## Power Failure Post-Mortem Meeting June 24, 2013

The tech staff met on June 24 to discuss the perfect storm of power, generator, and cooling failures that occurred on Thursday June 13, 2013.

### Synopsis

The Liebert was already operating at half capacity due to a failed compressor. The compressor was on order and is being replaced today. Fans were operating at both doors to use the house AC to help cool the machine room.

Power failed at the University and surrounding area on Thursday afternoon, June 13, 2013. All 3-phases were out to 3904 and all of CERI was without power. Both generators successfully started and the machine room was up and running.

The machine room was getting warm because the Liebert was only operating with one of two compressors. Steve B. turned off non-seismic network and non-critical loads to maintain temperature.

Steve B. came back for the estimated power restoration at 6:30 and he turned enigma and bofur back on because Holly was unable to ssh.

Steve had to come back later in the evening after a mudworm high temp alarm. He was joined by Jim and Dave to help with the generator. Power had failed again and the machine room was getting warm even with just the addition of bofur and enigma. The generator had failed. Power was going up and down.

Eventually, stable power was restored that evening (around 8pm?).

### Primary failure modes

Liebert was already at half capacity.

Generator initially worked but then failed, most likely from a weak part stressed further by multiple and prolonged power cycles. The fuel lift pump and electronic controller (what controller?) were replaced. Previous UM maintenance contract did not include deep 4-hour load test but it does now.

One ups failed -- not batteries.

### Secondary failures

AQMS does not auto switch from shadow to primary. This is inherent in the system and its intended to be operator initiated.

Smeagol logins depend on /usr/local mounted from enigma. That prevented Holly from getting ssh to process events

Deagol went down (in 3892 and not on generator) and request tracker did not come back clean.

Priam (AQMS primary) was unresponsive

Paris (AQMS shadow) was operational (Mitch switched roles to make primary)

Helen (AQMS postproc) had a raid failure.

Shakeworm got shutdown causing recenteqs to halt while processing. It left a file that prevented updates once mainworm was rebooted. Mitch removed the file and restarted.

Machine room was dark and difficult to navigate.

The transfer switch was not energized and not able to provide trouble conditions.

Procedures for emergency generator service are unclear.

**Remedial Actions**

A flashlight is now hanging next to the transfer switch.

A generator key is now hanging next to the transfer switch.

Emergency contacts (including generator after hours procedures) are posted next to the transfer switch.

The generator is repaired.

The Liebert is repaired.

The AQMS machines are repaired and we now know that we need to manually change roles.

Should we have backup cooling? (Mitch, Steve, and Jim).

Should we have alternate external Internet?

Backup to the generator is not feasible.

If Ricky Bibb is not available to contact Cummins, is Mitch authorized to allow emergency repair charges against the seismic networks? (Mitch, Michelle)

Remove smeagol dependencies on enigma (Steve and Bob).

Disentangle machine room spaghetti (Steve and Bob) -- a long-term goal.

Get colored adhesive dots to identify critical (red), non-critical (green), and in between servers (yellow) to help identify which servers may be powered off.

When AQMS postproc is in production, will we need a postproc hot backup (and dual archdb)?

**Other Observations**

This was a confluence of multiple, essentially unrelated failures.  Staff involved did an excellent job of maintaining as much functionality as possible while dealing with a challenging situation.  Short of having dual cooling systems and dual generators, we're not likely to be able to prevent a similar failure in the future.  The probability of a similar set of multiple failures is small.

NEIC is our backup if we are rendered inoperable.  NEIC has dual cooling and power systems each of which operates at 40% so that either one is capable of carrying the full load.  That's currently beyond our budget and we have many other priorities that compete for resources and are equally necessary for robust operation.