RELIABILITY ISSUES FOR LINEAR COLLIDERS*

N. Phinney#, C. Adolphsen, M.C. Ross, SLAC, Stanford, CA 94309, USA

Abstract

To deliver a high integrated luminosity over several years of operation, a linear collider must not only meet its energy and luminosity performance goals, but also have a very high hardware availability and operating efficiency. The first challenge is the size and complexity of the facility. If the typical reliability of existing High Energy Physics accelerators is simply scaled to the size of a 500 GeV linear collider, the overall system availability will be too low. The final design must incorporate a more rigorous failure analysis as well as built-in overheads and redundancy. An additional challenge is the complexity of the tuning procedures required to preserve a very small beam emittance. These include beam-based alignment of magnets and rf structures, automated trajectory correction, feedback, emittance and luminosity optimization, and more. Another issue is the inherently large power densities in the beams, which can damage any beamline components they intercept. An extensive machine protection system is necessary to inhibit beam in case of a fault and automatically execute a recovery sequence. This paper will present the important issues in the context of the proposed linear collider designs.

HISTORICAL PERSPECTIVE

The early generations of accelerators were high energy physics machines which were technically innovative. Their primary emphasis was on achieving breakthroughs in energy and luminosity, usually under tight cost constraints. Given the overhead of fills and ramping for storage rings, the luminosity uptimes achieved were in the range of 50%.

The relative importance placed on reliability has evolved with the advent of accelerator user facilities such as the synchrotron light sources, and with the new generation of high energy physics 'factories'. The large energy-frontier colliders such as the Tevatron at FNAL, HERA at DESY, LEP at CERN, and SLC at SLAC have achieved hardware availabilities in the range of 70-90%. In contrast, the B-factories at SLAC and KEK have closer to 95% availability for the colliders themselves. Synchrotron light or spallation sources have invested significant effort into improving reliability and now reach 98-99.5% [1].

While it is true that these facilities are often smaller than the energy-frontier machines, and in some respects less demanding as to performance, the reliability achieved does not appear to scale with the size of the complex. Rather, it appears that the user facilities and factories have higher standards for acceptable availability and therefore allocate the necessary resources to reach that target level.

*Work supported by the U.S. Dept. of Energy under contract DE-AC03-76SF00515.
#nan@slac.stanford.edu

AVAILABILITY GOALS

A reasonable goal for a future linear collider would be to have a hardware availability of 80-85%. Hardware downtime should include unscheduled repairs (something critical breaks), scheduled repairs (either at regular intervals or when enough problems have accumulated), and all associated cooldown, warmup and recovery times. Typically in the past, only the light sources have included maintenance periods in their downtime accounting, but this is really appropriate for all facilities. Modern accelerators do not require routine 'preventative' maintenance and interventions are only 'scheduled' when there is broken hardware. Hence, they take away from the overall beam time that might otherwise be delivered. Note that each maintenance intervention takes on the order of 3 shifts, including edge effects and recovery. A 'day' every 3 weeks represents already a 5% hit.

The overall operating efficiency or beam availability is typically significantly smaller than the hardware uptime. The integrated luminosity delivered is closer to half of what might be expected from the peak rate, even for the high performance 'factories'. Beam inefficiencies include Machine Development (time spent studying and improving the accelerator), the impact of tuning procedures, injection and the luminosity decay during a store (for storage rings), Machine Protection trips and recovery (for linacs), and last but not least, the simple fact that accelerators do not manage to deliver the same luminosity every pulse or every store. A reasonable goal for a linear collider would be a beam efficiency of 75-80%, which would produce a delivered luminosity equal to ~65% of peak performance.

Achieving this availability goal will be a challenging task for a facility the size of a linear collider, but it is necessary in order to integrate significant luminosity. Experience with the SLC and more recently with recommissioning the upgraded Tevatron and HERA has shown that poor reliability can impact the peak luminosity achievable as well as the integrated performance. If the hardware interruptions are too frequent, the machine is not up long enough to effectively make progress on the luminosity issues. It was only after the SLC achieved reasonable reliability that the many beam tuning challenges for a linear collider could be addressed. The more complex next generation of colliders must be designed from the start for high availability so that the inevitable new problems can be overcome rapidly and effectively.

