RELIABILITY SIMULATIONS FOR A LINEAR COLLIDER*

N. Phinney[#], T. Himel, M.C. Ross, SLAC, Stanford, CA 94309, USA
P. Czarapata, H. Edwards, M. Huening, FNAL, Batavia, IL 60510, USA

Abstract

A new flexible tool for evaluating accelerator availability was developed as part of the US Linear Collider Technology Options Study. The linear collider designs considered were based on the GLC/NLC X-band and TESLA Superconducting proposals, but modified to meet the US physics requirements. To better model some of the complexities of actual operation, a simulation program was written, which included details such as partial fixes or workarounds, hot-swappable repairs, multiple simultaneous repairs, cooldown periods before access, staged recovery from an outage, and both opportunistic and scheduled machine development. The main linacs and damping rings were modeled in detail with component counts taken from the designs, and using MTBFs and MTTRs from existing accelerator experience. Other regions were assigned a nominal overall failure rate. Variants such as a single tunnel or conventional positron source were also evaluated, and estimates made of the sensitivity to recovery or repair times. While neither design was predicted to be sufficiently reliable given present experience, the required improvements were estimated to increase the overall project cost by only a few percent.

INTRODUCTION

An evaluation of system availability is part of the detailed design of any new accelerator project. Typically, this has been done using a spread sheet (as for the Spallation Neutron Source) or a commercially available reliability software package. Neither approach is entirely satisfactory as they do not easily handle the complexities inherent in an accelerator design, such as built-in redundancies, soft failure modes, or partial work-arounds. They also ignore the ability to schedule multiple repairs during a single outage, the serial nature of tuning and recovery procedures, and the possibility of performing machine studies in one part of the complex while another region is under repair. Because of these considerations, a Monte Carlo simulation was developed for availability analysis as part of the US Linear Collider Technology Options Study (USLCTOS) [1].

One motivation for the pre-construction availability analysis is to identify those components of the project which pose the greatest risk to reliable performance. The project can then allocate engineering resources to improving reliability in the most cost effective way and concentrate on those components with the largest projected impact.

DESCRIPTION OF THE SIMULATION

The simulation starts from a file-driven description of the accelerator complex. The machine is broken into regions and both the regions and the order in which the beam passes through each region is described in an input file. For each region to be analyzed, a detailed list of components is generated, with counts for each item. The list should include all items identified as potential sources of failure such as rf components, magnets, vacuum pumps, power supplies, controllers, movers, diagnostics, control system elements, AC power circuits, water systems, etc. Each component is assigned a nominal Mean Time Between Failure (MTBF) and Mean Time To Repair (MTTR). These numbers are intended to represent steady state performance after several years of operation and ignore the typical reliability ‘bathtub curve’ with higher failure rates in early years due to infant mortality and in later years due to end of expected life.

Each component must then be characterized in terms of its accessibility, impact on operation, and possibility of partial fixes or work-arounds. An advantage of the simulation is the flexibility provided for describing complex component interrelationships.

Accessibility

An important consideration is accessibility for repair. Beam line components such as magnets or vacuum pipes require access to the accelerator tunnel, with associated overheads for radiation cooldown before entry and restoration of personnel security and possibly magnet cycling or temperature stabilization afterwards. These overheads are in addition to the actual repair time. Failure of a critical power supply located in a support building will interrupt operation but can be repaired much more quickly because it is readily accessible. Many other components located in the support housings can be designed to be replaceable without interrupting normal operation. These could include klystrons, modulators, low-level rf, vacuum pump supplies, controls electronics, water system pumps and controls, sub-units of modular supplies, and much more. Attention is required to ensure that repairs of such devices are truly transparent to operations. For the simulation, each component is labeled with one of three states *Requires tunnel access*, *Accessible but invasive*, or *Accessible and hot-swappable*.

Impact of Failures

The impact of a failed component varies widely and needs to be characterized in some detail. Some systems are essential and any failure interrupts operation. An example would be the bend or quadrupole magnets in a ring, where the beam cannot circulate until the repair is

---

[*] Work supported by Department of Energy contract DE–AC03–76SF00515 (SLAC) and DE-AC02-76CHO3000 (FNAL).

