FRIDAY, 30 NOVEMBER 2001

FRA - BEAM DIAGNOSTICS AND TIMING SYSTEMS

FRAT001
cs.AR/0111032

SNS TIMING SYSTEM

R. Nelson LANL, B. Oerter BNL, T Shea ORNL, C. Sibley ORNL

Abstract

A modern physics facility must synchronize the operation of equipment over a wide area. The primary purpose of the site wide SNS synchronization and timing system is to synchronize the operation of the LINAC, accumulator ring and neutron choppers and to distribute appropriate timing signals to accelerator systems, including the Injector, LINAC, Accumulator Ring and Experimental Facilities. Signals to be distributed include the ring RF clock, real-time timing triggers, machine mode and other informational events. Timing triggers and clocks from the SNS synchronization and timing system are used to synchronize hardware operations including the MEBT beam chopper, RF turn on, synchronous equipment state changes, as well as data acquisition for power supplies and beam diagnostics equipment. This paper will describe the timing equipment being designed for the SNS facility and discuss the tradeoffs between conflicting demands of the accelerator and neutron chopper performance due to AC power grid frequency fluctuations.

1 REQUIRMENTS

1.1 Synchronization of Neutron Choppers [1]

The neutron choppers are very high inertia mechanical rotors with blades that chop the neutron beam. To prevent modulation of the neutron energy spectrum, chopper operations must be phased with respect to the arrival of protons on the target, which is only a short transport delay from beam extraction. Overall operation of the facility is greatly simplified if the accelerator and storage ring can also be phased with extraction. For the several types of choppers, Fermi, T_0, and bandwidth limiting choppers, the most stringent desired timing accuracy for the synchronization is about ±0.5µs. The required accuracy is still to be determined, but is in the ±1 µs range.

1.2 Synchronization of LINAC

In order to inject approximately 1060 turns into the ring, the beam injection into the LINAC must begin about 896 µs before extraction. For a super conducting LINAC, the klystron modulators must be pulsed about 400 µs earlier than the beam. The beam diagnostics group have requested a pulse, approximately 5500 turns before the ring extraction time.

Klystons in the LINAC require synchronization with the power grid. Failure to maintain this synchronization results in non-linear performance and beam losses. The beam in the LINAC injector is chopped by a gated electrostatic focusing electrode in the LEBT and a traveling-wave deflector (beam chopper) in the MEBT. Each of these systems, and especially the beam chopper in the MEBT, must be synchronized to the ring period. The reason for chopping the beam is to produce a gap in the beam current that is synchronized to the ring period.

The beam occupies about 67% of the ring circumference, and the remainder of the ring (the beam gap) is free of beam (to about 1 part in 10^4) in order to accommodate "clean" extraction. The extraction from the ring must be synchronized to the beam to about ± 5 ns (the beam gap is about 300 ns long).

1.3 Ring Revolution period

During normal operation, tuning the ring by changing the B field (which changes the orbit radius), or tuning the LINAC to change the injection energy, or a combination of both, can change the revolution period by up to ± 1 ns. Although 2 ns is not a large number, the cumulative timing error over a normal accumulation cycle of 1060 turns exceeds 2 µs, more than one complete ring period. This is much larger than the desired synchronization accuracy of ± 5 ns for accumulating 1060 turns of beam in the ring. A beam synchronous timing system tied to the rotation frequency of the ring will compensate for this variation in fill time and maintain a ±5 ns timing accuracy. The system designed for SNS will operate over an energy range of 950 MeV to 1.3 GeV without modification

2 BEAM DIAGNOSTICS

2.1 Ring

The beam diagnostics in the ring measure the beam during the accumulation cycle. The ring beam diagnostic measurements include global systems such as a beam closed-orbit measurement (BPM) system, and local systems such as the azimuthal distribution of beam charge density (current monitor). Global and local refer respectively to systems that are widely distributed around the ring, or are needed in only one or two locations. These systems need to be synchronized to the beam in the ring to within about ±5 ns, even though the accumulation period may vary by 2000 ns or more. Diagnostics in the beam line between the LINAC and ring requires synchronization to the ring period. Diagnostics in the beam line between the ring and the spallation target need only a single pretrigger a few turns before extraction.

2.2 LINAC

Although the LINAC RF is not synchronized to the ring, the periodic beam gap in the LINAC is. Because of the effect of the beam gap on the performance of the LINAC beam diagnostics, the sampling of the beam diagnostics signals from all the diagnostics needs to be synchronized relative to the beam gap. In particular, beam current, beam position, and beam synchronous phase measurements would benefit from being synchronized relative to the beam gap. Thus the LINAC diagnostics also need to be synchronized to the beam in the ring. A 60-ns timing granularity (not jitter) of a timing signal relative to the beam gap is adequate (the beam gap is 300 ns long, and the beam mini pulse between beam gaps is 645 ns long).

3 TIMING SYSTEM ARCHITECTURE

3.1 Synchronization

Experience at LANL spallation source [3] has shown that the LINAC, ring, and neutron choppers can synchronize operation with each other and follow trends in the phase fluctuations of the ac power grid. Each subsystem operates as a slave to a master timing generator which smoothes phase fluctuations present in the power grid. At SNS the competing demands of the LINAC that limits phase differences with respect to the power grid and of the neutron choppers that limit the acceleration or deceleration of high-inertia rotors can be simultaneously satisfied within a “phase” window of ±500 µs measured with respect to the power grid.

3.2 PLL

The master timing generator will implement a phase-locked loop (PLL) that will follow the phase fluctuations of the power grid and produce the Cycle Start signal with “smoothed” phase fluctuations. The residual fluctuations should not exceed the ability of the neutron choppers to maintain phase lock with Cycle Start or with signals derived from the timing distribution system. Similarly the residual phase fluctuations must not exceed ±500 µs to ensure correct operation of the klystrons in the LINAC.

The ac-line synchronization requirement should be established by an operational parameter varied under computer control. This parameter characterizes the variance in the time interval between the zero crossing and the Cycle Start signal, even during line frequency transients. Over its full range this parameter will provide for phase coupling from that is essentially identical to the power grid through phase coupling with a standard deviation of 125 µs. While this parameter may be adjusted for certain measurements, normally its value would be fixed.

To synchronize the accelerator systems and the neutron choppers, a phase-stable timing signal is generated and distributed around the facility. **Cycle Start** will be defined. **Cycle Start** shall be generated some number, of ring-rf cycles before extraction. The actual number might vary depending on the number of micropulses injected into the ring. The phase of the ring RF will be resynchronized at **Cycle Start** while the accelerator is idle and the ring is empty. Once established at **Cycle Start**, the phase and frequency are held fixed until after the beam is extracted.

Should the master timing generator lose synchronization as indicated by the **Cycle Start** signal drifting too far from the zero crossing, then a sync lost pulse will be generated for each cycle until synchronization is regained. Conditions for defining lost synchronization are TBD.

To synchronize extraction of the beam from the storage ring with the rotation of neutron choppers, the master timing generator will first determine when beam should be extracted for the next beam pulse. Then the timing generator must schedule generation of the **Cycle Start** signal allowing for N_{cs} ring-rf cycles prior to extraction. The **Cycle Start** signal will reset the phase of the ring rf to zero.

3.3 Event Link Timing Distribution

The most convenient way to *broadcast* timing signals to many hundreds of clients distributed over a kilometer or more of accelerator and beam lines is to encode all the timing signals using a self-clocking scheme in which the timing signals are encoded on the carrier, and distributed to all clients. If the carrier frequency on which the events are encoded, is synchronous with the beam in the ring, rather than a fixed frequency, then timing triggers and delays derived from the carrier frequency maintain the proper relationship with the beam revolution period. This eliminates the need to adjust timing delays due to changes in the ring revolution period. SNS will use a carrier frequency of 16 times the ring revolution period.

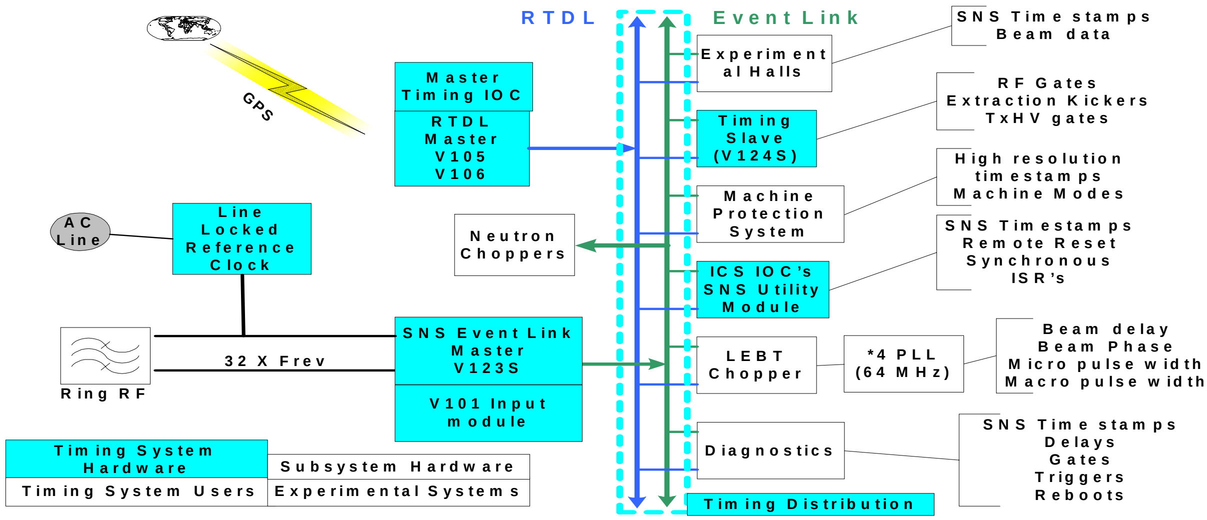


Figure 1 SNS Timing System Distribution

3.4 Real Time Data Link

The RTDL system is a master slave system used to broadcast data in real time. IOCs will have access to RTDL broadcasts through the Utility Module. A maximum of 256 data frames can be defined. The RTDL encoder contains a list of frames to be transmitted. All frames contained in the list are sent out prior to the start of each accumulation cycle. A list of the defined frames is shown below.

1. Time of day (IOC timestamps)
2. Ring revolution period (in ps.)
3. Operating Mode
4. 60 Hz phase difference
5. Beam Parameters for LEBT chopper
6. Previous beam pulse data
7. Data acquisition mode
8. Beam profile ID
9. Transmitter, RF Gates ON/OFF
10. IOC reset address

3.5 Time stamp

A commercial VME module resident in the RTDL chassis receives time of day information from a GPS receiver. At **RTDLStart** the time will be updated in the RTDL system for broadcast to all receivers. Thus each machine cycle shall have a unique time of day timestamp. For systems requiring a timestamp with greater resolution within a machine cycle, hardware counting at some multiple of the ring revolution frequency shall provide the additional fractional machine cycle component of the time stamp.

3.6 Utility Module

Each IOC will contain a Utility Module. The Utility Module provides a number of VME chassis services including receiver circuits for the Eventlink and RTDL links. The Utility Module can be configured to initiate VMEbus interrupts on the detection of specified events. Data transmitted on the RTDL are stored in local memory. One of the RTDL frames provides remote reset of VME chassis. All Utility Modules monitor this frame for their remote reset address. When the reset code matching an IOCs Utility Modules preprogrammed reset address is received, the Utility module asserts the VMEbusSYSRESET/ signal

4 STATUS—CONCLUSION

The SNS Timing System design is based on the RHIC timing system. Using this proven design minimizes new design risks and will result in a system with proven performance and reliability. Modifications for the SNS requirements have been straightforward. Prototypical systems are running at BNL and SNS. Integration Software (Run Permit System, see SNS Run Permit System Poster at this conference) in progress using hardware and software supplied by BNL.

REFERENCES

[1] B. Oerter "SNS Timing System Requirements Document" Internal Documentation (SNS 109020000-SR0001 Rev1)

[2] C. Sibley, "SNS Run Permit System", Paper TUAP023, This Conference

[3] R. Nelson, R. Merl and C. Rose "Timing Reference generators and Chopper Controllers for Neutron Sources" 15th meeting of the International Collaboration on Advanced Neutron Sources, November 6-9 2000, Tsukuba Japan

FRAT002
physics/0111049

BEAM DIAGNOSTICS SYSTEMS FOR THE NATIONAL IGNITION FACILITY

R. D. Demaret, R. D. Boyd, E. S. Bliss, A. J. Gates, J. R. Severyn
LLNL, Livermore, CA 94550, USA

Abstract

The National Ignition Facility (NIF) laser focuses 1.8 megajoules of ultraviolet light (wavelength 351 nanometers) from 192 beams into a 600-micrometer -diameter volume. Effective use of this output in target experiments requires that the power output from all of the beams match within 8% over their entire 20-nanosecond waveform. The scope of NIF beam diagnostics systems necessary to accomplish this task is unprecedented for laser facilities. Each beamline contains 110 major optical components distributed over a 510-meter path, and diagnostic tolerances for beam measurement are demanding. Total laser pulse energy is measured with 2.8% precision, and the interbeam temporal variation of pulse power is measured with 4% precision. These measurement goals are achieved through use of approximately 160 sensor packages that measure the energy at five locations and power at three locations along each beamline using 335 photodiodes, 215 calorimeters, and 36 digitizers. Successful operation of such a system requires a high level of automation of the widely distributed sensors. Computer control systems provide the basis for operating the shot diagnostics with repeatable accuracy, assisted by operators who oversee system activities and setup, respond to performance exceptions, and complete calibration and maintenance tasks.

1 INTRODUCTION

For the NIF to achieve its performance goal of 1.8 Mega-joules of ultraviolet light on target requires that each beam path is carefully controlled and accurately characterized. Each beam will be operated at up to 80% of its fluence damage threshold and still need to meet the System Design Requirement (SDR) of ≤8% rms deviation in the power delivered. This places stringent derived requirements on the energy, power temporal shape, and fluence measurement systems deployed in the NIF. Table 1 lists the principal requirements for measurements of pulse energy and power, and monitoring of the spatial profile of beam fluence.

Table 1: Tolerances for key diagnostics tasks

Beam Diagnostics Measurements Requirements	
Measure pulse power of all beams (power balance)	The rms deviation less than 8% of the specified power averaged over any 2-ns time interval [1].
Measure pulse energy at 1.053 and 0.351μm	2.8%
Measure pulse temporal shape versus time	2.8% with ≤ 450psec rise time
Record the spatial profile of beam fluence	2% fluence resolution,1/125 of beam spatial resolution

Sensor packages and diagnostics sensors complete the principal diagnostics hardware. Supporting front-end processors (FEP) and diagnostics electronics capture and process the data. The Integrated Computer Controls System (ICCS) [2] will provide the basis for completing shot preparations with repeatable accuracy on a timely basis.

These beam diagnostic functions are accomplished with optical-mechanical and electronic components distributed along each beamline. Figure 1 identifies these components and illustrates the fact that the beam control systems have interfaces with every part of the laser.

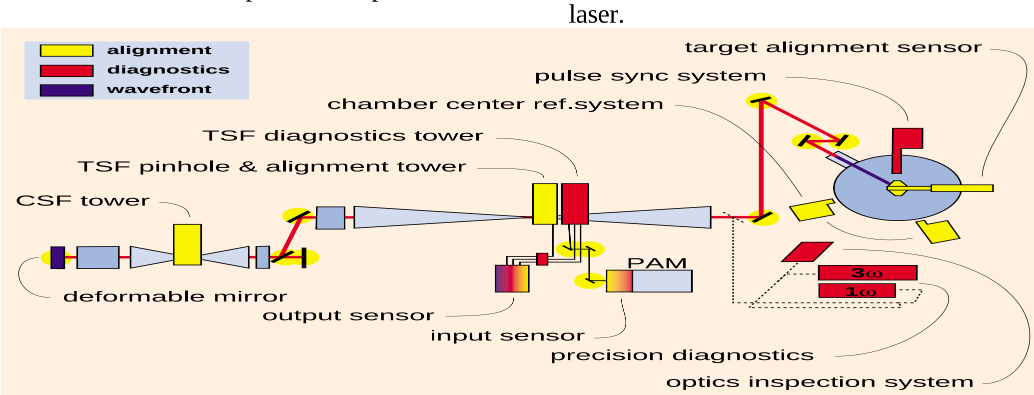


Figure 1: Optical, electronic, and mechanical components distributed along each beam perform beam diagnostic functions. Towers hold components for laser diagnostics functions on eight beams. In the figure, only one beamline is shown for clarity.

2 OPTICAL DIAGNOSTICS SYSTEMS

2.1 Input Sensor Package (ISP)

The input sensor package is located at the beginning of the laser chain and at the output of the preamplifier module (PAM), Figure 2. A charge coupled device (CCD) camera system is used to capture a nearfield or farfield image of the pulsed beam during a shot. A temporal sample of the beam is collected and launched into a fiberoptic cable to the power diagnostics system. The energy in the beamline is measured using an integrating sphere and collected by the energy diagnostic system. This energy diagnostic is calibrated against an insertable whole beam energy calorimeter.