SIGNIFICANT ISSUES

Several aspects of a linear collider make achieving high reliability particularly challenging. First is the sheer size of the facility and the number of components which must be functional if the machine is to operate. If the typical reliability of existing HEP accelerators is simply scaled to the size of a 500 GeV linear collider, then the resulting uptime will be unacceptably low. Fortunately, this is amenable to engineering solutions. Reliability with large numbers of components was studied extensively for the Superconducting Super Collider project in Texas and more recently for projects such as the Large Hadron Collider in Switzerland and the Spallation Neutron Source in Tennessee. Reliability must be addressed up front by failure analysis, and appropriate remedies must be implemented. Adequate engineering margins are essential if components are to perform reliably in the long term. The key issue is whether sufficient engineering and financial resources are actually allocated during development and construction to produce a reliable system.

Linac rf systems

The main linac rf system demands particular care because of the large number of components with relatively short lifetimes. Table 1 lists the component counts as given in the 2003 Technical Review Committee Report [2]. The klystrons (and some modulator components) must be replaced frequently and are considered a consumable expense. In addition, the modulators, klystrons, distribution system, and structures or cavities will experience brief faults or breakdown events where the hardware can be reset and continue operation after an appropriate timeout. Because each unit contributes only a small fraction of the total energy, a fault will typically not interrupt operation, but simply cause that pulse to be slightly low in energy. All linear collider designs plan to include spare rf units which can be switched in when a unit faults or needs repair. Critical issues are the frequency and impact of faults, the adequacy of the spares overhead, and the accessibility and duration of repairs.

	TESLA	**JLC-C**	**JLC-X/NLC**
Modulators	572	2138	508
Klystrons	572	4276	4064
Rf Distribution	n/a	2138	2032
Structures/ Cavities	20592	6784	12192
Spare rf units	2%	5%	5%

Table 1: Main linac rf components required for 500 GeV center-of-mass (both linacs)

In the linear collider designs based on warm rf technology, the klystrons and modulators are installed in a separate support housing where they are accessible for repair while the collider is delivering luminosity. Since they can be replaced more or less continuously, the number of spares required is determined by estimated fluctuations in the failure rate. In the present JLC-X/NLC designs, 5% overhead has been allocated to cover both faults and failures. The design based on superconducting rf technology described in the TESLA Design Report [3] has a single tunnel. The modulators are installed in support housings but the klystrons, transformers, and high-power pulsed cables are in the tunnel with the accelerator and can only be repaired during a shutdown. The stated goal is to have a maintenance intervention no more often than every three weeks. This would be difficult to achieve without substantially more overhead than the allocated 2% spares.

Tuning procedures

Another aspect which makes a linear collider particularly challenging is in the complexity of the tuning procedures required to preserve a very small beam emittance. In all areas of the collider from the damping rings to the interaction point, the component alignment tolerances are extremely tight (micron-scale) and cannot be achieved by traditional survey techniques. All of the designs foresee extensive use of beam-based alignment. In addition, the tight tolerances make the machines very sensitive to vibration (nanometer-scale) and to slow drifts due to temperature and ground motion effects. As a result, beam-based feedback systems are mandatory, and both invasive and non-invasive retuning will be required at intervals.

Regardless of the main linac rf technology, no linear collider can be considered a static machine and tuning is required on a variety of timescales. Feedback is essential to keep the beams in collision. Without it, they would drift apart between pulses of the machine by as much as tens of nanometers at a noisy site, such as Hamburg, to a fraction of a nanometer at a quiet site, such as the LEP tunnel. TESLA plans to bring the beams into collision and optimize the positions within a single long bunch train. NLC/JLC-X use pulse-to-pulse feedback at 100-120 Hz to damp motion at frequencies below about 10 Hz. Trajectory feedback is required to keep the beams centered in the strong final focus sextupoles or the luminosity degrades within minutes. Trajectory feedback is required elsewhere to damp transients and correct slow drifts. Energy feedback must compensate for fluctuations in the total linac energy due to rf faults as well as to a variety of rf phase or amplitude errors. Re-steering of the main linacs and damping rings will be needed on the time scale of hours and dispersion correction of the rings on the time scale of days.

Alignment tolerances

The alignment tolerances differ for the two technologies, as do the methods forseen to correct errors. The quadrupole and cavity tolerances are 10 and 100 times looser for the superconducting main linacs, but the X-band linac will have high precision position monitors on both structures and magnets, and movable stages on each magnet or girder to effect the required alignment. In the damping rings, the situation is reversed. The X-band damping rings are similar to third generation light sources and have tolerances which are no more than a factor of 3 tighter than what has already been achieved [4]. The damping rings for TESLA will have tolerances which are another 3-10 times tighter, and these must be maintained over a 17 km circumference. Meeting these tolerances may require movable stages for precision alignment and/or more rigid supports than presently foreseen. Table 2 compares the tolerances for the TESLA and NLC damping ring designs with those for the recently completed Swiss Light Source (SLS) as given in Ref. 4. In the final focus, the alignment tolerances are similar for both designs but the superconducting collider is more sensitive to vibration because the low repetition rate limits the frequency to which feedback can be effective.