[#] nan@slac.stanford.edu

complete. Some systems can degrade the luminosity without turning off the machine. Examples would be diagnostics such as beam position monitors, and correction elements such as steering dipoles or movable supports. Generally, operation will continue with some of these components out of service and they can be repaired at the next opportunity. Some systems have built-in redundancy and will not impact operation until all of the overhead or spare units have been exhausted. The most common example is the components of the accelerating rf systems. Large rf systems usually incorporate spare units which can be rapidly switched in to replace a failed unit.

For the simulation, each component is first identified as impacting either luminosity or a quantity such as energy with built-in overhead. It is then assigned a numerical value indicating how much it degrades the performance of that parameter. Failure of a critical component reduces the luminosity to zero, while failure of a diagnostic or correction element might reduce the luminosity by only a few percent. Failure of a device in the rf system would reduce the energy overhead by a given amount, depending on how many accelerating structures were affected. For example, a problem within a cryomodule might affect all structures in that module, while a klystron or modulator failure would affect all cryomodules attached to that rf source, and a water system failure would affect all rf sources on that supply.

Partial or Deferred Repairs

Accelerators often achieve reliability by using work-arounds for some failures which would be too time-consuming to repair completely. The simulation provides hooks for explicitly incorporating such repair strategies. An example would be a broken quadrupole or structure mover in the warm linac where repair would require tunnel access. Such a failure is modeled as causing a short interruption to match around the failed unit and then as reducing the luminosity slightly until the repair is made during the next access. Similarly, a problem with a coupler on an individual superconducting cavity might cause the entire module to be operated at reduced gradient until the next access when that cavity could be disconnected. The impact on energy overhead would be greater until the access was taken.

Downtime Planning

The facility must be shutdown for repairs whenever a critical component fails or when energy or other overhead is exhausted. If access to the accelerator tunnel is required, one hour is allowed for prompt radiation to decay before entry and one hour for lock up and other turn on procedures. These values are simply input parameters. The simulation attempts to optimize the repairs performed much as would be done in real life. In addition to repairing the component(s) that precipitated the shutdown, other devices are repaired as long as repair staff are available and their repair would not extend the downtime by more than 50% (without access) or 100% (with). The number of people available inside and outside of the tunnel are adjustable parameters, but individual repair specialties are not tracked.

A notable feature of the simulation is that all repair interruptions are treated equally whether they are unexpected or so-called *scheduled maintenance*. This reflects the experience at facilities like SLAC and FNAL where the accelerator components do not require attention during a 9-month run and maintenance days are only *scheduled* when there are too many accumulated failures. In all cases, they represent time when the facility is not delivering beam. If the available operation period is pre-determined, as is usually the case, any intervention subtracts from the integrated facility performance.

Recovery Time after Repairs

The common experience is that an accelerator takes considerable time to recover good performance after any interruption. The length of recovery time is roughly proportional to the downtime, with typical ratios on the order of ½ to 1. Rather than attempt to model recovery procedures in detail, the simulation simply assumes that the time to produce good beam from a region is proportional to the time off. The beam is recovered through each region sequentially with parameters set to typically 10 or 20% depending on the complexity of the tuning for the region. The attempt to model the recovery time is a unique feature of this simulation, but should more accurately reflect the true impact of a component failure on the integrated performance of the facility.

Machine Development

Machine Development (MD) is an essential tax on the operating efficiency of any accelerator. It is time used to better characterize the machine, develop new tuning procedures, and test possible future improvements. Typically, more MD is required in the early years of operation, but the time allocated for a more mature facility would be about 10%. As with recovery time, the MD was apportioned among the regions of the machine, with parameters set to 1 or 2%. The simulation assumed that some of the required MD could be done on an opportunistic basis, when beam was available in one region while another region was being repaired. Such time was counted and subtracted from the total MD allocation.

RESULTS FROM THE STUDY

The US linear collider study defined X-band and L-band reference designs with a starting energy of 500 GeV, in a tunnel long enough to support an upgrade to 1 TeV. The configurations were chosen to satisfy the physics requirements specified by the American Linear Collider Physics Group [2], which are essentially identical to those later set by the international parameters document. To facilitate comparison, the baseline designs were made similar wherever possible. Both options have a twin tunnel configuration and an undulator positron source. A similar level of conservatism was applied to diagnostics, overheads and emittance budgets for both designs. In addition to the baseline designs, several variants were considered, among them a conventional positron source and a single tunnel configuration for the L-band collider.