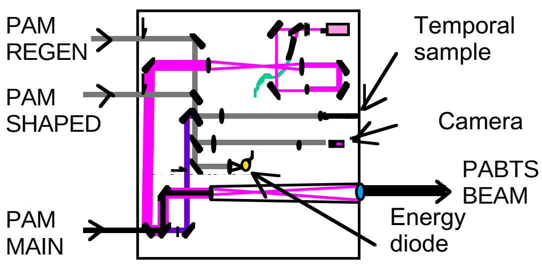


Figure 2: The ISP schematic shows how the PAM beams are sampled for alignment and diagnostics.

2.2 Output Sensor Package (OSP)

The output sensor is located below the transport spatial filter at the end of a set of relay optics as shown in Figure 3. This sensor provides diagnostic functions for the full beamline performance of a pair of beams. The diagnostic tasks performed by the OSP completely characterize the output of the main amplifier, measuring wavefront, and recording spatial and temporal pulse shapes. Multiplexing is utilized to record the performance at a reasonable cost. Two nearfield images are fit onto the 1ω CCD camera, so every beam is imaged every shot.

Energy measurement is accomplished for each beam in the relay optics, where a beam sample is directed onto a photodiode assembly. A temporal sample of one beam from every pair of 1ω beams is sent to the power diagnostic as described in the Input Sensor Package (ISP). The CCD cameras image the target plane by collecting light reflecting off of the final focus lens. Full-aperture roving calorimeters, which consist of an array of eight calorimeters that can move to intercept from 1 to 8 of the beams in a bundle, are located in the switchyard and complete the list of diagnostics. These calorimeters are used to calibrate the main amplifier photodiodes.

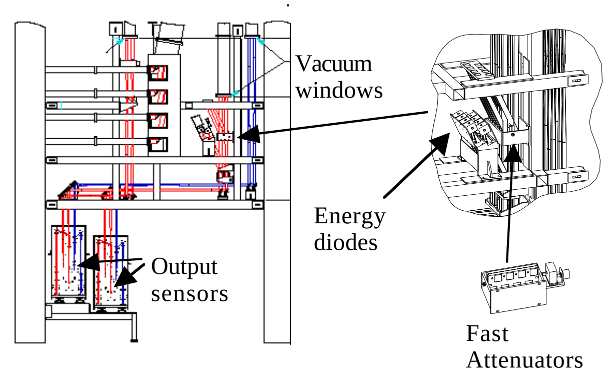


Figure 3: A model of the area under the transport spatial filter. The main beam samples are relayed to the output sensors. There is one energy diode for each beamline sample.

2.3 3ω Diagnostic

The 3ω diagnostics package is located at the target chamber following the frequency conversion crystals. The package measures the energy in the 3ω output beam with a calorimeter and the temporal waveform by directing a beam sample into a 3ω fiberoptic to the power diagnostics system as described in the ISP, see Figure 4.

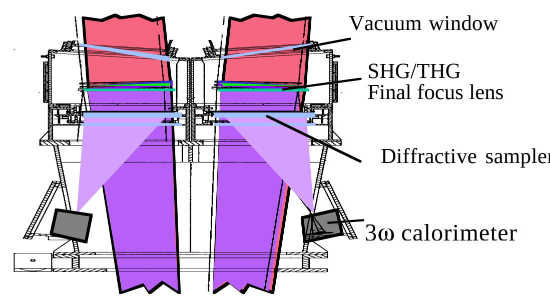


Figure 4: The absorbing glass calorimeter in the FOA gets a sample of the 3ω energy from an off-axis diffractive splitter.

3 DIAGNOSTICS MEASUREMENT SYSTEMS

3.1 Energy Diagnostics

The Energy Diagnostics system measures the laser pulsed energy to 2.8% accuracy using photodiodes and calorimeters at 432 locations distributed throughout the facility. The measurement electronics need to be accurate to better than 1.0% over a dynamic range of at least 10^3. An additional 24 calorimeters are used to calibrate the ISP and OSP photodiodes online, minimizing measurement errors induced by diagnostic component changes. These widely distributed sensors are interfaced using an Echelon™ LonTalk™ field bus-based embedded controller architecture to control the energy nodes located at the sensors. Centrally located VME-based PowerPC FEPs use a commercial off-the-shelf (COTS) VME-based Echelon™ LonTalk™ interface to communicate with the remote nodes. To simplify the design, the sensor power is also distributed on the network cable. The energy nodes are designed to minimize measurement errors and include maintenance functions to calibrate and detect energy node failures. Each LLNL-designed node has a signal source used to calibrate and test the node hardware in-situ. All calibration data, both system and hardware specific, is stored on the node in nonvolatile memory to ensure the calibration information remains with the node. Once the node is online and data is collected, the calibration information is used to calculate the absolute energy measured at this location.

3.2 Power Diagnostics

The temporal power measurement system (Power Diagnostics or PD) measures the pulse temporal shape to 2.8% accuracy using digital real-time oscilloscopes (digitizers) at 240 locations distributed throughout the facility. Meeting the measurement accuracy requires a digitizer with a dynamic range ≥3100, a record length of 22 ns, and a rise-time of ≤450 ps. In addition, an in-situ calibration system is used to periodically measure system performance for correction of the waveforms. The beam samples are collected from the ISP, OSP, and 3ω diagnostics and transported to centrally located Power Sensor Packages (PSP). Optically delaying each signal sequentially by 50 ns using 8-80 meter cables and converting them into electrical signals using a vacuum photodiode in the PSP multiplexes these signals. The electrical signals are input into a 2-to-1 transformer and recorded on a digitizer's four channels operated at different gains, allowing "overlaying" of the channels to extend the dynamic range of the measurement. These samples are then sequentially recorded on the digitizers to reduce the cost of the system. The 36 COTS digitizers are interfaced to the ICCS using 24 rack-mounted UltraSPARC processors and COTS Ethernet based GPIB interfaces. After the FEP reads the captured channel waveforms, they are corrected using digital signal processing (DSP) algorithms to correct amplitude distortion and normalize their bandwidth. Once the "raw" channel data is corrected, the four channel segments that correspond to a single "beamline" waveform are reconstructed using DSP algorithms to correct for the channel timing skew. This beamline data is then transferred from the FEP, and all beamline waveforms are analyzed offline with the ener-gy data to verify the NIF power balance for the shot.

3.3 Image Diagnostics

The measurement of the spatial uniformity of energy distribution within the beam is necessary to operate the facility consistently at 80% of the optical components damage threshold. This requires a 2% fluence (energy per unit area) measurement resolution at 1/125 of the beam spatial resolution. The principal means for monitoring the spatial profile of the beams is to capture images on CCD cameras located in the ISP and OSP described above. These images must then be digitized and processed through the ICCS Video FEP to determine the extent and locations of significant fluence modulation. The video FEP is a rack-mounted UltraSPARC processor and uses six four-channel COTS framegrabbers to capture the video images. All CCD cameras in the system are provided power and the synchronization signal by a Camera Interface Unit (CIU) located in proximity to the camera. The RS-170 video signal from the camera is returned to the CIU and digitized by a framegrabber within the FEP. The interconnection between a CIU and the FEP is either a coaxial cable or multi-mode fiber depending on the specific camera location within the NIF.

4 SUMMARY

The flow-down measurement accuracy for diagnosing the NIF created challenging design requirements for these beam diagnostics systems. Through a combination of COTS and LLNL-developed distributed systems combined with in-situ calibration and digital correction, we were able to accomplish these goals.

This work performed under the auspices of the U.S. DOE by LLNL under contract No. W-7405-Eng-48.

REFERENCES

[1] L. J. Lagin et al., "The Overview of the National Ignition Facility Distributed Computer Controls System," Lawrence Livermore National Laboratory, Livermore, CA, September 2001.

[2] O. Jones, B. Boyd, A. J. Gates, and S. Glenzer "Revised 3ω Measurement Error Budget," NIF-0053659, August 31, 2000.

FRAT003
physics/0111197

DISTRIBUTED CAN-BUS BASED BEAM DIAGNOSTIC SYSTEM FOR PULSE RACE-TRACK MICROTRON

F. Nedeoglo, O. Novojilov, S. Dudnikov, Department of Physics,
Moscow State University, 119899, Moscow, Russia
A. Chepurnov, I. Gribov, V. Shvedunov, Institute for Nuclear Physics,
Moscow State University, 119899, Moscow, Russia

Abstract

Very compact 70 MeV pulse race-track microtron is under construction now. To acquire outputs of beam-current transformers on every orbit and pulses of high voltage and RF field a distributed multi-channel beam diagnostic system was developed. Each acquisition controller consists of four fast differential amplifiers and one DSP-based micro-controller with on-chip ADC and CAN-bus controller. Each amplifier is coupled with beam-current transformer and has bandwidth of up to 150MHz and gain of up to 10. One of four channels is acquired during a measurement cycle. Another channel could be selected between two following pulses. All the controllers are connected via optically coupled CAN-bus with a host diskless PC running under Linux with the RTLinux extension. Dedicated software of the system consists of low level acquisition software for DSP, network software for controllers and host PC, application software for PC to present date for operator and control system. Standard CAN application layers were considered but refused because of the closed character of the whole system and centralised synchronisation of the whole system.

1 INTRODUCTION

First successful runs of the very compact pulse 70 MeV race-track microtron (RTM) have been provided and final tuning is carried out now. [1]. Parameters of the RTM are listed in the Table 1.

Table 1.

Injection energy	50 keV
Energy gain/orbit	5 MeV
Output energy	10-70 MeV
Number of orbits	14
Output current at 70 MeV	40 mA
Pulse length	~6-10 μs
Pulse repetition rate	150 Hz
Dimensions	2.2x1.8x0.9 mm
Weight	3200 kg

Because of the limited place between orbits, the original small size pulse beam current monitor (BCM) has been designed. The BCM is a passive wide-band current transformer with sensitivity up to 4,9 V/A and double-ended 50 Ohm-coupled output.

To measure amplitudes of the beam current in each orbit together with the amplitude of RF-field and high voltage pulse, a multi-channel distributed data acquisition beam diagnostic system has been created.

2 BEAM DIAGNOSTIC SYSTEM AS A PART OF CONTROL SYSTEM

The diagnostic system provides data necessary for control algorithms and human-machine interface (HMI) which are implemented in control system (CS) of the accelerator, therefore the system has been designed in such a way to be easily integrated with CS.

CS has a traditional three level structure [2]. X86-compatible computers are used. Front-end level consists of diskless PC with data acquisition boards. Middle level consists of diskless PC running under Linux together with real-time extension of the Linux - RTLinux. Linux is used to implement static and soft real-time algorithms whereas RTLinux is used to run hard real-time algorithms. HMI and the data bases are implemented in the third level. Ethernet over fibre optic is used to connect PCs in the accelerator hall with servers and HMI computers in the control room.

Beam diagnostic looks from top level of CS like one more dedicated acquisition subsystem but has different implementation architecture of front-end level.

3 STRUCTURE OF THE BEAM DIAGNOSTIC SYSTEM

The following technologies developed during the last few years have been used for the system: application of diskless PC running under Linux with real-time extension - RTLinux [3,4]; application of distributed stand-alone DSP-based smart controllers [5]; application of CAN-bus for accelerator control [6].

The output signal of BCM is measured by a stand-alone intelligent controller. Every controller has four inputs for the BCM. One of the four amplified signals could be digitised in a single acquisition cycle. The digitising process is synchronised by a dedicated pulse generated by the general synchronisation system of the

RTM. CAN-bus is used to connect controllers with the diskless x86-compatible host computer running under Linux together with real-time extension RT-Linux (Figure 1). BOOTP protocol is used to download the operating system to the host computer via Ethernet after switching power on. Host computer is equipped with an in-house designed CAN-bus adapter [5].

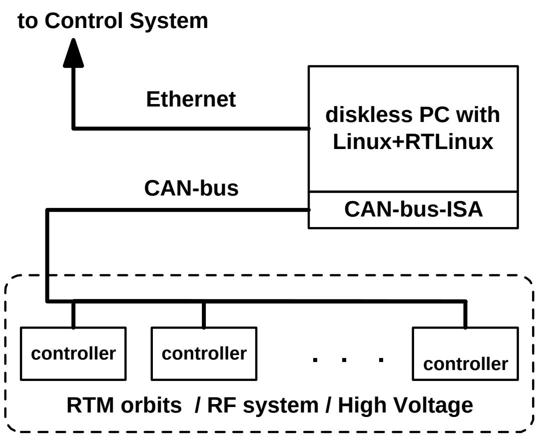


Figure 1. Structure of beam diagnostic system.

CAN-bus is a very popular fieldbus for accelerator control. Proceedings of ICALEPCS, PCAPAC and other conferences shows growing the popularity of CAN-bus for accelerator control with every next year. Maximum speed of CAN-bus is 1 Mbit/sec. But it is enough when the beam diagnostic system is used to measure values necessary for relatively slow static algorithms and HMI implemented in the high level of CS. The pulsed nature of the data allows transfer of the data in time gaps between two following pulses.

4 ACQUISITION CONTROLLER

The acquisition controller consists of analogue and digital parts (Figure 2). Four independent fast instrumental amplifiers (IA) are implemented in the analogue part. Every IA has unit gain bandwidth up to 200 MHz. Each of the IA could be separately enabled or disabled by the controlling DSP. All outputs of the IA are connected together to the inputs of two additional buffer amplifiers. They are used to couple the output of the IA with the ADC input and test analogue output simultaneously. The test analogue output allows us to use a digitising oscilloscope to measure and store the shape of the pulses in each orbit of the RTM.

The digital part consists of a digital signal processor (DSP) TMS320F241, an optically decoupled CAN-bus interface, an optically decoupled synchronisation input, a synchronisation and control schematic based on CPLD and an RS-232 interface. The DSP has an on-chip CAN-bus controller, fast ADC and other useful peripherals. The fast on chip ADC has 10-bits resolution and an 800 ns minimum conversion time.

Interrupt service mode of DSP operation allows us to utilise the high performance of DSP and ADC module.

The synchronisation pulse coupled with the beam pulse starts data acquisition process. One of the four channels is measured during one measuring cycle. The host computer sets the number of the channel to measure after the next synchronisation pulse comes. In addition, the host computer checks the state of the controller, defines the number of continuous measurements and initiates transmission of the results from measuring controller to the host via CAN-bus.

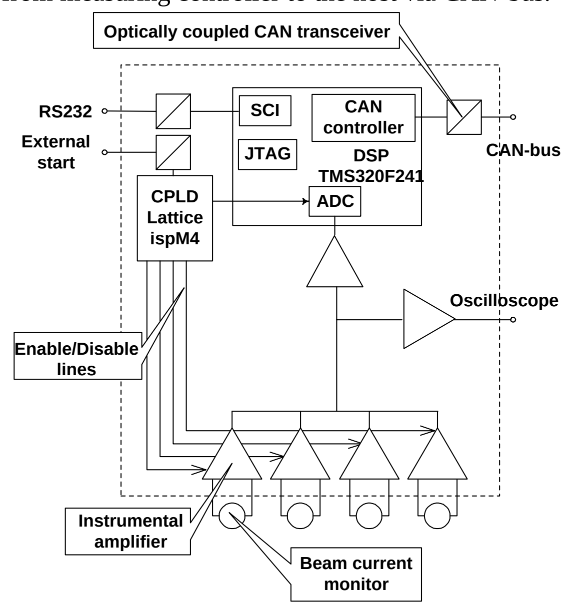


Figure 2. Structure of acquisition controller.

5 BEAM DIAGNOSTIC SOFTWARE

Standard CAN application layers such as CANopen and DeviceNet were considered as candidates for CAN application layers for the beam diagnostic system. Because of the following reasons the dedicated high level CANdiag protocol was created:

- the diagnostic system is closed to future extension, so a custom protocol is acceptable;
- DSP has limited size of on chip Flash-memory that is too small for standard protocols. Application of external memory is not reasonable;
- centralised synchronisation of the system and asymmetric flows of data makes application of standard protocols inconvenient.

The CANdiag protocol is based on a master-slave model of interaction. The master portion of protocol is implemented in host computer whereas all controllers are slaves.

Only the 11-bit CAN-identifier is used. Figure 3 represents usage of CAN-identifier bits.

10	6	2 0
Source MAC	Destination MAC	Message type

Fig. 3. Usage of CAN-identifier field in CANdiag protocol.

Each controller connected to the CAN-bus has its unique Media Access Control (MAC) identifier (ID), which identifies the device in the network and is used during the procedure of access to the bus. When the device sends a message to the bus, the first four bits of the CAN-identifier contain the MAC ID of sender (Source MAC ID). The next four bits contain the MAC ID of the device which expected to be receiver of this message (Destination MAC ID). The last three bits are used to identify the type of the message which define the semantic meaning of the message and format of data fields.

The CANdiag protocol supports simultaneous operation up to 15 devices in CAN-bus with addresses in the range between 0x00 and 0x0e. MAC ID 0x0f is used for broadcast messages to implement duplication MAC ID checking. Each device on CAN-bus working under CANdiag protocol starts its activity on the bus after switching power on with the duplication MAC ID checking procedure. The CAN-bus node sends broadcast message with Source MAC ID equal to Destination MAC ID and equal to 0x0f. Data field contains MAC ID of the node trying to connect to the bus. All nodes of the bus which are active in this moment receive this broadcast message and compare the MAC ID from data field with its own MAC ID. If the received MAC ID is equal to its own MAC ID, this node sends a broadcast reply which means that the requested MAC ID is occupied already and it means that the attempt to connect to the bus failed.

