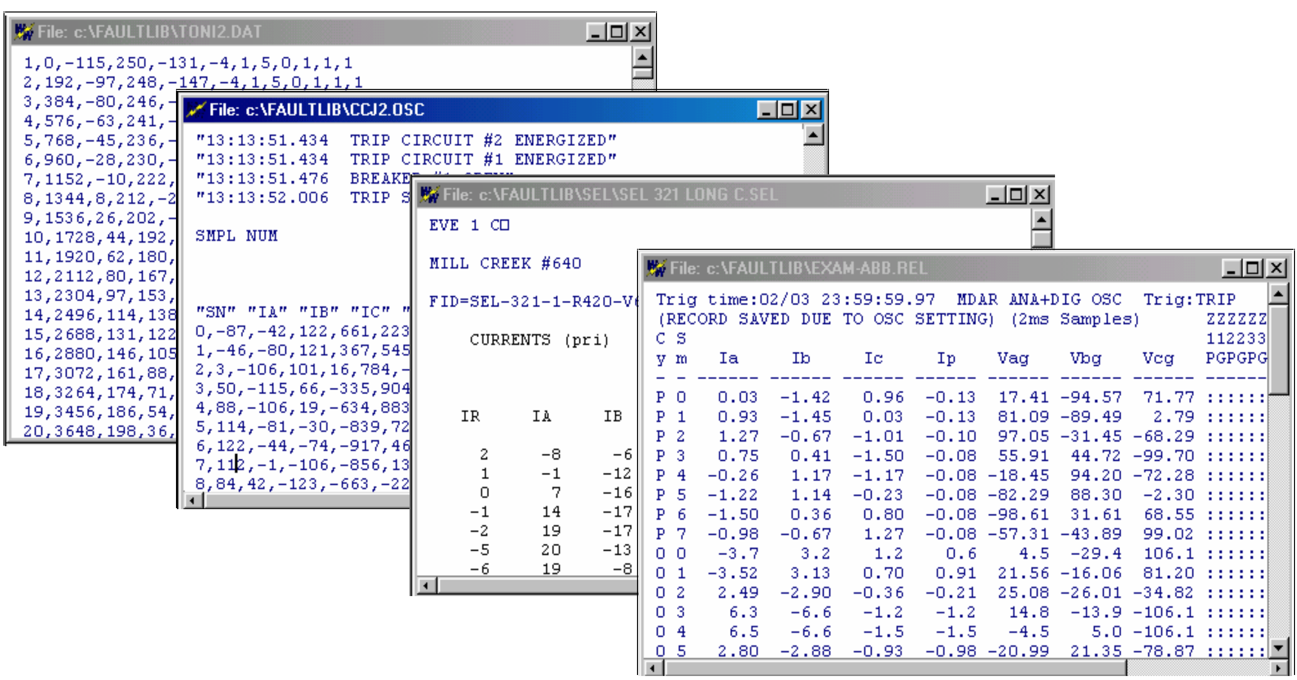

Figure-6: TSD File Examples Showing Various Types of Transient Data Formats

TSD Filenames: today there are multiple types of TSD file naming formats in circulation making it difficult to copy TSD files from various types of devices onto the same network folder. The names vary depending on the convictions of the manufacturer and the type of originating device. In general, the current naming conventions can be organized into three classes. The classes are: associated, coded, and sequenced.

Associated means the filename extension defines the type of storage format. For example, the extensions "HDR", "CFG", "DAT", and "INF" are used to indicate that the file contents are compatible with the 1999 COMTRADE standard. The non-extension part of the filename, the name, is left at the discretion of the user.

Coded means the filename contains information about the fault record. The contained information is usually manufacturer specific. The resulting filenames are not friendly and require special decoding software. For example, “G30BQ1EF.063” is the filename assigned by recorder number 163 on 09/18/1991 at 14:15:00.630.

Sequenced means that the filenames are incrementally assigned. This method is valid because the resulting filenames are unique, however, the total number of attainable filenames is limited to the maximum value of the numerical sequence. For example, a file named "MART1743.RCL" indicates that the event was recorded at Martin Station and the event number is 1743. The "RCL" extension means the data contents are from the first 16 analog input channels.

In order to facilitate the storage and exchange of TSD files, the IEEE-PSRC-PES has developed a “content addressable” naming convention where the filename contains a sufficient amount of information for users and automated analysis applications to be able to determine the extent of the TSD file’s contents by just reading the filename.

The IEEE naming convention specifies a content addressable format that can be used to uniquely tag and identify each TSD file within a database (or repository). The naming convention is as follows:

“Start Date, Start Time, Code, Station, Device, Company, User1, User2, etc. Ext”

An example filename is:

“000809,175215183,-4,Sta80,Line717,nyiso,AG-trip,024miles,Zone1.OSC”

There are many applications that can be realized given a unique, informative file naming system. Possible applications are: time-line manager, universal viewer, global data bank of fault records and so on. Utilities, independent operators, reliability commissions and manufacturers can now realize the benefits of sharing a common electronic filing system. Take any record from any device anywhere on the planet and seamlessly file it.