For the availability analysis, a goal of 75% availability was set. Experience from operating accelerators indicates that a significant fraction of downtime is due to design flaws or oversights, so 10% of the downtime was held as contingency and only 15% was assigned to known components. Due to lack of time, only the main linacs and damping rings were modeled in detail. Other regions were assigned an overall MTBF and MTTR per region.

MTBFs and MTTRs for individual components were taken as much as possible from data on operating accelerators at FNAL and SLAC. Some work had already been done to design more reliable devices, such as redundant magnet power supplies or solid-state modulators and these improvements were taken into account. In a few cases, it was necessary to extrapolate from present experience based on planned reliability engineering, as for the L-band cryogenics system which was specified to be a factor of 6 better than present day systems. A description of all the devices with assumed MTBFs and MTTRs was included in the report.

A summary of simulation results for different configurations is given in Table 1. Included are warm and cold designs with 2 tunnels and a conventional or undulator positron source. The predicted downtime with the initial MTBFs was similar for both designs, about 40%. This could be reduced to 25% by improving the component reliability for an additional cost which was crudely estimated to be about 2% of total project cost. With a single tunnel configuration, the Cold downtime increased by 10%. The additional cost to achieve the target availability was estimated at 5%. These numbers are a first very rough estimate, not the result of a careful cost benefit analysis of which components to improve.

Table 1: Summary of simulation results for the warm and cold 2 tunnel designs with initial and improved (imprvd) component reliablity and comparing undulator (und e+) and conventional (conv e+) positron sources.

Linear Collider Configuration	**% down, incl. 10% contingency**	**% Sched MD**	**% time Integrating Lum**
Warm, initial, und e+	38	7	55
Warm improved, und e+	25	10	65
Warm, imprvd, conv e+	21	1	78
Cold, initial, und e+	42	9	49
Cold, improved, und e+	26	10	64
Cold, imprvd, conv e+	22	4	74

The positron source configuration had a significant impact on performance. A conventional source was predicted to have 20% more integrated luminosity in steady-state operation. The impact was even greater during early commissioning, which was approximated by making all MTBFs a factor of 2 shorter, and both MD and recovery time a factor of 2 longer. Under this scenario, the time spent actually integrating luminosity decreased from 65% to 21% with an undulator and from 75% to 55% with a conventional source. Not only is the integrated luminosity more than a factor of 2 larger with the conventional source, the predicted uptime with the undulator source is so low that the collider is effectively never on. This difference is not due to any inherent unreliability of the source itself, which was not modeled in detail in either case. It is due entirely to the need for high energy electrons in order to produce positrons. With completely independent e^+ and e^- systems, MD can be performed in parallel on both systems, recovery is faster because both systems can be tuned up in parallel, and MD is possible on the e^+ system while the e^- system is down. A low-power alternate e^+ source would reduce the recovery time somewhat, but have limited impact on MD. It would be preferable for a linear collider to start operation with a conventional e^+ source and switch to an undulator source later when polarized positrons were needed.

Other sensitivity studies were performed to understand how dependent the results were on particular choices of parameters. If either the repair times (MTTR) or recovery times were reduced by a factor of two, the simulated downtime decreased by 5% (out of 15%), indicating that efforts to improve these are well worth pursuing. If the MTBF of any individual component was a factor of 10 worse than specified, the downtime typically increased by only 1-3%, indicating that missing one of these targets is not a disaster. The number of repair personnel needed inside and outside of the accelerator tunnel was not large for the 2-tunnel designs, 4-6 for each area. For the single tunnel design, these numbers were much larger, 25 in the tunnel plus 4 outside.

CONCLUSIONS

The simulation program developed as part of the US Linear Collider Technology Options Study provides a new flexible tool for evaluating accelerator reliability. Ideally it should be benchmarked against an operating accelerator, but differences in accounting procedures for maintenance and recovery make this difficult. Even without benchmarking, the simulation can be applied very effectively in comparative studies of different implementation options. The simulation can easily be applied to other facilities and a group has already started using it to study design alternatives for the Radioactive Isotope Accelerator, RIA. In the future, the simulation will be used to characterize other parts of the Linear Collider and develop cost-optimized solutions for achieving the desired availability.

REFERENCES

[1] US Linear Collider Technology Options Study, http://www.slac.stanford.edu/xorg/accelops/Full/LC_opts_full.pdf, Mar 2004.

[2] Design Considerations for an International Linear Collider, http://www.slac.stanford.edu/~hll/USLCSG/BidToHost/MachineScopeA30323.pdf, Mar 2003.