The CANdiag protocol supports the following types of messages defined by “Message type” field:

- configuration messages – are used to select dedicate measurement channel in slave device and reset controller remotely,
- status messages are used to check state of the controller;
- input/output messages are used to transfer stored digitised data.

Application software of the system was developed in ANSI C and consists of the low level software of the slave controller running on the DSP and the high level software of master.

Software for slave part of the CANdiag protocol was completely created, tested and debugged under Linux in emulation mode taking in to consideration features of the C –compiler for the DSP platform. Then pieces of code were ported very easily and quickly to the DSP.

Software of the master consists of a loadable module for RTLinux 3.0 and application software running under Linux on the same host computer. Application modules allow scanning of the CAN-bus to check state of all slave controllers, to provide cyclic polling of the controllers and so on.

A dedicated API is used between the host computer and the general CS of RTM to allow access to the beam diagnostic system from CS. The control program that is a part of the CS software uses two real-time FIFOs to communicate with the master’s software. One FIFO is used to transmit commands from the control program to the master’s program. The second FIFO is used to transfer results back from the master to the control program.

4 CONCLUSIONS

To simplify unification of the beam diagnostic system with the control system during start up and future operation the same architectural decisions should be used. Single platform of software development consisting of GNU C under Linux together with RTLinux was used. The platform was used to develop software for the CS, for the beam diagnostic system, as well as for high level, for embedded applications, and for real-time as well as for non real-time components. This approach is very convenient and could be recommended to develop control and beam diagnostic systems. One more application of CAN-bus for beam diagnostic systems is described. A disk-less PC running under Linux could be recommended as reliable and inexpensive solution for middle level of control systems.

REFERENCES

[1] V.I. Shvedunov, et. al., “70 MeV Electron racetrack microtron commissioning” Proc. of PAC, Chicago June 18-22, 2001

[2] I.V. Gribov., I.V. et. al., ‘‘RaceTrack Microtron Control System” Proc. of PAC, Chicago June 18-22, 2001

[3] A. Chepurnov, F. Nedeoglo, et. al., “Simple CAN-bus adapter for accelerator control running under Linux and RTLinux” CD-ROM Proceedings of PCAPAC’2000

[4] F. Nedeoglo, A. Chepurnov, D. Komissarov, “Linux and RT-Linux for accelerator control - pros and cons, application and positive experience” Proc. Of ICALEPCS’99, Trieste, Italy, ISBN: 88-87992-00-2, pp. 520-522.

[5] Chepurnov A.S., Dorokhin A.A., et. al., “Control System for Accelerator with distributed Intelligence Based on a "Family of Smart Devices", Proc. of the Vth Europ. Particle Accelerator Conference Sitges (Barcelona), Institute of Physics Publish. Bristol and Philadelfia, 1996, p.1794-1796

[6] A. Chepurnov, A. Alimov, et. al., “Control System for New Compact Electron Linac” Proc. Of ICALEPCS’99, Trieste, Italy, ISBN: 88-87992-00-2, pp.84-86.

FRAT004
physics/01111173

TIMING SYSTEM OF THE SWISS LIGHT SOURCE

T. Korhonen, M. Heiniger, Paul Scherrer Institute, Villigen, Switzerland

Abstract

The timing system of the Swiss Light Source provides synchronization for the accelerator and for the beamline experiments. The system is based on an event distribution system that broadcasts the timing information globally to all the components. The system is based on an earlier design [1] that came with extensive software support. We took the functionality from that design and implemented it on a new hardware platform. This paper describes the technical solution, the functionality of the system and some applications that are based on the event system.

1 INTRODUCTION

The minimum task of a timing system for a light source is to provide and distribute the reference pulses to control the injection and acceleration. However, from the beginning of design, it was clear that we wanted to have a system that would be more extensive.

First, good integration of the timing system into the control system allows the binding actions to timing signals and thus makes it possible to implement synchronous actions in a distributed system.

Secondly, it allows the ability to provide timestamps of collected data and performed actions. This makes the system a global timebase rather than just a pulse delivery network.

The possibility of extending the system beyond machine timing was one of the major objectives for us. For example, the requirement from early on was to provide the means for top-up operation. This implies that the timing has also to be available at the beamlines.

After surveying possible solutions, the APS event system seemed to provide the functionality that we were looking for. The only drawback was that the bandwidth was not sufficient for fast timing like injection control, but would have required additional hardware to implement the fast signals. However, the technology that would allow us to integrate all this into one subsystem was available. This prompted us to do a redesign of the APS system.

2 SYSTEM OVERVIEW

The injection system (linac and booster) of the SLS operates with a 320 ms (3.125 Hz) cycle. During one cycle the linac is triggered, the beam is injected to the booster, the magnets (and RF) are ramped to accelerate the beam and the beam is extracted from the booster and injected into the storage ring after which the magnets are ramped down before the start of the next cycle. The timing system (Figure 1) generates the synchronization reference signals and distributes these to the required components.

The reference generation is done with special purpose-built modules like a downconverter that divides the RF frequency to the fiducial signals of the machine. The event system is used to generate the sequence of events for the injection cycle and to distribute the timing to the relevant components. Practically the whole operation cycle is generated within the event generator from only one fiducial, namely one that tells when the number “0” buckets in the storage ring and booster are in alignment. The SLS injection sequence consists of about 20 events that are transmitted every 320 ms.

The event system is based on time-multiplexed transmission of event codes. The source is an event generator that transmits 8-bit event codes at a frequency of 50 MHz over an optical link. Thus, the time resolution is 20 ns. The bitstream is multiplied to several branches with fanout modules. At each destination VME crate there is an event receiver that decodes the received event codes and performs the appropriate actions that are programmed for each event of interest to the particular crate.

Some subsystems like the linac have their own internal timing. During construction the linac was required to have the ability to run standalone. The interface between the linac and the rest of the machine was defined to be the timing and RF signals. The synchronising timing to the linac is delivered with the event system; the linac cycle is triggered from an event.

The linac timing consists of the gun trigger system and the timing for the klystron modulators, prebuncher, gun RF amplifier and the 3 GHz frequency multiplier. This timing is achieved by using a pair of Stanford Research DG535 delay generators. The gun timing uses a number of TD4V delay cards [2], which also phase lock the trigger signals to the RF signal. The TD4V delay is programmable in units of RF cycles (2 ns) and is used to target specific RF buckets in the booster (and storage ring) and also to adjust the trigger pulse length. The trigger signal from these cards to the electron gun high voltage deck is delivered through an optical fiber using an Uniphase 51TA/RA transmitter/receiver pair.

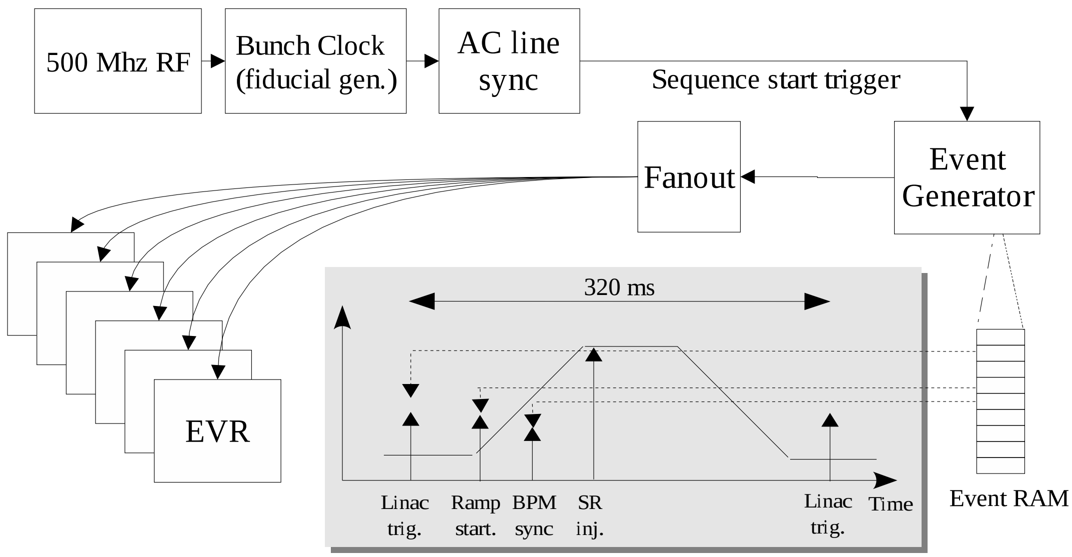


Figure 1: Structure of the timing system.

3 TECHNICAL SOLUTION

To integrate all the desired functionality into one system, we decided to take the APS system as a model, but to upgrade to a recent technology: the Gigabit Ethernet. The requirement was to preserve the software compatibility so that we could take advantage of the software that was existing for the APS system.

The functional design was done in VHDL. This brought a big advantage: we could separate the hardware solution from the software interface, in a way that allows us a smooth upgrade path for both of the hardware and the software; when the communication technology has made an significant advance, we can port our model to the new technology without being tied to any particular technology or a component. And vice versa: any new ideas from the control system side can be implemented in firmware without any physical modifications (within the limits of the capabilities of the hardware.)

3.1 The hardware

The event system is based on Gigabit Ethernet technology, using a standard VCSEL short wavelength (860 nm) transceiver (drop-in replacement for long wavelength transceivers exist) and a gigabit transceiver chip. The transceivers allow for cascading several cards in a daisy chain. This way we can reduce the number of fanout branches and also build a sub-branch where another event generator is added to an upstream branch.

The logic resides in two PLDs, one CPLD and one FPGA (Xilinx Virtex). The CPLD handles the VME bus logic and configuration of the second one from Flash ROM. Flash ROM allows us to upgrade the systems in situ, even without powering down the system. This was a big advantage during development.

The event generator and receiver use the same PCB, with some different components fitted in depending on the card type. The event generator has, in addition to the common hardware, two 512 KB RAMs for "store and playback" of an event sequence, external clock and external trigger inputs. The event receiver has three extra outputs in the front panel. Their use is defined in firmware; presently they are used for reference frequency output.

The main clock for the event generator is downconverted from the main (500 MHz) RF. The downstream receivers synchronize automatically to the incoming bitstream. All the components are thus phase locked to the RF.

3.2 The firmware

The firmware for the cards is about 2000 lines of VHDL code for both of the modules (EVG, EVR). Being VHDL, it is independent of the particular type of chip (FPGA) used in the hardware implementation.

The main features were already present in the APS version; we extended them to provide more functionality to suit our operation scheme.

The most essential features of the event generator firmware are event RAM sequence handling, event priority resolver and the possibility to send events from software by writing into a register. The EVR has 4 channels with width and delay possibility, 14 channels with width adjustment only. Both of these have also a clock prescaler to scale down the clock and polarity selection. The EVR also has timestamp counter with the facilities for a synchronous reset and event FIFO mechanism for latching the timestamps when an event is received. For details, see [4].

The firmware for the event receiver uses about 95 % of the available gates on the Xilinx Virtex 150.

3.3 Interface to devices

Different devices require different signal types for the trigger pulses. Most devices have a TTL level input, some may have an NIM or ECL level input. Rather than trying to cover every possible requirement on-board, we decided to put the interface logic to an transition module according to the same philosophy as for other SLS controls interfaces. As an example, the power supplies for SLS [3] have an optical trigger input to run an internal waveform. For the power supplies, a transition module with optical outputs was designed.

3.4 Distribution system

The signals from the main timing source are distributed using an optical fiber fanout tree. The fanout is done with arrays of VCSEL transmitters. These are built as VME size cards that plug in the crate, but have no data interface. One card has one input and eight outputs, limited by the card size. To create more outputs, one can connect the cards in a tree configuration.

4 APPLICATIONS

4.1 Filling control

The timing system must be able to precisely control the filling of the storage ring. With the event system we could create a simple application to control the filling, without the need for any additional hardware and a few parameters to handle. The harmonic numbers of the SLS booster and storage ring are 450 and 480, respectively. This means that for each successive turn in the booster, the storage ring lags 30 RF buckets, until after 16 booster turns the same buckets are aligned again. This means that by shifting the extraction delay from the booster we can select the target bucket in 30 bucket steps. For the smaller steps we need to adjust the linac gun timing. Thus all that is needed to control the injection are two parameters: extraction delay and the linac gun delay. With the proper combination of these, all buckets in the storage ring can be reached.

The injector timing sequence is generated with the event RAM. The RAM clock is selected to be exactly one booster turn (450/500 MHz = 900 ns) and the cycle is started from a pulse that signals that the SR and booster RF buckets are in alignment. Changing the position of the extraction event in the RAM by one shifts the extraction by one turn, giving us the 30 cycle steps. The smaller steps are done by adjusting the linac gun timing. The linac timing controller is synchronized with the main timing. Knowing the next target bucket it calculate the required delay for the next cycle. The calculation and setting is synchronised with the injection with the corresponding event: the event triggers a software sequence to be executed on every injection cycle.

4.2 Top-up

Ability to do top-up injection was one of the requirements for the SLS from the beginning. This also sets several requirements for the timing system. The timing system has to be able to do single injection cycles, has to have the capability of sending gate signals to the beamline experiments and obviously, to synchronize several controllers that otherwise are interconnected only through Ethernet. The network connection cannot be relied on to transmit the command sequences in applications that are related to injection in real time. Basically, top-up is just an extension of the normal injection control application; it has to monitor the beam current, allow for setting of proper injection intervals and to have a capability to send advance notice to the beamlines before an injection.

4.3 Beamline timing

With our structure, we can simply extend the timing to beamlines by extending one fiber branch to an interested beamline, making all the timing signals available at the beamline. If the beamline has some special requirements, we can add one event generator to create a new sub-branch to create a tree-like structure. In this way, the beamline can add its own events but still have the whole machine timing information available.

5 CONCLUSIONS

The advances in technology have made it possible to build a timing system that can handle most of the timing tasks from commercial off-the-shelf components that are standardized and in wide use. The SLS timing system integrates most of the required timing functions into a single subsystem. The system is layered so that we could benefit from other earlier development but at the same time move to a one generation newer technology. The functionality is captured into a device independent language (VHDL), which allows decoupling of the functionality from the underlying technology, providing a smooth upgrade path for the next generations. The integration of the functions into a global system enables us to glue a distributed system into a single entity. This can greatly enhance the capabilities and operation of a complex system.

6 REFERENCES

[1] F. Lenkszus, R. Laird. “The APS Event System”, ICALEPCS’95, Chigago, June 1995.

[2] T. Naito et.al. “Synchronization scheme for the KEKB”, ICALEPCS’97, November 1997.

[3] A.Luedeke et.al.,”Digital Power supplies for the Swiss Light Source”, these proceedings

[4] J. Pietarinen, “Event System”, available on-line at http://www.sls.psi.ch/controls/help/tutorials/subsystems/timing/

PERFORMANCE OF THE REFERENCE AND TIMING SYSTEMS AT SPRING-8

Y. Ohashi, T. Asaka, H. Ego, M. Hara, Y. Kawashima, T. Ohshima, T. Takashima and H. Yonehara, SPring-8, Hyogo 679-5198, Japan

Abstract

Reference RF and timing systems were developed at SPring-8 in order to handle beams in a stable and flexible manner under various conditions of machine operation. A synchronous universal counter plays an essential role in the timing system. Here, the method to realize high-precision timing is described, as well as the tools that were required to confirm that the system operates as designed. Furthermore, a new scheme for the synchronization of the different RF signals from linear and circular accelerators is described.

1 INTRODUCTION

The SPring-8 accelerator complex consists of an 8 GeV storage ring (SR), a booster synchrotron (SY), and a 1 GeV linac. A common master oscillator provides the 508.58 MHz RF reference signal to the SR and SY, whose harmonic numbers are 2436 and 672, respectively. The linac is operated at 2856 MHz and a repetition rate of 60 Hz. The SR is a third generation light source that provides high-brilliance synchrotron light to various experimental groups. Any kind of beam filling and refilling patterns are achievable by the successful development of a synchronous universal counter (SUC) [1] and a stable signal-transmission system.

The 508 MHz SUC must perform the following functions; 1) trace the location of the RF buckets by synchronous non-stop revolution counting, 2) act as a timing/delay pulse generator with synchronization of 508.58MHz. A SUC is widely used by the users who require precise timing relative to the beam. A phase-stabilized optic fiber (PSOF) [2] is commercially available and a temperature coefficient in the phase drift of about two orders of magnitude smaller compared to a phase-stabilized coaxial cable. In addition to that, electric-to-optic (E/O) and optic-to-electric (O/E) signal transmitters with small temperature coefficients became commercially available also. An optic signal-transmission system has the advantage over the conventional electric one with respect to phase stability. In combination with the synchronization feature of the SUC mentioned above, a high-precision timing-signal transmission system, for which time jitters are minimum and almost independent of the transmission distance and delay period, is established.

A development of reference RF and timing system was started in 1991, and its feasibility was confirmed by the end of 1993. Construction of the timing system was completed in 1996, and the system is running properly and stably since beam commissioning in March 1997.

A single-bunch beam is created from the linac beam of 1 ns width by RF knockout system installed in the SY [3]. Acceleration of a beam in a single RF bucket of 2856 MHz is desirable in order to make the single-bunch purity better. There is no simple relation between two RF frequencies of 508.58 and 2856 MHz, however, two RF signals had to be synchronized for precise beam injection. We paid attention to the fact that the linac is not operated in CW. A new scheme of synchronizing multiple RF frequencies has been implemented [4].

