

---

# Primary Node Failure and Recovery

## *MetaArchive Project, Extension Phase*

2008-02-08



### Summary

This document records the findings of the first of three network tests that the MetaArchive Cooperative will perform during the extension phase of its contract work with the Library of Congress (Sept. 2007 – May 2009). It reports on the actual failure of the primary node (at Emory University) which occurred on 29 Nov. 2007. Although this incident was an actual failure (not simply a test), it also provided an excellent test of the MetaArchive's Southern Digital Culture network and the LOCKSS-based recovery system that it employs.

This report includes both an incident report and troubleshooting recommendations that emerged from the successful rebuild of the primary node.

### Incident Report

#### ***First Power Failure***

On the night of 29 November 2007, the second floor server room at the Woodruff Library of Emory University experienced a power failure affecting both of the uninterruptible power supplies that service [ndiip.library.emory.edu](http://ndiip.library.emory.edu), the primary node of the MetaArchive Cooperative's LOCKSS network. The power failure also affected the AX100 storage array that is attached to the server. This array stores all of the data for the archival units managed by the server.

Working with the administrator of the other servers affected by the outage on 30 Nov., we found that the power outlets supplying the UPSs had failed. We also discovered that the UPSs had been plugged two outlets on a single circuit, rather than into separate circuits. Without physical access to the circuit breakers for the outlets, our only option to restore the most critical services immediately was to move the UPSs to outlets on our one remaining unoccupied power circuit.

We placed a work order for the failed circuit to be investigated and significantly reduced the load on the UPSs, disabling five servers and one storage array which were not critical to the continuance of business. After a few hours of monitoring, we concluded that the power supply to the UPSs would be sufficient until the failed circuit could be repaired.

While waiting for facilities maintenance to repair the failed circuit, we brought [ndiip.library.emory.edu](http://ndiip.library.emory.edu) and its storage array back online to examine their status. The storage array is divided into six separate virtual disks, and a cursory filesystem check revealed that three of those six had experienced some corruption. We initiated repairs on the damaged filesystems, a process that takes approximately 8 to 12 hours for filesystems of this size on such hardware. We left the repairs to run over the weekend.

The next week, the filesystems continued to show damage after the first round of repairs were complete, and we initiated a second round of repairs using a more detailed examination of the media, expected to take a few days to complete.

#### ***Second Power Failure***

While the filesystem repairs ran, the circuit we had selected to temporarily power both UPSs failed. Because we no longer had any circuits remaining to power our critical systems, we escalated our call to facilities maintenance, who restored power to both circuits by that afternoon. Once power was restored, we plugged each UPS into a separate circuit.

Once we had brought [ndiip.library.emory.edu](http://ndiip.library.emory.edu) and its storage array back online, the storage array hardware reported, in rapid succession, the failure of two hard disks out of the twelve that compose the array. Because the

array is configured using RAID Level 5, it can only withstand the loss of a single disk before irreparable corruption occurs. The second report of hard disk failure appears to have been a malfunction of the storage array.

Because a simultaneous failure of two disks meant we would have to rebuild the entire array from scratch, we shut down and powered up the storage array again to test whether or not the second disk failure was a malfunction. We allowed the controller hardware to determine again whether the disks had failed. Fortunately, once it was online, it reported the failure of only a single disk. We decided to replace this disk with a new one, and obtained and installed a replacement disk on 6 December 2007. The storage array accepted the new disk and began to rebuild its RAID, copying data onto the replacement disk. The rebuild continued through 7 December 2007.

Although the RAID was successfully rebuilt, the filesystems of the three affected virtual disks were so badly damaged that we had no option but to reformat them and re-crawl the corresponding archival units to recover the data. On the advice of Seth Morabito of the LOCKSS team at Stanford University, we applied configuration changes to our server on 8 December 2007 to aggressively rebuild the missing data. This process completed in only a few hours and we subsequently reverted the configuration changes.

### ***Present Status***

The disk has been in continuous service since 8 December 2007 without any report of failure from the array.

### **Troubleshooting Recommendations**

The corruption of filesystems housing archival units or the failure of the storage media containing those filesystems is among the most likely recoverable failures of any LOCKSS node.

Although our attempts to recover the filesystem were delayed by a second power failure, the process consumed a great deal of time, and these filesystems were likely beyond recovery even before the second failure occurred. Checking and repairing three large ext3 filesystems on AX100 hardware took far more time than our later solution of simply reformatting the disks and re-crawling the data from remote sources.

It is obvious that after severe catastrophes, in which hardware or hosting facilities are wholly destroyed, the best course of action would be to recover archival units from remote sources; this is a fundamental design principle of LOCKSS.

An unanticipated lesson of these incidents was that even less significant hardware or filesystem failures are sometimes more readily recovered from remote sources. Although the LOCKSS system provides for the automatic detection and repair of corrupted archival units, we recommend recovering damaged filesystems from other machines in a network rather than by attempting filesystem repairs, as it is likely to result in less node downtime.

### **Notes on this document**

This document was prepared by Chris Roddy on 2008-02-01 and was completed at the second MetaArchive All-Project Meeting held at University of Louisville on Friday 2008-02-08, by the Steering Committee and additional attending project participants.