Engineering in High Availability

Tom Himel

Contents

- Why High Availability is important
- General HA issues
- Controls specific HA issues

Why High Availability is important

- The ILC will be an order of magnitude more complex than most present accelerators.

- If it is built like present HEP accelerators, it will be down an order of magnitude more.
- That is, it will always be down.
- The integrated luminosity will be zero.
- Not good.

Availability Design Philosophy

- Design it in up front.
- Budget 15% downtime total. Keep an extra 10% as contingency.
- Try to get the high availability for the minimum cost.
- First stab was done for the RDR.
- Will need to iterate during the ED phase.
  - **Quantities are not final**
  - **Engineering studies may show that the cost minimum would be attained by moving some of the unavailability budget from one item to another.**
  - **This means some MTBFs may be allowed to go down, but others will have to go up.**

Will Need Improvement Program

- Must design to meet the budget on the first pass.
- Assume we are only partly successful and unavailability will be too high when we turn on.
- Will need operations budget and engineering to make the necessary improvements.

Top level approach

- Use Availsim a Monte Carlo simulation developed over several years.
- Given a component list and MTBFs and MTTRs and degradations it simulates the running and repairing of an accelerator.
- It can be used as a tool to compare designs and set requirements on redundancies and MTBFs.

MTBF goals

Device	Needed Improvement factor	Downtime (%) due to these devices	Nominal MTBF (hours)	Nominal MTTR (hours)
power supplies	**20**	0.2	50,000	2
power supply controllers	**10**	0.6	100,000	1
flow switches	**10**	0.5	250,000	1
water instrumention near pump	**10**	0.2	30,000	2
magnets - water cooled	**6**	0.4	3,000,000	8
kicker pulser	**5**	0.3	100,000	2
coupler interlock sensors	**5**	0.2	1,000,000	1
collimators and beam stoppers	**5**	0.3	100,000	8
all electronics modules	**3**	1.0	100,000	1
AC breakers < 500 kW		0.8	360,000	2
vacuum valve controllers		1.1	190,000	2
regional MPS system		1.1	5,000	1
power supply - corrector		0.9	400,000	1
vacuum valves		0.8	1,000,000	4
water pumps		0.4	120,000	4
modulator		0.4	50,000	4

Global availability work

- Will need to improve unavailability budget.
  - **Follow design changes as they occur (e.g. device counts and located in tunnel or not)**
  - **Re-apportion the allowed unavailability as we find cheaper ways to attain the budgeted 15% downtime. The present HA budget is just a guess at the cost optimum.**
- Provide a framework (FMEA package) for groups to use to evaluate their system’s availability.
- Push groups to do HA work that is being ignored (e.g. water instrumentation, collimators, and coupler interlocks).

Criteria for doing HA design during ED phase rather than waiting

- Decisions which couple multiple systems
- Designs which need a large availability enhancement compared to present designs (e.g. power supplies and water cooled magnets)
  - **These need to be prototyped and tested**
- Devices which will be produced in very large quantities (e.g. LLRF electronics, controls crates) or are difficult to retrofit (e.g. cryomodules, klystrons).

Example trade studies during ED Phase

- Redundancy of cavity sensors and interlock electronics vs. extremely high availability or extra energy overhead
- Design of vacuum manifolds, valves, pumps, pump supplies for couplers to be slightly redundant vs. extra energy overhead
- Use of waveguide switches and hot spare klystrons and modulators in large fractional energy gain regions vs. some other way of saving that ~1% downtime
- Locating extremely HA magnet power supplies in the beam tunnel vs. HA supplies in alcoves with long cables and the consequent heat load
- Do we have to do 1 vs. 2 tunnel yet again?
- Large redundant water pumps vs. small segmented pumps that don't kill the accelerator when one fails (e.g. for a single klystron).