2 SYNCHRONOUS UNIVERSAL COUNTER

Recent digital technology made it possible to perform a direct counting at a rate of over 500 MHz for accelerator applications. The 508 MHz SUC has a capability to count up to 30 bits in order to cover 1 sec of ramping time of the SY. The circuit block diagram is shown in Fig. 1. We describe some details of the SUC used at SPring-8. Application to a different environment is simple. The SUC counts the RF bucket number corresponding to the harmonic number (N=2435) of the SR, and it then provides a pulse corresponding to the revolution

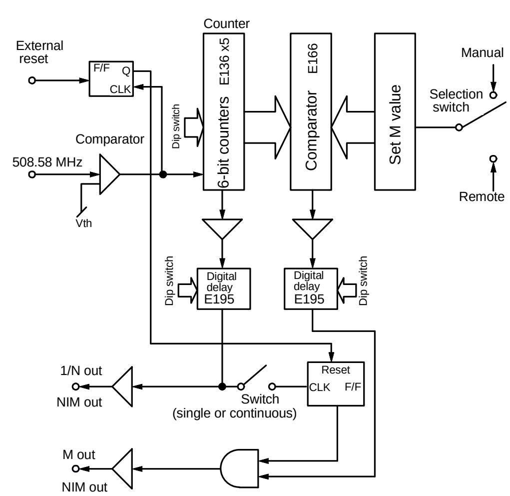


Fig. 1. Block diagram of a 508 MHz SUC.

frequency of SR from '1/N out'. The output 'M out' corresponds to the particular RF bucket within a revolution. Beams are injected into the targeted RF buckets from the SY by setting the number 'M' as desired. 'M' can be set either externally or manually. By resetting the SUC, a delayed pulse is generated from '1/N out' ('M out') with a delay time of N(M)/508.58MHz in a synchronous way. The device is packaged in a single width standard NIM module, and it can be applied whenever precise timing is required.

3 REFERENCE AND TIMING SIGNAL DISTRIBUTIONS

An optic signal-transmission system was adopted where phase/timing precision is crucial. Proper methods to distribute the signals were selected in consideration of the environmental factors. There are two ways to distribute the reference RF signal to the four RF control stations; 1) sequential transmission from station to station relying on the phase-locked-loop (PLL) feedback system between two stations, 2) single loop with optic directional couplers around the SR with a single PLL feedback between transmitted and returned signals. The SR-building is kept at a steady temperature of 25±1°C by air conditioning to meet with the strict experimental requirements from synchrotron-light-source users. Because of the excellent phase stability of the PSOF, we selected the latter method (Fig. 2), also because of the fact that the sequential use of PLL feedback piles up phase errors. On the other hand, we introduced a different method for the signal distribution to the injectors, since temperature along the fiber path of ~700 m is not controlled as well as in the SR case. A PLL feedback, using the signal returned in the

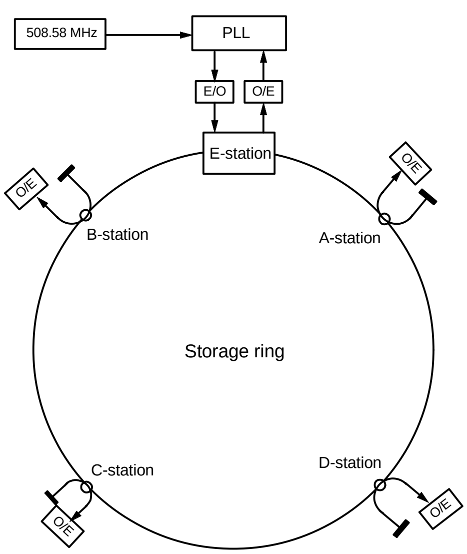


Fig. 2. Reference RF signal distribution for the storage ring.

same optic fiber reflected by the mirror located at the end point was applied [5].

Devices used for the timing-signal distribution are similar to the ones used for the reference RF signal. Time jitters due to the unstable output pulse amplitude from the E/O module are compensated by inserting a constant fraction discriminator (CFD). A more precise timing signal is delivered by synchronizing it with a reference RF signal in the place where precision is required. This is simply realized by introducing a signal concerned into the input 'External reset' of the SUC. Timing precision is restored to a few ps. Annual drift in the reference RF phase is about 5 degrees in the SR and 3 degrees in the SY without PLL feedback control. They become within 1 degree with PLL feedback control. Time jitters in the timing pulse transmission without synchronization vary depending on the system. For example, it was measured to be 7.4 ps (standard deviation) in the electron gun system of the linac [1].

4 PERFORMANCES

There are several ways to check the stability of the reference RF and timing system.

1) A picosecond pulsed laser used in a time-resolved experiment must inject laser beam in a sample exactly when synchrotron light beam hits the sample. Timing between the two beams was measured using a streak camera. A stability of ±2 ps was obtained for a few hours [6].
2) Impurity level of a single-bunch beam was measured in the SR and was less than 10^{-9} (below detection level) over several hours [7]. Reduced beam current during user experiment period is restored by the beam refills.
3) The beam orbit deviates due to the tidal forces. The beam energy shifts due to changes in the circumference of the SR caused by the deformation of the ground. The beam orbit correction system compensates such deviations, and it results in a shift of the RF frequency. Maximum peak-to-peak variation of about 70 μm in the circumference is properly corrected, where the correction step is currently set to 0.3 μm [8].

5 NEW SYNCHRONIZATION SCHEME BETWEEN TWO RF'S

A simple way to synchronize two RF frequencies is to generate those RF frequencies from a common sub-harmonic oscillator together with frequency dividers/ multipliers. Although this method should in principle be applicable, a compromise between hardware availability and performance has to be found. Since a conventional method was not applicable to the SPring-8, a new idea was sought by freeing the constraint between two RF's required for synchronization. A 2856 MHz RF is actively used for a short period that corresponds to a small duty factor of the linac. When the beam injection takes place, a pulse synchronized with 508.58 MHz triggers the 2856 MHz generation each time. Fig. 3 shows the block diagram of the new synchronization method [4]. The 2856 MHz RF generator is simple and consists of an arbitrary waveform generator (AWG), bandpass filters, and a frequency multiplier. The numbers in Fig. 3 indicate the current parameter setting. The duration of 2856 MHz generation is 290 μs. The spectra from the synthesizer and the present device are shown in Fig. 4. The difference of 18.7 kHz is small enough and does not affect practical operations. Suppression of time jitters in the 'start signal' in Fig.3 is essential to obtain a high precision synchronization of the two RF signals, and it is realized by using the SUC. Only one master synthesizer of 508.58 MHz is present after elimination of the one in the linac. The method described here can be applied widely and easily to similar facilities. One can choose the best RF components suited for the project without caring about their frequencies any more.

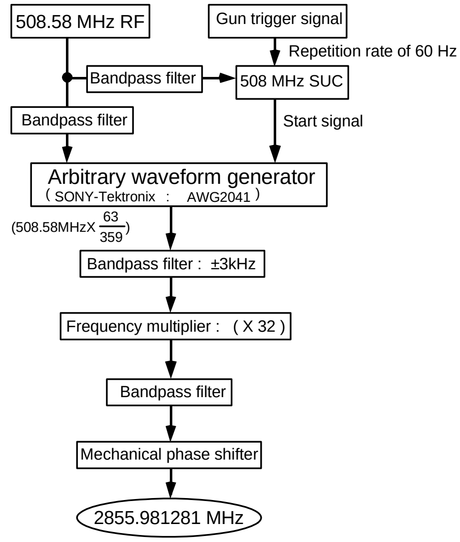


Fig. 3. Block diagram of the new synchronization method between 508.58 and 2856 MHz RF's. The system consists of an arbitrary waveform generator and a frequency multiplier. The 2856 MHz RF is generated only for a short duration of about 290 μs.

6 CONCLUSIONS

The SUC is a key tool in the SPring-8 timing system for flexible beam handling. It gives the revolution and RF bucket information with small time jitters. The reference RF signal distribution system was successfully constructed using PSOF, E/O and O/E with small temperature coefficient and time jitters. Annual phase drift is within 1 degree by a PLL feedback. A timing signal distribution system, which requires extremely small time jitters, was constructed by using the synchronization function of the SUC, and it is thus less sensitive to the transmission distance or delay. A new synchronization system for two different RF's performed well. We now have only one synthesizer as a master in the SPring-8 accelerator complex. Application of the tools and systems described in this paper is simple to other facilities.

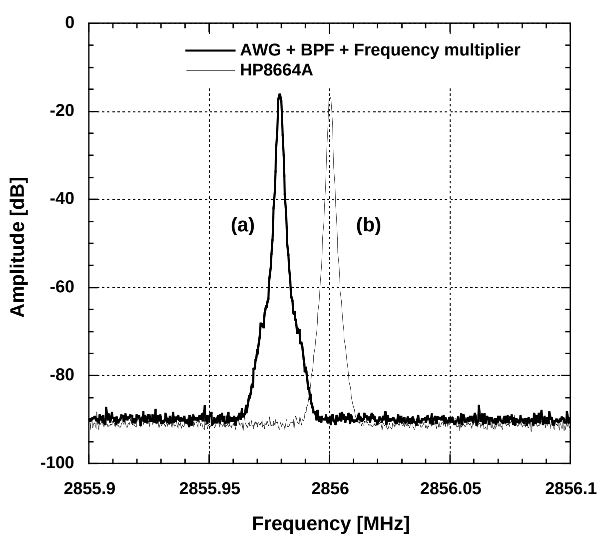


Figure 4: Spectra of 2856 MHz RF's: (a) generated by the new method; (b) generated by a synthesizer.

REFERENCES

[1] H. Suzuki *et al.*, Nucl. Ins. Meth. A 431, 294-305 (1999).
[2] J. Urakawa, Particle Accelerators, 1990, Vol.29, pp.251-256.
[3] H. Suzuki *et al.*, Nucl. Ins. Meth. A 434, 515-533 (2000).
[4] Y. Kawashima *et al.*, Phys. Rev. ST Accel. Beams Vol.4, 082001, 1-5 (2001).
[5] E. Peschardt and J.P.H. Sladen, IEEE, 1989, pp. 1960-1962.
[6] Y. Tanaka *et al.*, Rev. Sci. Ins. 71, 1268-1274 (2000).
[7] K. Tamura, Proc. of the 12th symposium on Accel. Sci. & Tech., Wako, Japan, 1999, pp. 546-548.
[8] S. Daté and N. Kumagai, Nucl. Ins. Meth. A 421, 417-422 (1999).

FRIDAY, 30 NOVEMBER 2001

FRB - OPERATION, COMMISSIONING, AND PROCESS TUNING, REMOTE OPERATION AND PARTICIPATION

FRBI001
cs.OH/0110029

HOW TO COMMISSION, OPERATE AND MAINTAIN A LARGE FUTURE ACCELERATOR COMPLEX FROM FAR REMOTE SITES

P. Czarapata FNAL, D. Hartill, Cornell; S. Myers, CERN; S. Peggs, BNL; N. Phinney, SLAC; N. Toge, KEK; F. Willeke, DESY; C. Zhang , IHEP Beijing

Abstract

A study on future large accelerators [1] has considered a facility, which is designed, built and operated by a worldwide collaboration of equal partner institutions, and which is remote from most of these institutions. All operation modes were considered including trouble shooting, development, commissioning, maintenance, and repair. Experience from existing accelerators shows that most of these activities are already performed 'remotely'. The large high-energy physics and astronomy experiments already involve international collaborations of distant institutions. Based on this experience, the prospects for a machine operated remotely from far sites are encouraging. Experts from each laboratory would remain at their home institution but continue to participate in the operation of the machine after construction. Experts are required to be on site only during initial commissioning and for particularly difficult problems. Repairs require an on-site non-expert maintenance crew. Most of the interventions can be made without an expert and many of the rest resolved with remote assistance. There appears to be no technical obstacle to controlling an accelerator from a distance. The major challenge is to solve the complex management and communication problems.

1 INTRODUCTION

The next generation of particle accelerators may require a new mode of international and inter-laboratory collaboration since they are too costly to be funded by a single nation and too large to be built by a single laboratory. The tremendous technical challenge of a new facility requires a critical mass of highly qualified and experienced physicists and engineers. These experts are presently distributed among the accelerator centers around the world and it is believed important to maintain and develop this broad base of expertise. The successful recent accelerator technology development depended on extensive exchange of people with complementary technical skills. Therefore, it is desirable that several accelerator laboratories will participate in any future project. A consequence of a multi-laboratory project is that the facility will be located a considerable distance from most of the contributing institutions which design and build it. Shared remote operation is a model that allows the experts who designed and built the machine to continue to participate in its operation. In order to make such a model work, the collaborating institutions must have a continuing commitment to the project. We discuss below a model for an international multi-laboratory collaboration to construct and operate an accelerator facility, which attempts to meet this requirement. The issues for far-remote operation are based on this model. The following questions are addressed: What is required for effective exchange of experience, ideas, parameters and data necessary for adequate discussion of the problems expected during commissioning, tune-up, failure analysis, and perfor-mance and reliability improvements? What local staff is required for operations, maintenance, and repair? What needs to be changed in the technical design of the hardware components to allow remote diagnosis and analysis? Are the costs of these changes significant? What are the requirements on the control system data transmission speed or bandwidth to support remote operation? Are presently available technologies a limitation requiring further research and development?

2 AVAILABLE EXPERIENCE

Existing large accelerators such as LEP and HERA are remotely operated where the controls architecture supports 'far-remote' control. Feedback loops that require fast response are implemented locally and do not require continuous intervention from the main control room. Analog signals are almost always digitized before transmission to the control room so that there is no loss of information through long cable runs. The enormous advances in computing and networking have made digital media the most convenient and inexpensive method for transmitting data, even over short distances. The large size of present accelerators and the limited access demands that interventions be well-planned and undertaken only after extensive remote diagnostics. Non-expert maintenance staff is able to handle most of failures and repairs. In difficult cases, they are assisted by experts via telephone or via remote computer access to the components. Unscheduled presence of experts on site is exceptional. Detailed reports and analysis from LEP and HERA that support these conclusions are available [1]. The commissioning, operation and optimization of the

SLC is perhaps the most relevant experience for a future linear collider. However, because the SLC was an upgrade of the existing SLAC linac, many of the technical components were not modern enough to support remote maintenance and troubleshooting. A significant presence of expert staff on site was required. The control system was designed to allow consoles to be run remotely from home or office. With proper coordination, they could be run from other laboratories. However, operators in the SLC control center relied on many analog signals from older diagnostics, which were not available remotely. Although extensive feedback systems were developed for the SLC to stabilize the beam parameters and optimize luminosity, some tasks still required frequent operator. This experience might seem discouraging for the feasibility of far-remote operations, but none of these technical limitations are fundamental given modern technology. The increase in SLC performance was often enabled by a good data logging systems and off line analysis. Such analysis could have been performed from anywhere in the world if the data were available to an external group. In fact, many aspects of SLC experience with feedback, automated procedures and complex analysis are encouraging for far-remote operation.

Many of the initial difficulties in commissioning accelerators are caused by insufficient diagnostics. Whenever comprehensive controls and diagnostics have been available in the control room at an early stage of accelerator commissioning, they have facilitated a rather smooth and quick turn-on, as seen at ESRF and PEP II. There are many more examples of facilities with insufficient initial diagnostics where progress was unsatisfactory. Any large accelerator must have remote troubleshooting capability, simply because of the distances involved. The conclusion is that any facility with adequate diagnostics and controls for efficient operation could also easily be commissioned remotely.

Non-accelerator projects also have extensive experience with remote operation of complex technical systems with restricted accessibility. The successful operation of space experiments and of distant telescopes demonstrates that efficient remote operation is possible, practicable and routinely performed. In particular, many observatories are built in rather inhospitable locations and operated with only very little technical support on site. Troubleshooting and consultation with experts is almost exclusively performed remotely. The European space agency ESO has remotely operated telescopes in Chile from a control center in Germany for a decade. Their operational experience [2] is encouraging but demonstrates that the main difficulties lie in handling exceptional situations. These institutions maintain a strong presence of experts on site despite the unfavorable conditions in order to mitigate these problems. The collaborators on a remote accelerator project should carefully analyze and learn from the ESO experience.

3 MODEL OF REMOTE OPERATION

3.1 General Organization

The accelerator is built and operated by a consortium of institutes, laboratories or groups of laboratories. Each collaborator is responsible for a complete section of the machine including all subsystems. This responsibility includes design, construction, testing, commissioning, participation in operations, planning and execution of machine development, maintenance, diagnosis and repair of faulty components. These machine sections are large contiguous parts of the machine such as for example injectors, damping rings, main linacs, beam delivery, for a linear collider. This eases the design, construction, and operation of the accelerator, if the responsibility for large systems is assumed by a group of institutions from one region. To minimize the variety of hardware types to be operated and maintained, collaborators are also responsible for a particular category of hardware spanning several geographic regions.

Central management is needed to coordinate the design and construction of the machine and later supervise operation and maintenance. Responsibilities include the overall layout of the accelerator with a consistent set of parameters and performance goals. This group ensures that all components of the accelerator fit together and comply with the requirements for high performance and efficient operation (definition of naming conventions, hardware standards, reliability requirements, quality control, control system standards and interfaces as well as of the real-time and off line database). The central management coordinates the construction schedule and has responsibility for the common infrastructure (roads, buildings and tunnels, power, water distribution, heating and air conditioning, cryogenics, miscellaneous supplies, site-wide communications and networks, radiation and general safety).

