The NetarchiveSuite Archive

- and why it *isn't* a secure archive (yet)

Access for researchers

- Researchers are showing increasing interest for our archive. In our original design we envisioned two ways to access the archive.
  - Browsing the archive in a browser
  - Running batch jobs on the archive
  - *But...*

Batch job security

- If researchers can ru nany batch job, how do we make sure it is safe?
  - We should add a security model to the archives. Batch jobs should not be allowed to
    - Edit any files
    - Read files outside the archive and temporary areas
    - Communicate with the network?
  - A security model would also protect us against dangerous bugs in our own code

Batch job possibilities

- Currently we can only run known batch jobs
- We should be able to send any batch class file to the bitarchives
  - This involves a class loader and some reflection
- It would be desirable to have a method called when batch results are merged on the bitarchive monitor
- Current batch performance is bad – mostly FTP

Archive authorisation

- Currently, the archive is protected for access at the access points, i.e. the browser and shells
- However, it is also protected at deletion by a password. This password is written in cleartext in settings, and password for both archives is known by BJA (and probably the entire dev. Group)
- We need a better solution!

The power of the code

- Currently, a developer could without much trouble sneak code into the archives, that did damage, intentional or not.
- Suggestion from IT dept at SB is that full file systems are mounted read only. Thus the archive is only accessible for updates on request. This would give a much desired distribution of responsibility.

Impractibility of bit preservation

- Our bit preservation algorithms are impractical, and we need to take steps soon.
- Running checks takes weeks!
- Update on errors is contra-intuitive, slow, and for large numbers of files extremely impractical.
- Problem addressed often before.

Solution identified by NHC, LC, BJA

- Need to be able to run automatic smaller checks
- On bit preservation it **is** desirable to know the position of trouble files
- We should consider replacing admin data with a database, accessible by bit preservation
- User interface needs to be rewritten and rethought.

GUI Mockup

- Attempted mockup on GUI is here.