CONTROL SYSTEM AVAILABILITY FOR THE SPALLATION NEUTRON SOURCE[*]

S. M. Hartman[†], Spallation Neutron Source,
Oak Ridge National Laboratory, Oak Ridge, TN, USA

Abstract

The Spallation Neutron Source (SNS) is continuing its ramp up of beam power, while simultaneously increasing production hours and striving for reduced unplanned downtime. For the large, highly-distributed EPICS-based control system of the SNS, this demand for increased availability is combined with the need for ongoing system maintenance, upgrades and improvements. Causes of recent control system related downtime will be reviewed along with experiences in addressing the competing needs of availability and system improvements.

INTRODUCTION

The SNS is an accelerator based neutron source located at Oak Ridge National Laboratory. The integrated control system (ICS) along with the rest of SNS, was built by a partnership of six Department of Energy National Laboratories. The scope of the ICS includes accelerator systems, a cryogenics plant for a super-conducting radio-frequency (RF) linac, conventional facilities, machine protection system, and personnel protection system. The control system input/output consists of approximately 150 VME based Input-Output Controllers (IOCs) running VxWorks, approximately 100 softIOC instances running on 8 commodity servers running Linux, over 100 programmable logic controllers (PLCs) (primarily Allen-Bradley ControlLogix), and a variety of additional devices (generally networked devices with a softIOC for interface to the control system). In addition, the beam instrumentation group manages more then 300 additional IOCs. The network is distributed over a number of service buildings across the SNS site, with over 100 network switches. Several dozen servers are located in a central server room and additional servers and workstations are located throughout the site.

The system is highly distributed such that there are few faults which can interrupt the entire control system, limited primarily to electrical power and the networking infrastructure. Electrical issues are mitigated by the extensive use of uninterruptible power supply systems and automatic transfer switches to provide multiple power feeds to critical devices. The networking infrastructure also offers some redundancy (redundant network switches in key places with redundant fiber connections).

However, in terms of availability of the neutron source, any fault in the control system which interrupts operation of the accelerator complex results in machine downtime. For the control system, there are hundreds of devices (primarily IOCs and PLCs) whose failure will interrupt operations. These devices are generally assemblies of components (power supplies, processors, input-out modules, communications modules, firmware, software, etc.) all of which must be functional for the assembly to work.

Availability Goals

For the fiscal year 2009, the overall availability goal for the SNS is 80% (actual run time divided by scheduled run time). This translates to a "downtime budget" goal for the controls group of only 45 hours of unscheduled downtime. With projections for increased availability in coming years, only 4% of the budgeted system downtime is allocated for controls (not including protection systems). With a goal of 90% availability for a 5000 hour production schedule, the proposed downtime budget for controls (not including protection systems) is only 22 hours (99.56% ICS availability). For a plan of 95% availbility and a 5000 hour production schedule, only 11 hours is budgeted for controls (99.78% ICS availability).

The budgeted downtime includes time for troubleshooting, repair and recovery, and during non-business hours or during holidays, possibly transit time for an engineer or technician to arrive on-site. Recovery time may be extensive for some systems, such as a personnel protection system requiring re-certification, or a cryogenics system which results in the trip of a cold box. It becomes quite clear that component level reliability must be exceedingly high and software must be quite robust.

CATEGORIZATION OF DOWNTIME

At SNS, machine downtime is recorded by the operations group as part of beam time accounting procedures. The downtime is recorded per twelve-hour operations shift, assigned to a category (accelerator division group) and subcategory (subsystem) with a short description describing the fault. The data is stored for later retrieval from a relational database.

A summary of yearly downtime for fiscal years 2009, 2008 and 2007 for the controls group is shown in Tables 1, 2 and 3.

[*] SNS is managed by UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the U.S. Department of Energy

[†] hartmansm@ornl.gov

Table 1: Controls Downtime Summary, FY 2009

Category	Freq.	Time (h)
ICS PLCs and Remote I/O	5	25.1
ICS Computer Systems	6	19.9
Machine Protection System	28	15.8
PPS Radiation Monitors	9	11.8
PPS Target Protection System	4	10.4
Cooling Systems - Accelerator	7	9.9
RF - Low Level	4	4.4
PPS	4	4.2
Power Supplies	4	3.7
Cryogenics	3	3.6
ICS Timing System	3	2.4
RF - High Power	3	2.4
ICS IOCs	3	1.7
MPS Fast Protect, Latched	3	0.5
Target Moderator Systems	1	0.4
ICS Software	1	0.4
ICS Network	1	0.1
Test Facilities and Networks	1	0.1
Total Towards Down Time		116.8

Table 2: Controls Downtime Summary, FY 2008

Category	Freq.	Time (h)
Cooling Systems - Accelerator	11	9.7
MPS Fast Protect, Latched	26	8.3
PPS	2	7.6
ICS IOCs	10	7.3
MPS Fast Protect, AR	1	7.0
ICS Timing System	4	6.8
ICS Software	11	5.2
Vacuum System	4	4.6
ICS Network	1	3.4
MPS	8	2.0
PPS Radiation Monitors	2	1.5
ICS PLCs and Remote I/O	3	1.5
ICS Computer Systems	1	1.0
RF - Low Level	2	0.6
Power Supplies	4	0.6
PPS ODH	1	0.3
Total Towards Down Time		67.4

Table 3: Controls Downtime Summary, FY 2007

Category	Freq.	Time (h)
ICS Network	19	66.3
MPS Fast Protect, Latched	54	36.8
PPS Target Protection System	4	14.9
ICS Software	11	13.7
ICS Computer Systems	7	13.0
MPS	15	12.4
MPS Fast Protect, AR	9	6.4
Target Moderator Systems	1	5.7
ICS Timing System	7	4.8
PPS Gamma Blocker	1	2.0
ICS Software	1	1.5
PPS ODH	2	1.1
PPS Radiation Monitors	1	1.0
Test Facilities and Networks	1	0.7
Total Towards Down Time		180.3

Downtime Data

The collected downtime data offer some useful information as to frequency and duration of faults for the subsystems, and give some value towards prioritizing areas for improvement. But the utility of the data is limited by the criteria used for categorization and lack of integration with other tools for documenting troubleshooting and fixes.