Central management also plans and coordinates the commissioning, as well as supervision and training of local maintenance crews. A central operation board (in coordination with the experiments) would be responsible for the mode of operation, operational parameters, machine study periods, and interventions, planning of maintenance periods, organization of machine operation, and training of the operations crews. All remote operation crews would report to the central operation board. Regardless of where the active control center is located, high performance operation of the accelerator will depend on a continuous flow of information and input from all of the collaborators. They must maintain responsibility for the performance of their components for the entire operational lifetime of the machine.

3.2 Machine Operation

In the multi-laboratory model, there are several fully functioning control centers capable of operating the entire accelerator complex, one at the site and one at each of the collaborating institutions. The operations crew is decentralized and can operate the accelerator from a far-remote center. At any given time, however, the machine is operated from only one of these control centers. The current control center has responsibility for all aspects of accelerator operation including commissioning, routine operation for physics, machine development studies, ongoing diagnosis, and coordination of maintenance, repairs and interventions. Control is handed off between centers at whatever intervals are found to be operationally effective. Supporting activities may take place at the other locations if authorized by the active control center.

3.3 Maintenance

The collaborators remain responsible for the components they have built. They must provide an on-call service for remote troubleshooting. The current operations crew works with the appropriate experts at their home institutions to diagnose problems. It has the authority to respond to failures requiring immediate attention. An on-site crew is responsible for exchanging and handling failed components. Their responsibilities include putting components safely out of operation, small repairs, disassembling a faulty component or module and replacing it by a spare, assisting the remote engineer with diagnosis, shipment of failed components to the responsible institution for repair, maintenance of a spares inventory, putting the component back into operation and releasing the component for remotely controlled turn-on and setup procedures. Some tasks such as vacuum system interventions or klystron replacement will require specialized maintenance staff which must be available on site to provide rapid response time. Decisions about planned interventions must be made by the operations board in close collaboration with the laboratory responsible for the particular part of the machine.

3.4 Radiation and other Safety Issues

Any accelerator can produce ionizing radiation. Its operation must be under strict control. In addition to the laws and requirements of the host country and its overseeing government agencies, there are also the internal rules of the partner laboratories. Usually it is required that on-site staff be supervising beam operation to guarantee responsibility and accountability and permit safe access to the accelerator housing. There are also concerns about the activation of high-power electrical devices and other potentially hazardous systems requiring interlocks and tight control. We believe that there exist straightforward technical solutions to ensure the required safety and security. The legal and regulatory issues are more difficult and will need careful investigation. Most likely, a near-by laboratory will have to assume responsibility for radiation and other safety issues. Similarly, unusual events like fires, floods, accidents, breakdown of vital supplies or general catastrophes will require a local crisis management team available on call to provide an effective on-site response. There must be a formal procedure to transfer responsibility to the local crisis management in such instances. This function could be provided by nearby collaborating institutions.

4 REQUIREMENTS FOR FAR-REMOTE OPERATION

4.1 Organizational Requirements

Operation via a rotating control center requires good documentation and a mechanism to assure continuity when the operations are handed over to another laboratory. Electronic logbooks will be necessary, including a comprehensive log of all commands and events with time stamps, information on the originator and comments, with a powerful intelligent browser to make the logged information useable for analysis. All control rooms should be close to identical with a no local dialect and specialization. A comprehensive system of tokens or permissions to coordinate between the different control centers is needed. These are granted by the operators in charge. The tokens must have sufficient granularity to allow active remote access to a specific component and to regulate the level of intervention. Some organizational measures will have to be taken to avoid misunderstanding and loss of operation time. This includes a more formal use of language with strictly and unambiguously defined elements. Formal names are required for accelerator sections, lattice elements, technical components and even buildings. This means that a comprehensive dictionary has to be written and maintained.

In order to keep all collaborators well informed and involved and in order to maintain active participation of distant institutions, a special effort should be made to make the control center activities visible and transparent for distant observers, in particular the current status of machine operations. Monitors should be available to follow the operations progress and discussions in the active control center. They should easily allow regular 'visits' to the control center. All operations meetings (shift change, ad hoc meetings for troubleshooting, operation summaries, coordination with experiments, etc.) should be open to remote collaborators. Virtual 'face-to-face' communications can support multi-party conversations, including shared 'blackboards' and computer windows, perhaps using virtual 'rooms' to accommodate specialists with common interests. A permanent videoconference type of meeting may serve as a model. We expect that growing commercial interest in this sector will promote the needed development.

To avoid having the accelerator teams cluster in local groups with limited exchange of information, one should prepare for regular exchange of personnel, for plenary collaboration meetings, for common operator training across the collaboration and for regular exchange and rotation of leadership roles in the project.

A relatively modest staff appears to be required on site for operations, maintenance or repair. Most of the activities of operation, troubleshooting and failure diagnosis can be performed remotely by off-site personnel. Extrapolating from the experience of existing facilities, expert intervention on site is required in only about 1% of the failures. If one assumes a rate of 2000 incidents per year, there should be not more than 20 occasions where expert help has to be available on site even without relying on further improvements in remote diagnostics, increased modularity and more maintenance friendly future designs. Extrapolating from HERA experience, the size of the on-site maintenance crew is estimated to be about 75 persons for a large future facility. For efficient operation of the accelerator, regular maintenance is required in addition to the repair of failed components. This work must also be performed by on-site staff that is supported by a nearby laboratory or by industrial contractors. The collaborator responsible for the components would plan and manage these efforts under the coordination of the operation board. A small coordination team of about ten would be needed to provide the necessary micromanagement. In addition, there must be staff on site for security, for radiation safety, and for maintenance of infrastructure, buildings and roads. The number of persons needed depends very much on the specific circumstances of the site and the type of accelerator and it is hard to predict a number. In a large laboratory, the staff for these tasks is typically 50-100. In conclusion, we estimate that a local staff of about 200 would be required to maintain the facility and assure operations.

4.2 Technical Requirements

The control system must optimize the flow of information from the hardware to the operations consoles to provide remote accessibility of the hardware from remote sites without excessive data rates. The layered approach of modern control systems comfortably supports these requirements.

The console applications at the control centers would essentially only build up displays and relay operator commands. These activities require a slower data rate commensurate with human response times, which should not be a problem over any distance on earth. The requirements for console support are well within the reach of existing technology. The most significant bandwidth demand is for real-time signals, which are used for continuous monitoring by the operations staff. Most of the existing accelerator control systems use Ethernet LAN technology for data communications at the console level. In present facilities, 10Mbit/sec Ethernet technology is sufficient to accommodate the required data rate with an overhead of a factor of ten. The technology for ten times this bandwidth is already available and further development can be anticipated. This should be more than adequate for any future console communication requirements.

The intercontinental data connections have been revolutionized by the recent progress in fiber optics systems providing data rates in the multi-Tbit/sec range, or nearly inexhaustible capabilities. Future needs for data communications at the particle laboratories are in the range of several Gbit/sec [3]. They are driven by the exchange of experimental data. The need for remote accelerator control is in the order of a few 10Mbit/sec which doesn’t constitute a significant fraction of the anticipated connectivity. Thus the network is not expected to impose any limitation to remote operations.

High performance accelerators rely extensively on automated procedures and closed loop control. These functions often require high speed or high bandwidth and therefore would all be implemented in the on-site layers of the control system, as would time-critical machine protection algorithms, extensive data logging and execution of routine procedures.

The evolution of computer hardware and networks has allowed a migration of computing power from large centralized systems to highly distributed systems. This evolution matches well the growing accelerator complexes. Networks with Gigabit speeds and processors with clock speeds approaching one GHz have pushed far greater control autonomy to lower levels in the controls architecture. These developments favor a 'flat' (non-hierarchical) network structure with intelligent devices that would be directly accessible over the network. Such devices essentially coincide with the current catchword 'Network Appliance', and there will be an immense amount of commercial activity in this direction which will be useful for future projects.

The intelligent device model also implies that the devices be directly on the network rather than hanging on a field-bus below some other device. Traffic can be localized in this structure using 'switches' which forward packets only to the port on which the destination device hangs and whose 'store and forward' capability essentially eliminates Ethernet collisions.

On the accelerator site, the majority of repairs would involve the exchange of modules. This requires that all components be composed of modules of a reasonable, transportable size which have relatively easy to restore interfaces to the other constituents of the component.

On the other hand, the requirements for the hardware components of a remotely operated accelerator are essentially identical to the requirements for any large complex technical facility. The general design criteria are: redundancy of critical parts, if cost considerations allow it; avoidance of single point failures and comprehensive failure analysis; over-engineering of critical components to enhance mean time between failure; standardization of design procedures; quality assurance testing; documentation; standardization of components, parts and material wherever technically reasonable; avoidance of large temperature gradients and thermal stress; and control of humidity and environmental temperature extremes.

Specific features connected to remote operation are foreseen: high modularity of the components to ease troubleshooting and minimize repair time; more complete remote diagnostics with access to all critical test and measurement points necessary to reliably diagnose any failure; and provision for simultaneous operation and observation. If a device is to be fully diagnosable remotely, it is important that a detailed analysis of the requirements be an integral part of the conceptual design of the component.

A survey of engineers and designers in the major accelerator laboratories indicates that all of these design goals are already incorporated in planning for future accelerators. Due to the large number of components, even with an extremely high mean time between failures, one must expect several breakdown events per day. Even for an accelerator that is integrated into an existing laboratory, comprehensive remote diagnostics are obviously necessary to minimize downtime. This will be one of the crucial technical issues for a large new facility. The mean time between failures has to improve by a factor of 5-10 compared to existing facilities like HERA. This is the real challenge and any additional requirements for remote operation are minor by comparison.

The conclusion is that the major technical challenges for the hardware of a future accelerator are due to the large number of components and the required reliability and not to the possibility of remote operation and diagnostics. The additional costs for compatibility with remote operation appear negligible.

5. SUMMARY AND CONCLUSION

We consider a facility which is remote from most of the collaborating institutions, designed, built and operated by a worldwide collaboration of equal partner institutions. Expert staff from each laboratory remains based at its home institution but continues to participate in the operation of the machine after construction. We consider all operation activities. As far as maintenance, troubleshooting and repair is concerned, the experience from existing laboratories is encouraging, indicating that most activities are already performed 'remotely', or could be with properly designed equipment. The experts are only rarely required to be physically present on site. Repairs require an on-site maintenance crew. Most of the interventions are made without an expert or with only telephone assistance. For a future large accelerator facility, we conclude that it should be possible to perform most of the tasks remotely. Maintenance, troubleshooting and repair by non-experts do require comprehensive remote diagnostics, modular design of components, and a high level of standardization. An accelerator could be operated far-remotely. Modern control systems use a layered approach, which appears to be adequate. The rapid rate of development of communications technology should easily support the demands of future accelerator operation. Considering this we conclude that there appears to be no technical obstacle to far-remote control of an accelerator.

Operation of the accelerator is not an easy task. Frontier facilities are inevitably pushing the limits of accelerator technology and present unanticipated difficulties that require intense effort from a dedicated team of experts to diagnose and solve each new problem. Past experience has shown how critical it is for these experts to have offices near each other to facilitate exchange of ideas and information. Equally important is contact between the experimenters and the accelerator physicists, and between the physicists, engineers and operations staff. To encourage an effective interchange between these disparate groups, it will be necessary to have a critical mass of experts located in at least one, if not several, of the laboratories.

During normal operation, the on-site staff required could be much smaller than are typically at existing large laboratories. A reliable number for the minimum staff depends very much on the details of the remote facility but experience from large machines indicates that is could be as small as 200. There would be a much greater number of technical staff of all descriptions actively involved in the accelerator operation remotely.

The major challenge of a remote operation model lies in solving the complex management and communication problems.

ACKNOWLEDGEMENTS

The authors acknowledge that much of the material presented in this report was supplied by Prof. M. Huber, ESA; Dr. M. Ziebell, ESO, M. Jablonka and B. Aune, CEA Saclay, G. Mazitelli, INFN, C. Pirotte, CERN, S. Herb and M. Clausen, H. Freese, M. Bieler, J.P. Jensen, M. Staak, F. Mittag, W. Habbe, M. Nagl, and J. Eckolt, DESY.

6 REFERENCES

[1] P. Czarapata FNAL, D. Hartill, Cornell; S. Myers, CERN; S. Peggs, BNL; N. Phinney, SLAC; M. Serio, INFN; N. Toge, KEK; F. Willeke, DESY C. Zhang,

IHEP Beijing: Report of the Inter-Laboratory Task Force on Remote Operation, to be published
[2] M.Ziebel, ESO,private com., January 2001
[3] HEP Nework meeting, Indianapolis, June 2001 http://lexus.physics.indiana.edu/griphyn/Fermilab_ Connectivity

FRBT002
cs.SE/0111021

SYSTEM INTEGRATION OF HIGH LEVEL APPLICATIONS DURING THE COMMISSIONING OF THE SWISS LIGHT SOURCE

A. Lüdeke, PSI, Switzerland

Abstract

The commissioning of the Swiss Light Source (SLS) started in Feb. 2000 with the Linac, continued in May 2000 with the booster synchrotron and by Dec. 2000 first light in the storage ring were produced. The first four beam lines had to be operational by August 2001. The thorough integration of all subsystems to the control system and a high level of automation was prerequisite to meet the tight time schedule. A careful balanced distribution of functionality into high level and low level applications allowed an optimization of short development cycles and high reliability of the applications.

High level applications were implemented as CORBA based client/server applications (tcl/tk and Java based clients, C++ based servers), IDL applications using EZCA, medm/dm2k screens and tcl/tk applications using CDEV. Low level applications were mainly built as EPICS process databases, SNL state machines and customized drivers. Functionality of the high level application was encapsulated and pushed to lower levels whenever it has proven to be adequate. That enabled to reduce machine setups to a handful of physical parameters and allow the usage of standard EPICS tools for display, archiving and processing of complex physical values. High reliability and reproducibility were achieved with that approach.

1 INTRODUCTION

The construction and commissioning of the Swiss Light Source was done in a very tight time schedule. The top priority was to deliver all required applications in time and do enhancements when needed on the fly during the commissioning. External companies delivered subsystems including the controls, like for the Linac[1] and the 500 MHz RF-system[2]. The requirements for the graphical user interfaces for these systems were done by the system responsible and therefore no effort on a standardization of the interfaces was spend. As a result the high level applications are built in a variety of different languages and even using several different intermediate access methods to the same data.

A careful integration of the high level software was required to limit the negative long term effects on the maintainability and on user interface standardization. Since the man power did not allow to rewrite all high level applications with a standardized interface, the chosen strategy was to smoothly migrate the functionality into lower level until the actual user interface could be replaced by GUIs built with a generic EPICS GUI builder.

[1] Delivered by Accel http://www.accel.de, Controls done by Puls-Plasma Technik (PPT)

[2] Delivered by Thomcast AG, Switzerland

2 APPLICATION ENVIRONMENT

An excerpt of the application environment scheme of the SLS control system is shown in figure 1. The graphical

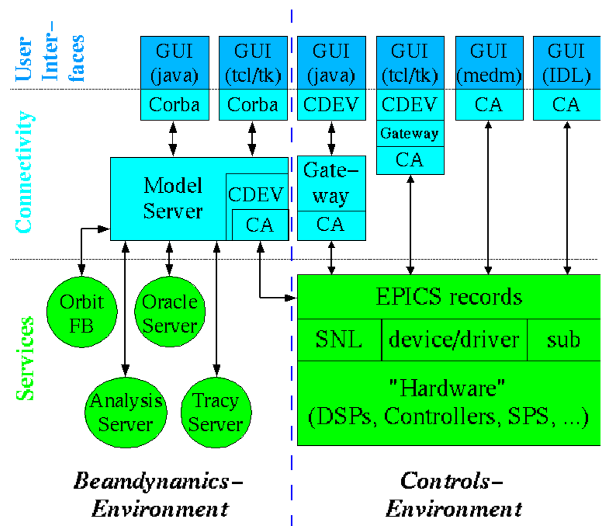


Figure 1: Application model

user interfaces (GUI) are written in a variety of different programming languages. The choice was mostly left to the developer, only under the constrains of connectivity to EPICS and the portability of the programming language. Java, Tcl/Tk (including itcl and the BLT library) and the commercial interpreter language Interactive Data Language (IDL) were used for applications.

The beamdynamics group developed a Common Object Request Broker Architecture (CORBA) environment to support their needs for multiple connectivity. The so called *model-server* provides access for the GUIs as well as for other client/server applications to the EPICS control-system, the Oracle database, to the tracy accelerator simulation tools[1] and an event server. Details about this environment are presented in [2].

The Controls environment uses the Common DEVice (CDEV)[3] middle-ware as primary connection method of the applications. Nevertheless several standard EPICS applications using still the standard channel access methods to connect to the EPICS databases. CDEV also allows additional access methods, like an access to special tables in the Oracle database[4].

3 FUNCTIONALITY MIGRATION

For the development of new applications it has certain advantages to incorporate all data processing into the GUIs.

