GreenFS: Making Enterprise Computers Greener by Protecting Them Better

Nikolai Joukov and Josef Sipek

Presented by Carsten Weinhold

Paper Reading Group, 2008-04-15
Motivation

Observations:
- Data is irreplaceable asset of high value
- Disks are fragile
- Disks consume energy
- Disks are noisy

Current solution: **Spin up / down**
- Disks wear out:
  - 50,000 cycles for desktop disks
  - 600,000 cycles for laptop disks
- Ineffective against unexpected shocks
- User inconvenience
GreenFS Design

Key idea: **minimize number of active disks**
- Reduce shock susceptibility
- Reduce energy consumption
- Use flash memory to hold working set (+ version history)

Design goals:
- Usable for laptops, desktops, servers
- Designed for enterprise environment:
  - Centralized servers
  - High-bandwidth network
- Support disconnected operation
GreenFS Architecture

GreenFS-client

Versioning

Flash

... 

GreenFS-server

Redundancy Elimination

Figure copied from GreenFS paper.
Buffered All-Time Protection

- Run-time backup (Continuous Data Protection)
- All updates sent to remote server
- Remote server may keep versions
- Flash memory used as cache:
  - Recently used files
  - Version history of updates, if disconnected
- Flash memory + remote server store complete history
Reversed roles of backup server and local disk:

• Remote server:
  – Receives all data updates
  – Keeps version history

• Local disk:
  – Inactive most of time
  – Keeps backup of data on server

• Local disk synchronized on:
  – Shutdown, regular intervals
  – Memory pressure (disconnected operation)
  – Large large writes
All data of all machines stored on remote server.

- Energy efficient, because server is shared
- How to reduce costs for storage?
  - Redundant data elimination
  - Hierarchical backup
  - Not deeply discussed in paper
Reliability

• Connected Operation:
  – All data stored redundantly on remote server
  – All updates sent immediately
  – Updates may / may not be written to disk
  – Crash: local disk gets synchronized on boot

• Disconnected Operation:
  – Local disk spins up on demand
  – Servers / desktops: keep disk spinning
  – Laptops: minimize time disk is spinning

• Flash memory is robust
Authors claim: *power efficiency, reliability and user convenience are not* mutually exclusive!

- Local disk in standby:
  - Consumes little power
  - Totally silent
  - Shock resilient

- Disk spins up for constant reads / writes

- Break even: power for disk / network
GreenFS Implementation

Figure copied from GreenFS paper.
Application Launch Benchmark

Figure copied from GreenFS paper.
OpenSSH Compile Benchmark

Elapsed Time (seconds)

- Wait
- User
- System

Figure copied from GreenFS paper.
### Power Efficiency

**Estimation of power savings:**

<table>
<thead>
<tr>
<th>System</th>
<th>Original Power (W)</th>
<th>GreenFS Power (W)</th>
<th>Savings (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLADE</td>
<td>90</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>SERVER</td>
<td>113</td>
<td>101</td>
<td>12</td>
</tr>
<tr>
<td>DESKTOP</td>
<td>54.1</td>
<td>46.8</td>
<td>14</td>
</tr>
<tr>
<td>NOTEBOOK</td>
<td>20.1</td>
<td>19.6</td>
<td>2.5</td>
</tr>
<tr>
<td>NOTEBOOK’</td>
<td>13.9</td>
<td>13.4</td>
<td>3.6</td>
</tr>
</tbody>
</table>

*Figure copied from GreenFS paper.*
Shock Exposure

Benchmark:
- Notebook carried around
- Disk is active

<table>
<thead>
<tr>
<th></th>
<th>elevator</th>
<th>stairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>trip time (sec)</td>
<td>85</td>
<td>58</td>
</tr>
<tr>
<td>network available (%)</td>
<td>94</td>
<td>100</td>
</tr>
<tr>
<td>disk is shock protected APS (%)</td>
<td>29</td>
<td>63</td>
</tr>
<tr>
<td>disk is shock protected GreenFS (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>disk stop operations APS</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>disk stop operations GreenFS</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(Samsung HM500LI: shocks up to 750G when inactive, up to 325G when spinning)

Figure copied from GreenFS paper.
Summary

- Modular design
- Stackable file system
- COTS hardware
- Base on existing components:
  - UNIONFS
  - Cryptfs, VersionFS, Ext3cow, …
  - NFS
Discussion Points

• Interesting solution to important problem ... to become obsolete with SSDs?

• Evaluation:
  – Size of flash memory? What if too small?
  – Power consumption of laptop USB ports?
  – Redundant Data Elimination?

• Power saving & laptop disks

• Windows Vista & hybrid disks (“Ready Drive”)
References

- http://www.heise.de/newsticker/Flache-Notebookplatten-mit-500-GByte--/meldung/106339/from/rss09
Energy consumption: \[ U = U_{up} + P \times \frac{S}{B} \]

<table>
<thead>
<tr>
<th>Device</th>
<th>$U_{up}$ (J)</th>
<th>$P$ (W)</th>
<th>$B$ (MB/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server HDD</td>
<td>75</td>
<td>15</td>
<td>71</td>
</tr>
<tr>
<td>Desktop HDD</td>
<td>41</td>
<td>11</td>
<td>56</td>
</tr>
<tr>
<td>Notebook HDD</td>
<td>5</td>
<td>2.5</td>
<td>44</td>
</tr>
<tr>
<td>gigabit ethernet</td>
<td>0.12</td>
<td>1.5</td>
<td>86</td>
</tr>
<tr>
<td>100Mbps ethernet</td>
<td>0.12</td>
<td>1.2</td>
<td>11</td>
</tr>
<tr>
<td>802.11g</td>
<td>0.14</td>
<td>1.4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Typical values of start up energy ($U_{up}$), power ($P$), and bandwidth ($B$) while transferring the data.