	SLS	NLC	TESLA
Energy [GeV]	2.4	2	5
Circumference [m]	288	300	17,000
Sext vert align [μm]	71	31	11
Quad roll align [μrad]	374	322	38
Quad vert jitter [nm]	230	75	76

Table 2: Alignment tolerances for the TESLA and NLC damping rings compared with tolerances at the Swiss Light Source (SLS)

MACHINE PROTECTION

The small, very intense, beams in a linear collider require a new approach to machine protection (MPS) untested at any existing or soon to be completed machine. The pulsed time structure of the beam, as opposed to the CW nature of storage rings like the Tevatron or LHC is an additional difficulty. A single, nominal intensity, bunch will damage almost any accelerator hardware it happens to strike downstream of the damping rings. Since it is not possible to stop a given beam bunch once extracted from the damping ring, there is little fundamental difference in the MPS exposure or design strategy for the different machines. The long inter-bunch interval in TESLA allows the beam to be switched off somewhat more quickly than in JLC/NLC. The minimum time required to turn off the beam is one full interpulse period for the JLC/NLC short train and about 1/10 of the train length (~100 μs or 300 bunches) for TESLA.

Protection system schemes have been proposed for both TESLA and JLC/NLC which appear feasible [5]. They must automatically control changes in beam power, both by halting operation when a fault is detected and by restoring operation when the fault is cleared. They rely heavily on the use of a pilot bunch and a fast permit system. The permit signal is derived from beam data taken on the previous pulse and from a system that monitors fast devices. Before operation can be resumed after a fault, the MPS must provide for the production of a sequence of pilot and low power pulses that prove the fitness of the downstream systems for high power operation.

TUNNEL CONFIGURATION

The TESLA Design Report [3] proposed a collider built within a single tunnel. This tunnel would contain the beam lines for the superconducting main linacs, damping rings, injectors, injector linacs, positron production, and beam delivery systems. The linac klystrons with pulse transformers, rf controls and high power pulsed cables, as well as many power supplies and electronics, would also be installed in the same tunnel. In contrast, the X-band rf machine would have separate tunnels for the injectors and damping rings and separate accessible support housings for klystrons, power supplies, electronics, etc. to facilitate repair during operation.

A single tunnel would require interrupting operation at frequent intervals to access the tunnel to replace failed klystrons and repair other components. Great care would be needed to ensure that all in-tunnel components had extremely high reliability. Because a single tunnel would house almost all beamlines, linac access would also impact the rings and injectors. The single tunnel also limits flexibility in initial commissioning. All of these issues would need to be carefully assessed with regard to reliability and efficiency. The single tunnel choice was driven by cost considerations and constraints of the DESY site, but could well be reconsidered for a superconducting linear collider built elsewhere.

CONCLUSIONS

To deliver the integrated luminosity demanded by the physics goals, a linear collider will need to be designed for very high hardware availability and beam efficiency. Nominal goals of 80-85% for hardware availability and 75-80% for beam efficiency will not be achieved without considerably more effort than has often been devoted in the past. A robust design requires rigorous failure analysis, generous built-in overheads and redundancy for critical components. Complex tuning procedures will demand an unprecendented level of automation. Overall these goals should be achievable, but only if sufficient attention and resources are allocated from the earliest design stage through commissioning and operation.

REFERENCES

[1] L. Hardy, “Accelerator Reliability - Availability”, Proc. of EPAC02 (Paris, France, June 3-7 2002)

[2] International Linear Collider Technical Review Committee Second Report 2003, SLAC-R-606, January 2003.

[3] R. Brinkmann, et al., (eds.), TESLA Technical Design Report Part-II, DESY-2001-011, March 2001.

[4] T. Raubenheimer, A. Wolski, "Comparison of Alignment Tolerances in the Linear Collider Damping Rings with those in Operating Rings," LCC-112, January 2003.

[5] C. Adolphsen, et al., “The Next Linear Collider Machine Protection System,” SLAC-PUB-8130, Proc. of PAC99, new York, NY, 29 Mar-2 Apr 1999.