A variety of development environments are available for the console hosts like Visual Age for Java and scripting languages like tcl/tk or IDL with easy means to test data processing algorithms and to incorporate several views on the same data for debugging purposes. The application should allow to view the raw data, to compare results of different algorithms and to optimize the data filtering. Especially when the data is retrieved by newly developed hardware, these features can be very useful.

On the other hand the maintenance of these customized GUIs is difficult and time consuming in the long term. Therefore it is desired to separate the functionality from the user interface. At the SLS we follow the strategy to migrate functionality that is needed for operation into low-level applications as soon as it works reliable. Either it will be implemented as a server application in the CORBA environment or it is migrated to the VME level: as an EPICS database, a device/driver support or a SNL program.

4 EXAMPLES

In the following some examples are presented to outline the advantages of the delayed functionality migration from the user interface to the service level. Migration of functionality is most often a collaborative task, where the implementation in the different levels are done by various persons. The described examples were realized by the SLS controls, beamdynamics and diagnostic groups.

4.1 Lifetime Calculation

A simple example of a data processing application is the lifetime calculation. The first step was to calculate the lifetime from a precise current measurement with an update period of two seconds, done by a Voltmeter readout via GPIB[3]. A Java application (GUI shown in fig. 2) collected the data and calculated the lifetime by four different algorithms. An EPICS soft channel was used to export the calculated lifetime for other applications. Therefore standard EPICS applications could be used to archive the lifetime for later analysis or to have real-time strip charts together with other channels.

The drawback was, that the data was only generated while the GUI was running. Therefore the algorithm was ported to C as an EPICS device support, after its reliable operation was proved. All parameters of the lifetime calculation

[3]GPIB: IEEE-488 parallel bus

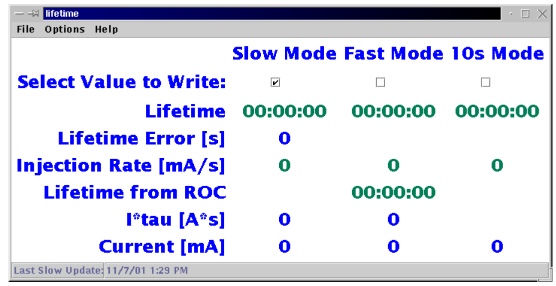


Figure 2: Lifetime application

are now channels and can be controlled and read by standard applications. This has the additional advantage, that no maintenance effort for the customized lifetime GUI is needed.

Tests of an alternative faster readout of the current measurement are now in progress. The increased sample rate will allow to measure the lifetime between the continuous injections in “top-up” mode.

4.2 Magnet Optics Control

A particularity of the SLS storage ring is the individual powering of all 174 quadrupole magnets. This allows very flexible adjustments of the focusing but also contains the risk of a huge parameter space. Right from the start of the ring commissioning a special IDL GUI (see fig. 3) was used to set all elements of the magnet optics according to theoretical calculations with just a few physical parameters for the adjustment to the real machine. The optics was selectable by a menu button and the adjustment parameters were: horizontal- and vertical tune shift, horizontal- and vertical chromaticity shift, a sextupole- and a global energy scaling. The nominal optics values were coded in the GUI and all matrix parameters to calculate the magnet set-currents from the nominal settings and the adjustment parameters. This approach allowed to store and accumulate beam within the first days of commissioning.

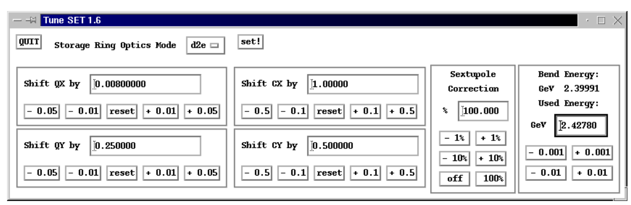


Figure 3: Magnet optics application “Tset”

A clear drawback was that the actual machine state was only known to the application. If the GUI was closed, there was no easy way to deduce the actual used optics and adjustment parameters from the magnet current settings. This was solved by migrating the functionality to an EPICS database. The nominal magnet currents of the optics and the adjustment matrices are now generated by the optics simulation application in a standard EPICS snapshot format. This can be downloaded to the machine using standard save and restore tools. Therefore newly developed optics can be easily applied to the machine.

The actual settings of the machine for a chosen optics are now reduced to very few physical parameters. They are now EPICS channels and all EPICS standard tools can be used to control, save, restore, archive and view them (see fig. 4.) An important advantage, compared to storing the

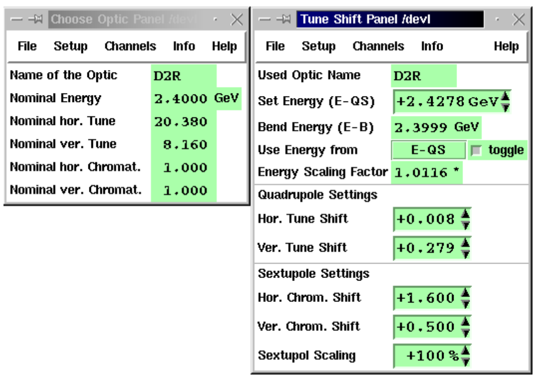


Figure 4: Generic panel.tcl applications to control the magnet optics of the SLS storage ring

set-currents of the magnets, is that these actual set-ups do explain the machine adjustments in physical terms, directly understandable to the accelerator physicist.

4.3 *Orbit Feedback*

The SLS storage ring orbit feedback is designed to allow a global feedback with a regulation loop period of 1 μs. The correction matrix is calculated by a central server using the Singular Value Decomposition (SVD) method. Local DSPs will calculate the actual corrections from the BPM data and write the set-current to the corrector magnet power supplies (see also [5].)

The implementation of this rather complex application was approached in several steps. One of the first steps was to implement an orbit correction by the GUI “oco”. This application connects to the model server to read the BPM data, calculate corrections using the tracy server and writes the corrections to the corrector PS again via the model server. Useful enhancements like the introduction of the accelerating frequency as an 73th corrector and the possibility to disable low weighted eigenvalues for the reduction of the total correction strength were found at that stage. After successful tests of the orbit correction method the sequential control was migrated to a “slow orbit feedback server”. This server is configured, started and stopped by the GUI oco but otherwise runs independently with a loop period of up to a second. First tests were done in a passive mode, where the calculated correction were not applied. Again EPICS soft channels are used to have a standardized interface to watch the activity of the server. Figure 5 shows the archived activity of the slow orbit feedback during a top-up run at 150 mA. The saw tooth behavior of the applied hor-

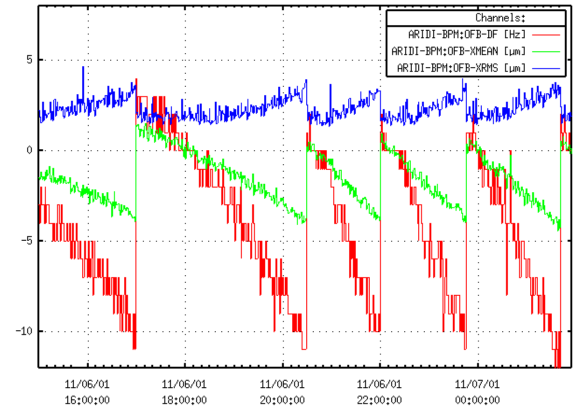


Figure 5: Archived activity of the slow orbit feedback. Standard EPICS tools like the Channel-Archiver and Channel Archive CGI Interface were used to debug the feedback algorithm.

izontal RMS kick (OFB-XRMS) and the horizontal mean kick (OFB-XMEAN) is due to the orbit length correction by minimum frequency steps (OFB-DF) of 10 Hz.

In the next step, the functionality of the feedback server will be partially migrated to the local orbit feedback DSPs.

5 SUMMARY

The successful commissioning of the light source in time was the main goal for the application development at the SLS. All desired functionality was delivered timely and worked satisfactory. The main focus now for the high level applications is to improve the maintainability of the system by separating the required functionality for the operation from the GUIs and provide standardized user interfaces.

The intermediate usage of EPICS soft channels for the data export from the applications proved to enable a transparent migration of the functionality to the low level applications later on.

6 REFERENCES

[1] M. Böge “*Update on TRACY-2 documentation*”, June 1999, PSI, Switzerland, SLS-TME-TA-1999-0002

[2] M. Böge, J. Chrin, “*On the Use of CORBA in High Level Software Applications at the SLS*”, ICALEPCS 2001, San Jose, USA

[3] J. Chen et al., “*An Object-Oriented Class Library for Developing Device Control Applications*”, ICALEPCS 1995, Chicago, USA

[4] T. Pal, “*Storing and transparently accessing CDEV Device properties in a relational database*”, EPICS workshop 1999, Triest, Italy

[5] M. Böge, M. Dehler, T. Schilcher, V. Schlott, R. Ursic, “*Fast Closed Orbit Control in the SLS Storage Ring*”, PSI, Switzerland, PAC 1999, New York, USA

FRBT003
physics/0112065

THE BUNCH INJECTION CONTROLLER FOR THE PEP-II STORAGE RINGS

R. Chestnut, A. Fisher, SLAC, Stanford, CA 94025, USA

Abstract

The PEP-II storage rings at SLAC each have 3492 'buckets' into which electrons and positrons can be injected into the high- and low-energy rings. Equipment to measure the currents of all the individual buckets was originally provided by the Lawrence Berkeley Laboratory and is implemented in VXI-based hardware. Data from this equipment as well as high precision direct current measurement provide the hard data for the Bunch Injection Controller. A large number of parameters determined by injection design considerations as well as set by operators for different circumstances are also used by the software algorithms to determine the desired bucket injection order and charge quantity for each injection pulse. These requests are then passed on to the venerable SLAC master pattern generator, which provides beams for other applications as well. This highly visible and highly successful system is implemented using the EPICS toolkit, and fits well into the merged SLAC EPICS/SLC control system. The Bunch Injection Controller hardware is a VME-based EPICS IOC, which makes extensive use of shared memory for communicating with the VXI measurement equipment and the SLAC master pattern generator.

1 PEP-II INJECTION REQUIREMENTS

The individual 3492 "buckets" in each ring circulate at a frequency of about 136 kHz, and are separated by 2.1 nanoseconds. The SLAC Linac can inject into both rings at 30Hz simultaneously or one ring at 60Hz. There are four "quanta" or injection chunks selectable for each individual injection request. The largest is about 10^10 particles, corresponding to an increase in current of about 218 microamperes; smaller quanta are approximately 2/3 of the next larger.

The most important constraints on injection are:

- Allow arbitrary fill patterns.
- Observe the bunch damping times of 35ms (LER) and 60ms (HER) when considering where to inject next.
- Fill each bucket to within 2% of the requested amount where possible.
- Fill the rings as evenly as possible.
- NEVER overfill a bucket.

2 PRE-EXISTING HARDWARE

2.1 Bunch-by-Bunch Current Monitor (BxBCM)

LBL was in the process of implementing hardware when the injection requirements were under discussion. The BxBCM implementation consists of analog down-converters and a set of VXI-based multiplexers and 8-bit fast ADCs, controlled by a PC-based slot-0 controller running NT. Each ring has its own analog and VXI support. The NT-based software prepares one second sums for each bucket. There is, additionally, control and read-out of attenuations and crate health.

2.2 Direct Current Current Transformer (DCCT)

For each ring the pre-existing DCCT provides a stable voltage derived from the total ring current. These voltages are each digitized twice a second by individual Keithley 2002 Digital Volt Meters.

2.3 Master Pattern Generator (MPG)

The MPG is a well-established iRMX-based machine, which has been running the SLAC beams for two decades. Since a secondary requirement foresees interleaved operation between PEP-II injection and other linac operations, the injection implementation needed to fit seamlessly into the old system.

3 IMPLEMENTATION

We chose to implement the Bunch Injection Controller (BIC) as an EPICS VME-based IOC using shared memory to communicate with the pre-existing computers and GPIB for the Kiethley DVM readout. The first SLAC EPICS project, PEP-II RF control, had been very successful and we were enthusiastic about further use of EPICS. The BIT-3 shared memory provides access with the lowest degree of coupling and fewest software and hardware changes to the pre-existing subsystems.

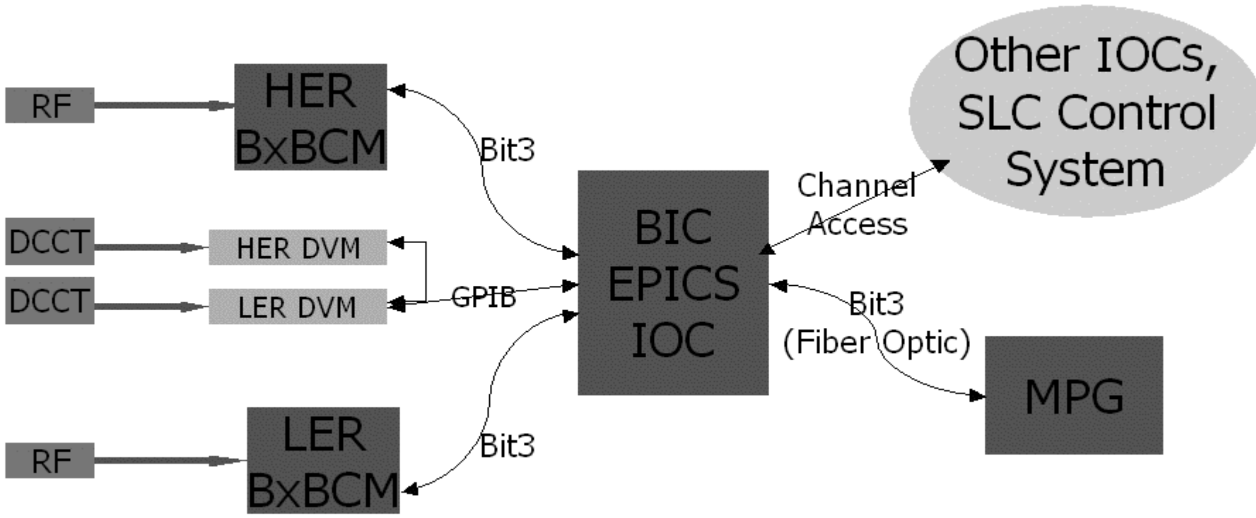


Figure 1: Cartoon of Bunch Injection Controller Interfaces

3.1 Bunch Current Monitor BIT-3 Interface

Each BIT-3 memory is partitioned into a small control segment, a small diagnostic read-back segment, and a large data array segment. Pre- and post-data array update counters are used to guarantee that the BIC reads consistent data. These counters also serve as a heartbeat. There is no other synchronization.

3.2 Master Pattern Generator BIT-3 Interface

The BIT-3 memory is dedicated to two small ring buffers. One contains request information written by the BIC for the MPG and the other contains responses and status from the MPG for the BIC. Both the BIC and MPG code require no synchronization.

The actual injected current is measured by comparing the stable bucket current before injection with stable current after injection. This incremental charge is used to maintain a dynamic calibration of the injection quanta.

The Pulse-ID in the MPG response is used by monitor code in conjunction with Pulse-ID-tagged Toroid measurements in the injection lines for efficiency calculations.

3.3 Kiethley DVM GPIB interface

One GPIB controller serves both the LER and HER Digital Volt Meters. Each DVM is read out at 2 Hz.

4 BUNCH CURRENT PROCESSING

4.1 Hardware Processing

The signal from the beam pick-off is mixed with an RF carrier to form two periods of an amplitude modulated sine wave at 1428 MHz, three times the 476-MHz RF frequency. The high frequency allows an independent current measurement for each bucket with little cross-talk. A pair of fast 8-bit ADCs operating in alternation and each clocked at half the RF frequency generate two streams of digital data, for the even and odd buckets. Each stream goes to a “decimator” board, which distributes the data among 12 Xilinx gate-array processors. Since this data rate is still too high for further processing, data from every eighth bucket are used on each turn. Thus each of the 3492 buckets are measured 250 times each 1/60 second period The data are then copied from the VXI module into the Slot-0 Controller computer memory. After every data accumulation, the phase of the incoming 1428-MHz reference signal is shifted 90 degrees.

4.2 Software Processing

The computer code accumulates two 3492 bucket arrays, one containing 0 degree data minus 180 degree data, the other containing (90-270) degree data. This subtraction removes the background and provides clean “sine” and “cosine” phase data. Each second, after taking 15 samples at each phase, the data are transferred to the shared memory, where the BIC can access the two arrays.

In the BIC, these data are used to generate the amplitudes and phases of each bucket. Since the hardware phase shifter does not shift a true 90 degrees, to first order the actual amplitude is given by:

Amplitude=sqrt(S*S+C*C-2*S*C*delta),

Where delta is the average deviation of the phase shifter from 90 degrees, S is the “sine” value, and C the “cosine” value. This correction provides a wide flat region where the amplitude is insensitive to the phase setting. The value of **delta** is calculated by measuring **S**, **C**, and **amplitude** over a range of phase settings and finding a **delta** that provides the widest flat response.

These data are then normalized to the total ring current and the resulting array of microampere per bucket is used in subsequent processing.

4.3 Filling Algorithm

The fill goal is specified by a pattern description and the total desired ring current. Simple pattern descriptions can be typed directly into a control variable, or complex patterns can be entered into a data file, pointed to by the control variable. The pattern "language" provides shortcuts for specifying whole ranges of buckets and for easily specifying a "ramp", where the desired fill value increases smoothly over a range of buckets. Combining the pattern and total desired ring current creates an array of desired bucket currents (microampere).