From an operational point of view, categorizing the fault in terms of system/subsystem (e.g. controls/timing or controls/RF) is descriptive in terms of observed impact to the machine and is effective in terms of system owner notification. However, for a reliability analysis, it offers little insight to the types of faults and recurring patterns [1]. Of greater interest would be categorization based on the failed device. For software, that may be type of bug (e.g. buffer overflow, lack of error handling, logic fault, etc.) or type of software (e.g. PLC ladder logic, IOC configuration, etc.). For hardware, the category of interest may be type of component failure (e.g. VME power supply, PLC discrete output module, ethernet module firmware). However, this level of information is rarely available at the time of the downtime report and may not be readily apparent even after troubleshooting and recovery.

For instance, an extensive downtime period (resulting from a long recovery process) in early 2007 had initially been categorized as “ICS Network” based on observations made at the time of the incident and in recovery afterward. It was more then two years later after the same fault occurred a second time that the problem was traced to a bug in a PLC ethernet module which manifested itself every 825 days [2].

For the categories currently in use at the SNS, another layer of confusion is added by mixing subsystems and components in the categorization. A fault in an IOC controlling magnet power supplies may be categorized as a problem in the controls power supply subsystem on one occurrence, but categorized as an IOC problem on a later occurrence, obscuring the relationship between the two events. For problems involving multiple systems or for bugs in the inter-relationship of systems, such as a fault in one system generating an error which is not properly handled in a different system, the relationships become even more complex.

ADDRESSING POINTS OF FAILURE

From the data that are collected, the task is to identify trends in types of failures and ideally find remediation. In some cases it may be a component which has aged to the point of an increased failure rate, or a systematic design issue.

The summary of downtime for 2008 shows a marked improvement over 2007, with the sum of controls related downtime decreasing from 180.3 hours to 67.4 hours during a time period when operational hours were increasing. However, 2009 showed an unexpected increase in controls related downtime. A review of the top three contributing categories in Table 1 shows several types of failures with different causes and solutions.

The Machine Protection Category included a number of events related to noise leading to false trips of the protection system. The increase in noise was directly related to the increase in beam power as compared to the previous year. As the limitation in the existing system was identified, noise mitigation techniques were employed to reduce the occurrence of the faults.

For the ICS Computer Systems category, four of the six events, totaling 19.4 of the 19.9 hours, were actually the same occurrence. Most of the downtime was for the recovery after an IOC fault tripped the cryogenic cold box. An extensive review of the fault eventually led to the discovery of software bug which led to memory corruption on the IOC. The bug had actually been in place on approximately 100 systems for several years without causing any major impact. Several previously unexplained IOC faults with less downtime impact were likely the result of this bug, but with no pattern noticed until this incident was reviewed.

The most dramatic rise in downtime contribution for 2009 was seen in the PLC category, with 25.1 hours plus several more hours for downtime categorized by subsystem rather then hardware. Several of the instances involved extended recovery time after a fault in a cryogenic system PLC. Review of each incident revealed no strong pattern of location, type of module or type of fault. A few units had what appeared to be a hardware fault, but attempts to have the modules analyzed by the manufacturer proved to be unhelpful.

After expanding the review to include related incidents in prior years and to include non-downtime related incidents two dominant types of faults were revealed. The first type was failures of processor modules, and included a manufacturing defect, loss of program from volatile memory, and single-bit memory errors. All three of these issues have been addressed in the newer generation of processor modules for this line of PLCs. The second type of fault concerned communication modules. While there was no consistent fault found, a number of observed events proved to be associated with bugs fixed in more recent firmware. The SNS controls group decided to standardize on the newer generation of processors and newer revisions of firmware for processors and communications modules. During a scheduled maintenance period in the summer of 2009, processors and firmware revision levels were upgraded on approximately 40% of the PLC systems in use with plans to continue the upgrade in later maintenance periods. At this time, approximately two months after the first round of upgrades, there is insufficient data to determine the impact of these upgrades to control system availability.

CONCLUSION

In order to improve control system availability for an operational accelerator or other large experimental physics project, it is necessary to analyze any faults or failures of the system. While there may be urgency in getting the system recovered as quickly as possible to resume operations, the effort to take data and understand the fault, much of which can be done in parallel or after recovery, pays dividends towards long term availability goals. In conducting a more thorough review of recent downtime events, several bugs which had been in place for years causing occasional unexplained faults were identified and fixed.

It also becomes important to categorize downtime events in a way which is useful in finding patterns or trends and to correlate events over time. Having data available describing similar events can greatly speed up troubleshooting and identify issues requiring greater scrutiny and follow up.

The SNS downtime reporting tool currently lacks a way to show relationships between downtime events beyond the broad subsystem category, or to re-categorize events after troubleshooting. It also lacks a way to document response and subsequent improvements made afterward. Ongoing efforts to improve availability would benefit from having these tools available.

Acknowledgment

Thanks to Dave Purcell for providing the relational database reporting tools for viewing the collected downtime data.

REFERENCES

[1] S. M. Hartman, “Failure Mode Effects Analysis for an Accelerator Control System,” these proceedings.

[2] Rockwell Automation Tech Note, “After 825 Days or More of Continuous Operation, EtherNet/IP Devices May Stop Communicating for 20 Minutes,” TN No. P164757202 (2006).