Example trade studies during ED Phase

- Placing some LLRF (maybe just down mixers) in tunnel to reduce the cable plant while making them extremely HA or increasing the energy overhead
- Fast e+ target replacement vs. extra e+ target region beam-line.
- What temperature should the electronics be cooled to?
- Is the 0.5% downtime allocated to AC power source and distribution enough? Is there some way we can prevent 0.25 second power dips from causing 8 hours of downtime? Should we get power from two places on the grid?
- Use of Uninterruptible Power Supplies vs. more extra redundancy in the AC power distribution.
- Is the allocation of 1% downtime for cryo problems enough (10 times better than CERN does)? How will it be achieved?

Example trade studies during ED Phase

- What are the right operating margins for cryo, magnet supplies, RF, utilities to optimize cost and availability?
- Is it better to add RF units or make each more reliable?
- Should there be redundant beam instrumentation for critical feedback loops?
- How much electronics should be put in the tunnel?

Example Trade-off studies that can probably wait

- Should individual channels of electronics (e.g. timing generator, RF amplitude and phase detector) be hot swappable or is OK to put many of them on a single board?
- Detailed designs of items that do not need major availability improvements e.g. beam loss monitor readouts, laser wire electronics, ADC boards, magnet supports

Two aspects to controls and availability

- Controls itself should not go down often. (Covered in talk by Krause)
- Controls needs to provide tools to help discover what is wrong in other systems.

Tools to help other systems

- Network access for laptops and diagnostic equipment near all hardware
- Readout and recording of diagnostic information built into other systems (e.g. a power supply may record its voltage and current at a megahertz)
- Either record everything very often or allow flexibly triggered readout of everything or both.
- Provide analysis tools of the data that is recorded.

Example of need for sync readout Based on SLC “flyer pulses”

- Infrequently a single bunch causes very high backgrounds. Need to figure out why.
- Only know few seconds after the fact that a bunch was bad.
- Could be caused by bad kicker pulse. Need to know kicker strength on each bunch.
- Could be caused by DR phase instability (saw tooth). Need to know orbit and phase of that bunch on many turns prior to extraction

Help Solve Subtle Problems

- Phase drifts – compare redundant readouts
- Lying BPMs – chisquared? Redundancy?
- Drifting BPMs (both mechanical and electrical)
- Difficult to localize problems (normal module swaps don't fix it). E.g. noise coupling in on a long cable or a flakey connector.
- Vacuum bursts in DRs (present PEP problem) – read 1/sec, provide good analysis package

Conclusions

- Designing for High Availability is vital for all systems.
- A good fraction of it must be done during the ED phase.
- Controls has the added responsibility of providing the tools to help other systems diagnose their problems quickly.

Backup slides

The Simulation includes:

1. Effects of redundancy such as 21 DR kickers where only 20 are needed or the 3% energy overhead in the main linac
2. Some repairs require accelerator tunnel access, others can't be made without killing the beam and others can be done hot.
3. Time for radiation to cool down before accessing the tunnel
4. Time to lock up the tunnel and turn on and standardize power supplies
5. Recovery time after a down time is proportional to the length of time a part of the accelerator has had no beam. Recovery starts at the injectors and proceeds downstream.
6. Manpower to make repairs can be limited.

The Simulation includes:

7. Opportunistic Machine Development (MD) is done when part of the LC is down but beam is available elsewhere for more than 2 hours.
8. MD is scheduled to reach a goal of 1 - 2% in each region of the LC.
9. All regions are modeled in detail down to the level of magnets, power supplies, power supply controllers, vacuum valves, BPMs …
10. The cryoplants and AC power distribution are not modelled in detail.
11. Non-hot maintenance is only done when the LC is broken. Extra non-essential repairs are done at that time though. Repairs that give the most bang for the buck are done first.

The Simulation includes:

12. PPS zones are handled properly e.g. can access linac when beam is in the DR. It assumes there is a tuneup dump at the end of each region.
13. Kludge repairs can be done to ameliorate a problem that otherwise would take too long to repair. Examples: Tune around a bad quad in the cold linac or a bad quad trim in either damping ring or disconnect the input to a cold power coupler that is breaking down.
14. During the long (3 month) shutdown, all devices with long MTTR's get repaired.