There are four different injection chunks (quanta) available for injection into each ring, each about 2/3 of the next larger. These allow for fast bulk injection and "fine tuning" near the end.

To provide a smooth fill, intermediate goals of 1.2, 2.2, etc. times the value of the largest quanta are generated in turn as the previous intermediate goal is filled. Only the largest quanta are used in this phase. At the end of filling, when using the true goal, all quanta are considered.

A candidate bucket is checked to verify that an injection will not overfill. If a candidate qualifies, a request is placed on the queue to the Master Pattern Generator (MPG). The next bucket considered is at least 197 buckets distant, so that the newly injected bucket has time to damp to the nominal beam size. If a candidate does not qualify, the next adjacent bucket is considered. Each second only sufficient requests are made to the MPG queue so that the MPG is never idle. Right near the end of filling, spacers are often inserted in the queue to avoid filling buckets too closely spaced.

This approach is very successful in filling empty rings, topping off, and in filling with many "drop outs", i.e. buckets which lose disproportionately high current due to tune or beam-beam interaction problems.

4.4 Efficiencies and Calibration

The four individual injection quanta are dynamically calibrated by observing the change in current for each satisfied injection request. Each measurement is passed through an anomalous value filter and used as input for a simple smoothing function. Further, the injection efficiency is calculated by comparing injection line toroid measurements to measured bucket current increases, using the MPG linac pulse-id as the correlator, as described above. The numerology describing the linac damping rings and PEP-II ring result in a difference of four buckets for each damping ring turn. For a difference of fewer than four buckets, the linac is rephased. Therefore injection efficiency is separately calculated for each bucket number modulo four.

4.5 Control Interface and Displays

Injection is controlled through a standard EPICS Display Manager control interface. There are some standard control room overhead displays of daily luminosity and currents as well as expanded currents and lifetimes for the last twenty and two minutes. The most useful for filling, however, is a display of actual bucket currents normalized to the goal currents. After a successful fill, all visible points cluster near a value of unity. If buckets not represented in the goal pattern actually have some current, those buckets have the value 1.2. Thus timing problems, resulting in filling the wrong bucket, are easily visible.

5 EXPERIENCES AND PROBLEMS

Three distinct types of problems were noted:

1) GPIB malfunction – multiplexing one digital voltmeter for both rings caused many deadlocks. We now use two digital voltmeters.
2) Bad cables – loose or intermittent BIT-3 cables caused difficult problems.
3) Timing shifts – by far the most frequent problem. Evolving diagnostic displays now enable quick operational response.

REFERENCES

[1] V. Bharadwaj et al., "PEP-II Injection Timing and Controls", SLAC-PUB-7600, Jul 1997. Talk given at 17th IEEE Particle Accelerator Conference (PAC 97), Vancouver, Canada, 12-16 May 1997.

[2] Gregory A. Lowe, "Flexibility of the PEP-II Injection Scheme", SLAC-PEP-II-AP-NOTE-3-94, Feb 1994.

[3] Fieguth, et al., "Injection System for the PEP-II Asymmetric B Factory at SLAC", SLAC-PUB-5775, Mar 1992. Also Berlin 1992, Proceedings, EPAC92, vol. 2 pp. 1443-1445, presented at the 3rd European Particle Accelerator Conference, Berlin, Germany, Mar 24-28, 1992.

FRBT004
physics/0111066

MODEL DRIVEN RAMP CONTROL AT RHIC*

J. van Zeijts, Collider-Accelerator Department,
Brookhaven National Laboratory, Upton, NY 11973, USA

Abstract

At the Relativistic Heavy Ion Collider (RHIC), magnets are ramped from injection energy to storage energy in several minutes where it is to remain for several hours. The path up the ramp is marked by 'StepStones' where the the optics of the machine, which can change dramatically when we perform a beta*-squeeze, is given in units like quadrupole focusing strengths or corrector-dipole angles. The machine is tuned at these Stepstones, and at injection or storage, by specifying physics properties like tunes and chromaticities. An on-line model server handles conversion to magnet strengths, and predicts the optics along the whole ramp.

We will describe the underlying principles, the client-server environment, including on-line model servers, Ramp Manager and Editor, and present operational experience with the system.

1 INTRODUCTION

The RHIC magnets are driven by about 1000 Wave Form Generators (WFG). Most quadrupole magnets are hooked up through a nested power supply scheme, which minimizes the number of high current cryogenic feed-throughs, but complicates their programming considerably. A more detailed description of the ramp control is given in [1], here we concentrate on the physics control and modeling sections.

2 MAGNET CONTROL

Magnets are programmed in physics units like KL (integrated strength), and angle. The WFG's execute formulas at 720Hz that read the machine magnetic rigidity from the real time data link (RTDL), look up the interpolated requested magnet strength, calculate the required field strength, and use the magnetic transfer table to calculate currents for the associated power-supplies.

2.1 StepStones

StepStones are placeholders for a set of magnets and their associated strengths. The strengths is split up in a 'Design', and 'Trim' part. The machine is set to the design level by modifying the trim settings, client applications usually use the design part of the strength for model calculations, since it more closely resembles the real machine. StepStones are sparse, in a sense that only some magnets need to be set explicitly, all other are interpolated as a function of the relativistic gamma. The interpolation scheme is critical for proper power supply performance, and involves cubic splines for quadrupole and sextupole magnet strengths. Other types of magnets use linear interpolation of strengths.

2.2 Ramps

Ramps are placeholders for a set of StepStones. The ramps in use at the moment for RHIC accelerate, and Beta*-squeeze at the same time. The model server does optics simulations at many points along the ramp, giving tunes and chromaticity predictions that can be compared with measured numbers. The model can contain multiple named ramps simultaneously, each containing tens of StepStones (see Fig. 1 for a typical ramp layout, Fig. 2 for a graph of the main quadrupole strength).

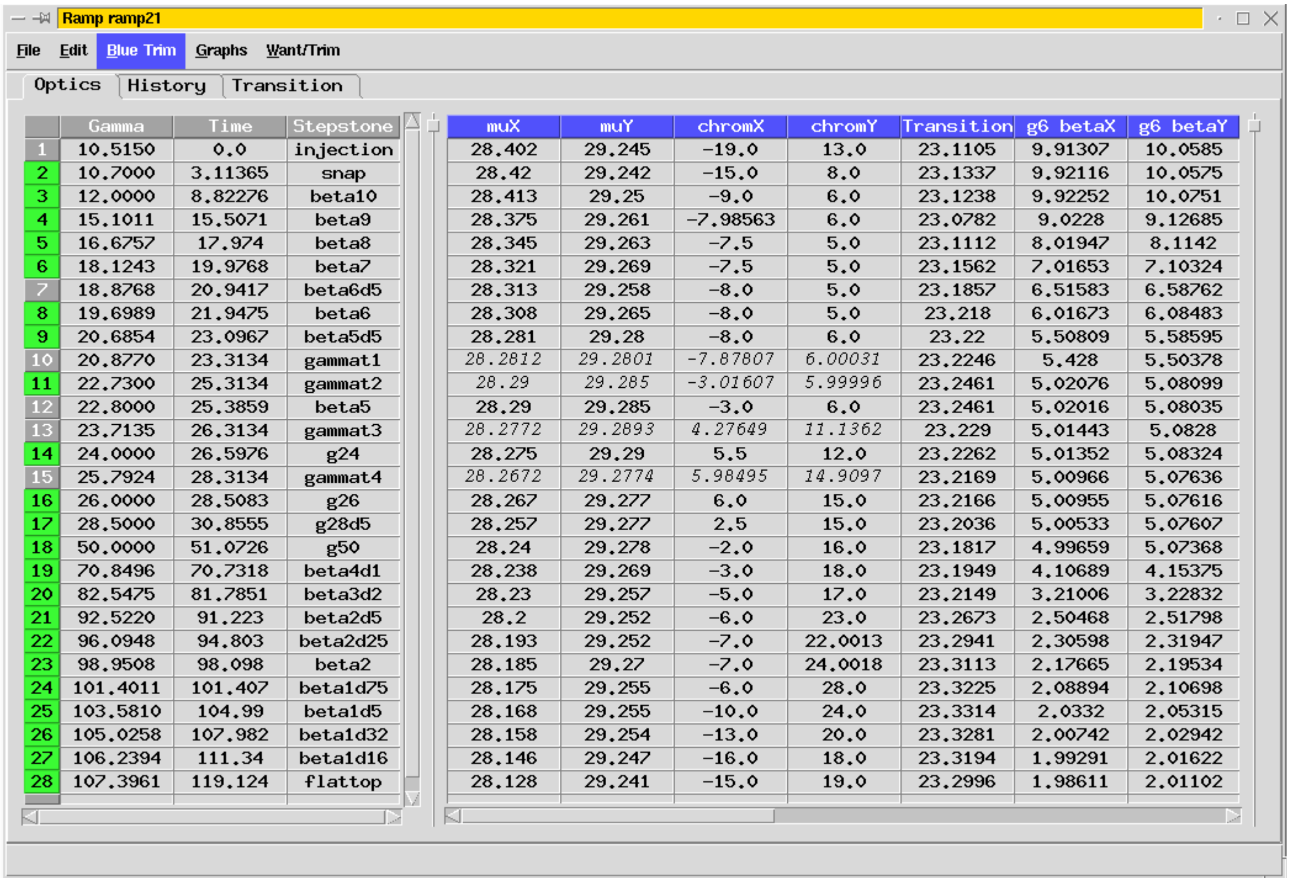
3 MODEL SERVERS

Multiple model servers are available, each presenting an identical interface. The differences are in speed and accuracy. The fast model only considers linear un-coupled optics. There are on-line models available which consider full coupling, non-linearities etc. [2], but with the associated longer execution time. For regular machine operation the linear model is preferred, for studies we can switch to a more complete model.

The model server is implemented using the CDEV [3, 4, 5] generic server framework, which allows for rich data structures to be passed between client and server. Ramps and StepStones are accessible as CDEV devices, and present properties which can be monitored by the client applications. The clients receive updates when magnet strengths are modified. All typical optics properties are exported, the most commonly used ones include:

- 'LatticeFunctions', the clients specify a beam line (Blue or Yellow) and a list of element names. The server by default returns a full set of lattice functions. The context can be modified to only request certain lattice functions.
- 'OpticsFunctions', the clients specify a beam line. The server returns a list of tunes, chromaticities, etc.

*Work supported by U.S. DOE under contract No. DE-AC02-98CH10886.



	Gamma	Time	Stepstone	muX	muY	chromX	chromY	Transition	g6 betaX	g6 betaY
1	10.5150	0.0	injection	28.402	29.245	-19.0	13.0	23.1105	9.91307	10.0585
2	10.7000	3.11365	snap	28.42	29.242	-15.0	8.0	23.1337	9.92116	10.0575
3	12.0000	8.82276	beta10	28.413	29.25	-9.0	6.0	23.1238	9.92252	10.0751
4	15.1011	15.5071	beta9	28.375	29.261	-7.98563	6.0	23.0782	9.0228	9.12685
5	16.6757	17.974	beta8	28.345	29.263	-7.5	5.0	23.1112	8.01947	8.1142
6	18.1243	19.9768	beta7	28.321	29.269	-7.5	5.0	23.1562	7.01653	7.10324
7	18.8768	20.9417	beta6d5	28.313	29.258	-8.0	5.0	23.1857	6.51583	6.58762
8	19.6989	21.9475	beta6	28.308	29.265	-8.0	5.0	23.218	6.01673	6.08483
9	20.6854	23.0967	beta5d5	28.281	29.28	-8.0	6.0	23.22	5.50809	5.58595
10	20.8770	23.3134	gammat1	*28.2812*	*29.2801*	*-7.87807*	*6.00031*	23.2246	5.428	5.50378
11	22.7300	25.3134	gammat2	*28.29*	*29.285*	*-3.01607*	*5.99996*	23.2461	5.02076	5.08099
12	22.8000	25.3859	beta5	28.29	29.285	-3.0	6.0	23.2461	5.02016	5.08035
13	23.7135	26.3134	gammat3	*28.2772*	*29.2893*	*4.27649*	*11.1362*	23.229	5.01443	5.0828
14	24.0000	26.5976	g24	28.275	29.29	5.5	12.0	23.2262	5.01352	5.08324
15	25.7924	28.3134	gammat4	*28.2672*	*29.2774*	*5.98495*	*14.9097*	23.2169	5.00966	5.07636
16	26.0000	28.5083	g26	28.267	29.277	6.0	15.0	23.2166	5.00955	5.07616
17	28.5000	30.8555	g28d5	28.257	29.277	2.5	15.0	23.2036	5.00533	5.07607
18	50.0000	51.0726	g50	28.24	29.278	-2.0	16.0	23.1817	4.99659	5.07368
19	70.8496	70.7318	beta4d1	28.238	29.269	-3.0	18.0	23.1949	4.10689	4.15375
20	82.5475	81.7851	beta3d2	28.23	29.257	-5.0	17.0	23.2149	3.21006	3.22832
21	92.5220	91.223	beta2d5	28.2	29.252	-6.0	23.0	23.2673	2.50468	2.51798
22	96.0948	94.803	beta2d25	28.193	29.252	-7.0	22.0013	23.2941	2.30598	2.31947
23	98.9508	98.098	beta2	28.185	29.27	-7.0	24.0018	23.3113	2.17665	2.19534
24	101.4011	101.407	beta1d75	28.175	29.255	-6.0	28.0	23.3225	2.08894	2.10698
25	103.5810	104.99	beta1d5	28.168	29.255	-10.0	24.0	23.3314	2.0332	2.05315
26	105.0258	107.982	beta1d32	28.158	29.254	-13.0	20.0	23.3281	2.00742	2.02942
27	106.2394	111.34	beta1d16	28.146	29.247	-16.0	18.0	23.3194	1.99291	2.01622
28	107.3961	119.124	flattop	28.128	29.241	-15.0	19.0	23.2996	1.98611	2.01102

Figure 1: High-level display of a ramp in the Ramp-Editor. Tunes and Chromaticities are modified from this page.

- 'Orbit', the clients specify a beam line, and a list of element names. The server returns the predicted orbit using the dipole corrector set points.

4 CLIENT APPLICATIONS

The on-line model server is the hub for lattice and optics information. Magnetic element strength are handled in a separate Ramp-Manager. Applications routinely retrieve and monitor element strengths and lattice functions at specific StepStones, and (at a higher resolution) along the ramp. Below is a subset listed of the client applications connected to the model.

4.1 Ramp Editor

The main ramp control GUI allows modification to tunes, chromaticities, and individual element strengths. On each change the model recalculates the predicted optics at each stone, and along the ramp.

4.2 Injection Application

Injection into both the RHIC rings is facilitated by the 'Injection Application'. This application retrieves the transverse lattice functions in the transfer line and the first sextant of the rings from the model server. Dipole corrections for optimized injection, and closed orbits are calculated and sent to the Ramp Manager. Predicted and measured orbits are displayed.

4.3 Orbit Correction

Global Ring Orbit-Correction, Local Correction, 3 and 4 Bump construction etc. are supported in this application. Dipole correctors strengths are calculated and set though this application. Lattice function information, including phase advance between correctors and Beam Position Monitors (BPM) is retrieved from the model. Predicted and measured orbits are displayed.

4.4 Transverse Profile Manager

Lattice functions at the Profile pickups are monitored by the 'Profile-Manager', measured profiles are then converted to normalized emittance at injection, up the ramp, and at storage energies.

4.5 Luminosity Monitor

Beta functions at the interaction regions are monitored by the 'Luminosity-Monitor', which combines this information with beam intensity and compares measured and predicted luminosity.