The Data Overload Problem

Clearly, the data overload problem is a visible downside of the automated data concentration experiment. In WWII the British Navy declared: “we can’t be everywhere, but we can be anywhere at anytime”. Same thing here, we have too much data from these smart devices and it is impossible to manually examine every byte trying to discover where the trouble spots are. We need to use artificial intelligence (AI) to help analyze circuit conditions and classify circuit behavior. Solving the data overload problem is about using AI to locate where the trouble spots are and automatically notify the proper personnel.

However, as shown above, these smart devices have too many types of data formats, communication protocols, and networking topologies. Universality, therefore, is also an essential component in data concentration. Universality is needed to allow for seamless access to TSD files from the microsecond level to the yearly level. The aim is to solve all of the deficiencies in a universal and intelligent way. Accordingly, the term “smart” in SDC was specifically chosen to echo universality and intelligence. The successful SDC program will automatically collect TSD files, analyze them, locate trouble spots and report them. The successful program should also have a universal viewer that provides a common platform for manual display and analysis of TSD files as shown above in Figures 2 & 3. After all, any allegations brought forth by the AI networks should always be manually verified and the universal viewer is essential for that.

One a recent project, we used 5 SDCs and a Master Station to collect hundreds of mega bytes worth of periodic load data from all circuits in 5 substations covering the periods 2003 and 2004. The collected data was analyzed by using the SDC AI networks of sensors, filters, rules and methods that are collectively aspiring to classify circuit behavior. The SDC program was able to automatically spot and classify abnormal conditions such as imbalance, overload and inefficiency at the circuit and at the substation levels. The system ran without error and the performance was accurate and repeatable. The issued priority values for each circuit were evenly distributed ranging from 0 to 646. The maximum possible priority value was 1,000. Only (1) circuit scored above 600 and six (6) circuits scored above 400. A total of 95 circuits were considered in this project. The system planners were able to quickly identify that 1 out 95 circuits that needed to be examined and were also able to use the SDC’s universal viewer to manually scrutinize the data and verify the results of the AI networks.

Comments

The main concept behind the SDC is to provide a specialized data collection and analysis system that can assist in the process of extracting critical reliability information from the vast wealth of available TSD files. The SDC program is used for fast line restoration, fault and disturbance analysis, dynamic relay testing, real time monitoring of evolving loads, contingency planning, and intelligent maintenance applications.

The main SDC objectives are to increase system reliability, reduce engineering time, and minimize data loss. By continuously polling such devices, the SDC program helps manage the risk of losing valuable data. The program interfaces help engineers expose faulty wiring, defective devices, bad configurations, false relay operations, nasty harmonics, unbalanced circuits, overloaded assets and so on.

As Utilities continue to replace their aging infrastructure with smart devices the need to integrate all substation devices onto a user-friendly platform becomes imperative. The ability to interrogate DFR's, microprocessor relays, digital meters, and monitors quickly and efficiently becomes a high priority in a competitive, customer oriented and down sized environment. Integrating and utilizing the present microprocessor arrangements for capturing data such as relay targets, or oscillography data will add value to the corporation. Utilizing this method of integrating devices from various manufacturers allows for secure substation automation using off the shelf technology and provides operators with the choice of selecting from manufacturers that produce the best equipment for the application.

The number one beneficiary is maintenance. The SDC program is designed to run unmanned for very long periods of time. The program works with various types of proprietary and standard formats and protocols. This eliminates the need for having to learn a diverse mixture of products and individual operating nuances. Captured data is automatically archived and system abnormalities are reported to the office immediately. And, all of the employed hardware and software components are based on market available open architecture technology. If you need more space or speed, buy a bigger or faster computer. If you need more ports, buy an extra panel, plug and play. Such technology has never been made available before and will have a considerable impact on the way in which substation equipment is operated in the future.

References:

“Universal Data Retrieval”, Pond & Makki, Proceedings of the Distributec Conference, Jan 2005

"Transient Information Management Efficiency,” A&M Makki, M. Taylor, Proceedings of the Fault and Disturbance Conference, Georgia Tech, May 2001

"Naming Convention For Time Sequenced Data Files," H8 working group report for the IEEE Protective Systems Relaying Committee (PSRC), April 2001

"Survey of Event File Saving Schemes," A&M Makki, M. Taylor, L. Johnson Proceeding of the Fault and Disturbance Analysis Conference, Atlanta, Georgia, May 1999

"Survey of Event File Naming Schemes," A&M Makki, M. Taylor, Proceeding of the Fault and Disturbance Analysis Conference, Atlanta, Georgia, May 1998

"IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems," IEEE Standard P37.111, 1999

"Overview of FAT, HPFS, and NTFS File Systems," Microsoft Windows NT Workstation 4.0 Resource Kit, Chapter 18, V.1997

"Low Cost Data Concentrator for Serially Accessed Devices;" Proceedings of the Fault and Disturbance Analysis Conference, Washington, D.C., November 1996

"Performance Considerations for High-End PC A/D Boards" Evaluation Engineering, App. Note A, February 1995