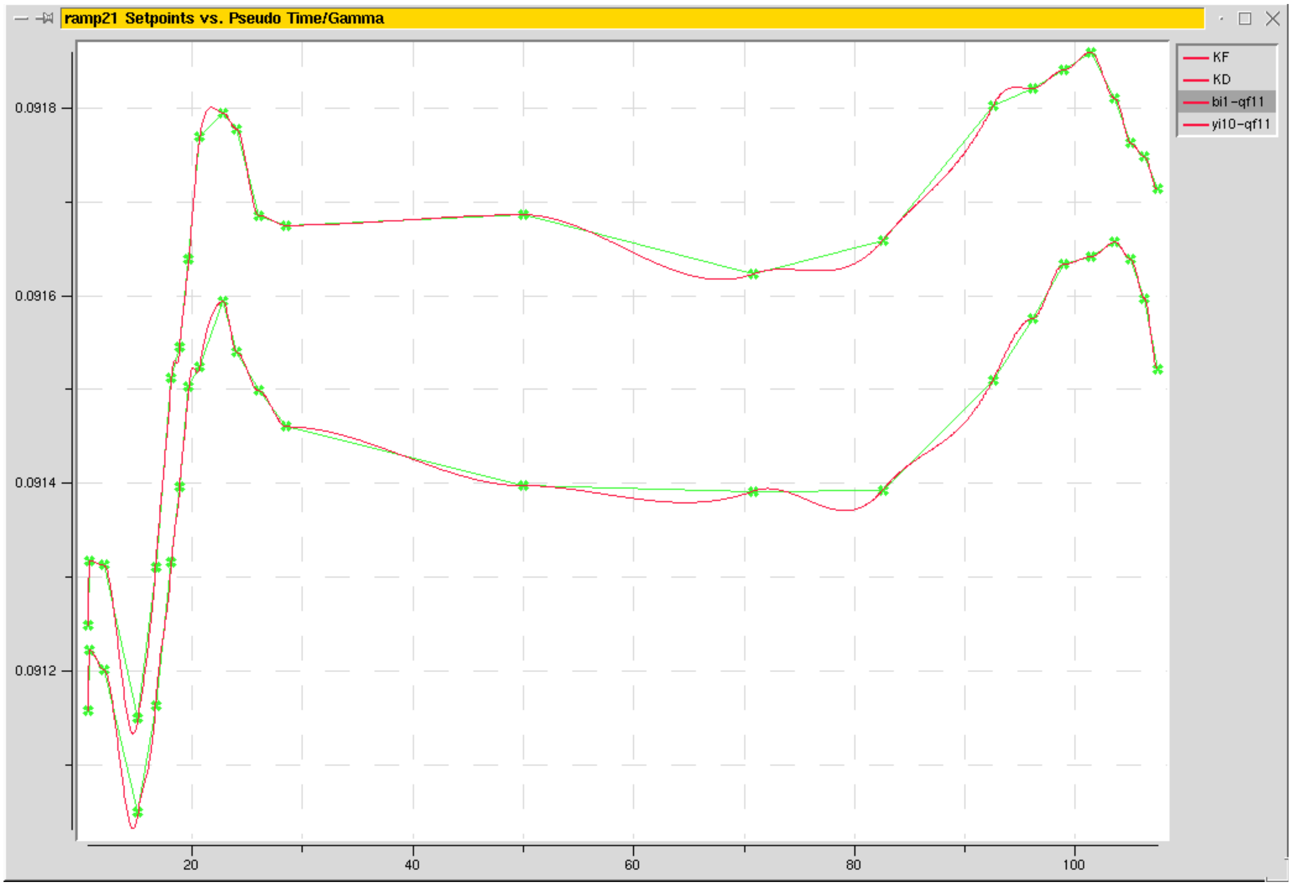


Figure 2: Main quadrupole magnet strength vs. gamma. The green markers are at the location of the StepStones, the smooth lines are the cubic-spline interpolation.

4.6 Coupling-Correction Application

In order to correct transverse coupling in the machine the tune set-points are swept over a given range while plotting the measured tunes vs. set-points. The correction application utilizes the model to calculate the required set points.

4.7 Sequencer

Progress though the many steps required to run the RHIC through its machine cycle is choreographed by the 'Sequencer' program [6, 7]. This program sets the 'liveRamp' and 'liveStone' CDEV devices to their appropriate value during the cycle. The client applications usually use these aliases to get updates on the current optics, instead of named stepstones.

5 OPERATIONAL EXPERIENCE

Having a consistent source of optics information is critical for commissioning a complex machine. The on-line model servers provide such a source. The servers have been in operational use for several years serving client applications routinely used to run the machine. The interface to the servers is through a well defined CDEV interface, which much simplifies the client application programming. The system of servers is flexible, and performs reliably even under simultaneous load of tens of client applications.

6 REFERENCES

[1] J. Kewisch, J. van Zeijts, S. Peggs, T. Satogata, "Ramp Management in RHIC", proceedings of PAC 1999, New York.

[2] T. Satogata, K. Brown, F. Pilat, A. Alai Tafti, S. Tepikian, J. van Zeijts, "The RHIC/AGS Online Model Environments: Design and Overview", proceedings of PAC 1999, New York.

[3] J. Chen, G. Heyes, W. Akers, D. Wu, W. Watson, "CDEV: An Object-Oriented Class Library for Developing Device Control Application", proceedings of ICALEPCS 1995, Chicago.

[4] W. Akers, "An Object-Oriented Framework for Client/Server Applications", proceedings of ICALEPCS 1997, Beijing.

[5] J. van Zeijts, "CDEV Generic Servers for RHIC Commissioning and Operations", proceedings of ICALEPCS 1999, Trieste.

[6] J. van Zeijts, T. D'Ottavio, B. Frak, "The RHIC Sequencer", proceedings of PAC 2001, Chicago.

[7] T. D'Ottavio, B. Frak, J. Morris, T. Satogata, J. van Zeijts, "Description of the RHIC Sequencer System", these proceedings.

FRBT005
physics/0111174

SIGNAL ARCHIVING AND RETRIEVAL: ESSENTIAL LONG TERM PERFORMANCE TUNING TOOL*

R. Bakker, T. Birke, R. Müller, BESSY, Berlin, Germany

Abstract

The first two years of user service of the third generation light source BESSY II emphasized the importance of a reliable, comprehensive and dense logging of a few thousand setpoints, readbacks, status and alarm values. Today data from sources with various characteristics residing in different protected networks are centrally collected and retrievable via an uncomplex CGI program to any desktop system on the site. Data post-processing tools cover Windows applications, IDL, SDDS and custom programs matching users skills and preferences. In this paper illustrative sample data explorations are described that underline the importance of the logging system for operations as well as for the understanding of singular events or long term drifts. Serious shortcomings of the present installation and focus of further development are described.

1 INTRODUCTION

Like other third generation light sources BESSY II exceeds many of the primary design goals. Users not only appreciate the additional potential of the excellent beam definition and stability — an increasing number of experiments simply depend on the high and reliable beam quality. Especially important aspects are minimal beam center of mass drifts, well defined beam energy with minimal spread and high beam intensity with a long lifetime. It is difficult to prevent drifting of these parameters over the several days necessary for an experiment. Many effects originating from facility operating conditions and user activities can contribute. Since there is never enough time to isolate all possible effects by dedicated accelerator development studies, archives of logged data are most important sources of information.

2 SETUP AND STATUS

Despite the eminent importance of archived data the archiving system at BESSY is far from being well settled. Adequate carefully done is the collection both of snapshot files and long term monitoring data [1]. Loss or omission of essential data would destroy unrecoverable knowledge about past behaviour of the facility. Retrieval tools are still cumbersome, immature and subject of maloperation frequently resulting in loss of work time. Configuration is mainly hand-work, thus not fault free. Surveillance of data source availability and data integrity is done occasionally. Only the collector programs themselves are systematically supervised by watch-dog or stop/restart procedures.

2.1 SDDS based Data Store

Initially the BESSY archiving configuration was based on the SDDS toolkit. Storage formats are compressed SDDS files spanning a device class and a full day, sorted into a calendar mapping directory structure. A TclTk glue application combines navigation, SDDS data retrieval, correlation and export [1].

This data store is still a good compromise even though not optimal with respect to data format and size, network resources and CPU requirements: channel selection, previewing facility, available post-processing tools cover most of operators requirements. The SDDS archive is not discontinued, collects 20 GB/y and serves as valuable backup system. A more or less frozen and easy maintainable list of signals essential for the understanding of basic operation parameters are monitored. Major obstacle for a site-wide usage of the archive is the (intended) in-accessibility of the data store residing in the protected accelerator control production area.

2.2 Central Channel Archiver

Since mid 2000 a *Channel Archiver* [2] instance has been set up in addition. It is intended to overcome the self-containment of the (accelerator) SDDS archiver and serve the whole site. Any major development and configuration effort goes into this system. Data collector engine(s) and `CGIExport` retrieval tools are installed in a dedicated environment[1]. A six-processors HP N-class server (*archive server*) in a non-routable private network stores the data on a RAID system that is backed up to a tape robot. It is planned to migrate mass storage to a fibre channel system attached to a tape library this year.

2.3 Data Flow

In an attempt to minimize adverse effects on the system caused by unexpected activities and to maximize uptime neither user accounts nor NFS access to the archiver network are provided. For data collection all data sources residing on dedicated networks are connected by two multi-homed CA-gateway computers (8 network interfaces each). Presently a

* Funded by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) and the Land Berlin

single archiving engine (process) stores 50 GB/y accelerator relevant data. A second engine has been set up early this year for the beamline area and auxiliary data presently collecting about 15GB/y.

Common retrieval method is `HTTP` invocation of `CGIExport` [2] via the central network router. Typically the available `gnuplot` presentation of the data requested is used as a preview ensuring that the data selection provides the desired information. Then the data are retrieved in spread-sheet or Matlab format and stored on a local disk. Favourite postprocessing tools are PC Windows tools (Origin, Excel) or UNIX applications (IDL, Matlab). A small program (`caa2sdds`) converts the spread-sheet output to SDDS format enabling data analysis with the full data selection, post-processing and display power of SDDS.

3 TYPICAL UTILIZATION

3.1 Identification of Singular Events

Probably tracking down sudden perturbations to its causes is the most common usage of the archive. Examples for this application are e.g. an unusual large drift that corresponded to the failure of a water pump or the sudden onset of orbit jumps that was due to an improper motor reset resulting in a constant rotation of strong chicane magnets.

3.2 First Hints on Unexpected Effects

Archived data help to get a first idea of possible explanations: mid. 2001 for example a strong, periodic orbit perturbation has been reported by the operators. By phase analysis it was possible to locate the problem source with a few meters precision at a ring segment where no active elements are installed. The time pattern of perturbation onset and disappearance (see fig. 1) suggested an unknown correlation with user activities. Targeted investigation found out that one user group reversed the field of a 1 [T] magnet twice a minute several meters apart from the beampipe.

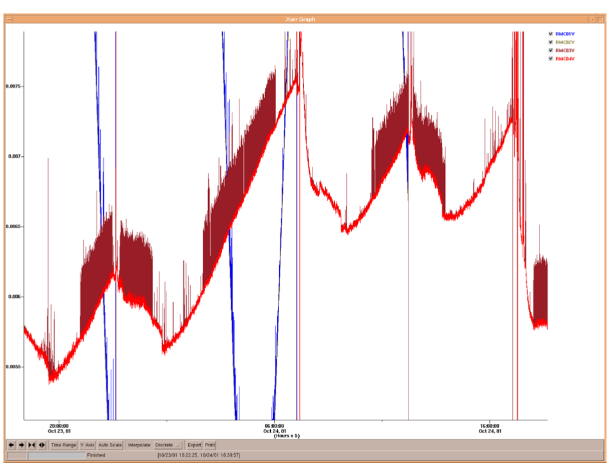


Figure 1: Orbit perturbations due to switched user magnet outside the storage ring tunnel. Phases of experimental activity are clearly visible. Viewing tool: `Xarr`.

3.3 Analysis of Changes

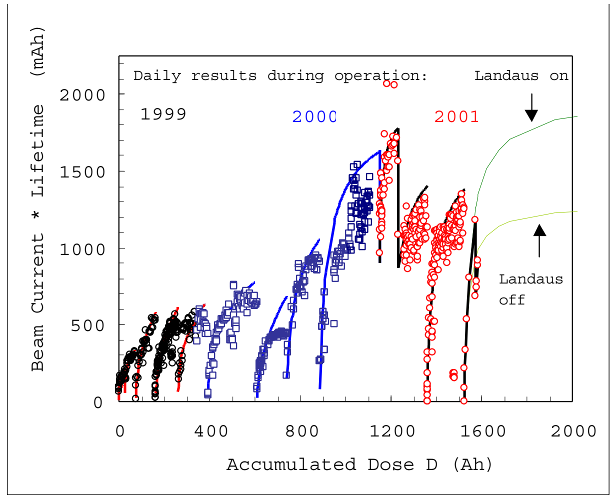


Figure 2: Vacuum effects on lifetime shown as a function of accumulated beam dose. Postprocessing tool: `EXCEL`

On the long term extreme the archive provides the data needed to make evolutions visible that are hardly perceptible on a fill to fill basis. Plotting e.g. the normalized lifetime [mAh] against the accumulated dose [Ah] over the full operating time of the facility is a powerful mean to find out very fundamental factors: From fig. 2 it can be concluded that the vacuum related lifetime reduction is basically overcome by beam scrubbing 1000 [Ah] after start up of the accelerator. Every venting due to installation requirements needs another 100 [Ah] to reinstall the previous performance. On top of these basic conditions global lifetime improving effects of Landau cavities (mid. 2000) as well as reducing effects of imperfectly corrected insertion devices (beginning 2001) can be seen.

4 DEMANDING REQUIREMENTS

4.1 Uptime, Reliability

Requirements on uptime, reliability and consistency of the archive are substantial. The archive data has to contain signals of very different importance. Beam intensity is analyzed and correlated in any thinkable way e.g. integration (dose), differentiation (beam loss), pattern analysis (user runs) etc. Here a loss of data would be serious, but recognized within minutes. Other signals are monitored as a precaution. They could potentially help to find candidates for sources of performance degradation. Dispensable for the all day business they are not under human surveillance. Regardless they have to contain reliable data when needed.

4.2 Data Density, Aging

The most common approaches to prevent growing of the archive to unmanageable dimensions are removal of 'old'

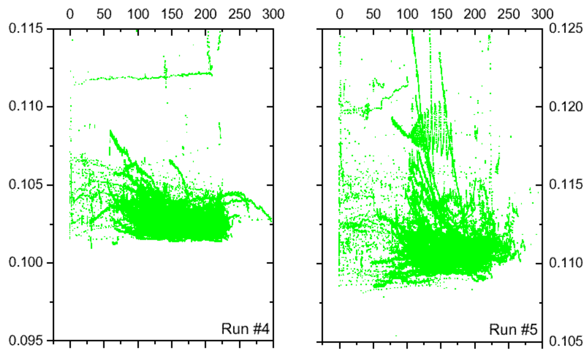


Figure 3: Raw data of (uncorrected) vertical orbit stability (+/- 5 μm RMS) during all user fills (220 mA - 80 mA) at user run #4 (Aug. 2000, left). General degradation and spurious exotic drifts can be clearly identified at run #5 (Oct. 2000, right). Postprocessing Tool: `Origen`

data (tape, deletion) or a progressive reduction of data density. Fig. 3 and 4 are examples of the opposite requirements for a dense and long term archive. In Fig. 3 spurious observations and user complaints could be quantified after serious hardware modifications. The comparison of performance and influence of a new operation mode required per fill details (8h) months apart for fig. 4.

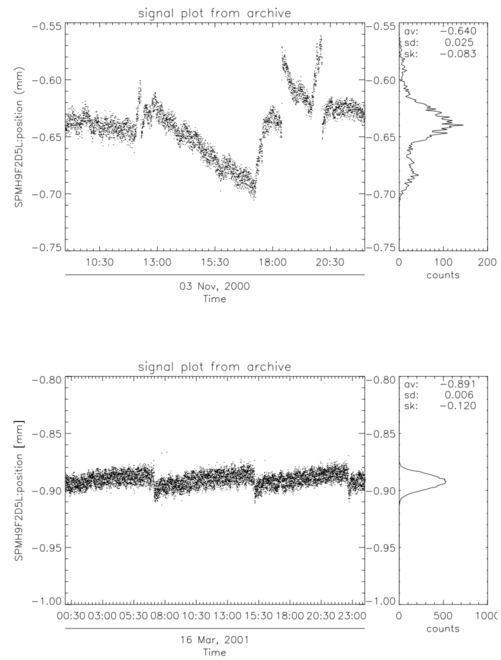


Figure 4: Stability comparison of 'Uncorrected' with 'Drift Corrected' Fills. Data Postprocessing Tool: `IDL`.

5 PRESENT FOCUS OF ACTIVITIES

5.1 Data Collector

Today management and configuration of collector engines is further robustified. Usage of the system is simplified by GUI administration tools. Signal configuration management based on the reference RDB is still missing.

5.2 Retrieval

Performance of data retrieval from large and multiple archives has been drastically enhanced. Channel detection method for a given time interval is improved. Volume of intermediate data needed for previewing is reduced to the minimum allowed by the anticipated `gnuplot` resolution.

5.3 Data Partitioning

From the iterator model and the hash table directories the binary data format of the *Channel Archiver* is optimized for retrieval of data from archives containing a moderate number of channels and starting e.g. from 'now' going backwards in time. Retrieving a dozen of channels out of the 'middle' of a continuous archive holding several thousands of channels requires patience.

As a first improvement approach the huge monolithic data block is split into a moderate number of weekly ordered chunks holding certain fragments of the whole signal collection. Adjustment of the I/O routines results in orders of magnitude retrieval acceleration. But however home grown data formats are optimized: ultimately the retrieval of arbitrary data selections out of huge data stores is best done with commercial RDB systems. Consequently the utilization of a RDB storage format has to be re-considered.

6 SUMMARY

Ideally one would like to be able to 'replay' any controllable and measurable parameter out of the signal archive with the reasonable time resolution of a few seconds. For a BESSY size facility this would require data stores of several TB/y. The *Channel Archiver* provides a robust data collector and retrieval toolkit but the archive itself has to be reduced to manageable dimensions.

The challenges today are configuration (select relevant signals, grouping, choose proper archiving frequencies), correlation detection (identify signals) and data organisation (optimized search). Plotting options and postprocessing requirements have to be provided by the end-user according to his specific skills and varying needs.

7 REFERENCES

[1] T. Birke, R.J. Bakker, S. Bub, R. Müller, "Logging of Supervisory Data at BESSY", Proc. of the 7th Int. Conf on Accelerator and Large Experimental Physics Control Systems (ICALEPCS99),Trieste, Italy, p. 431 (1999).

[2] K. Kasemir, L. Dalesio, "Overview of the Experimental Physics and Industrial Control System (EPICS) Channel Archiver", these proceedings, THAP019.