TECHNISCHE UNIVERSITÄT DRESDEN

Department of Computer Science, Institute for System Architecture, Operating Systems Group

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

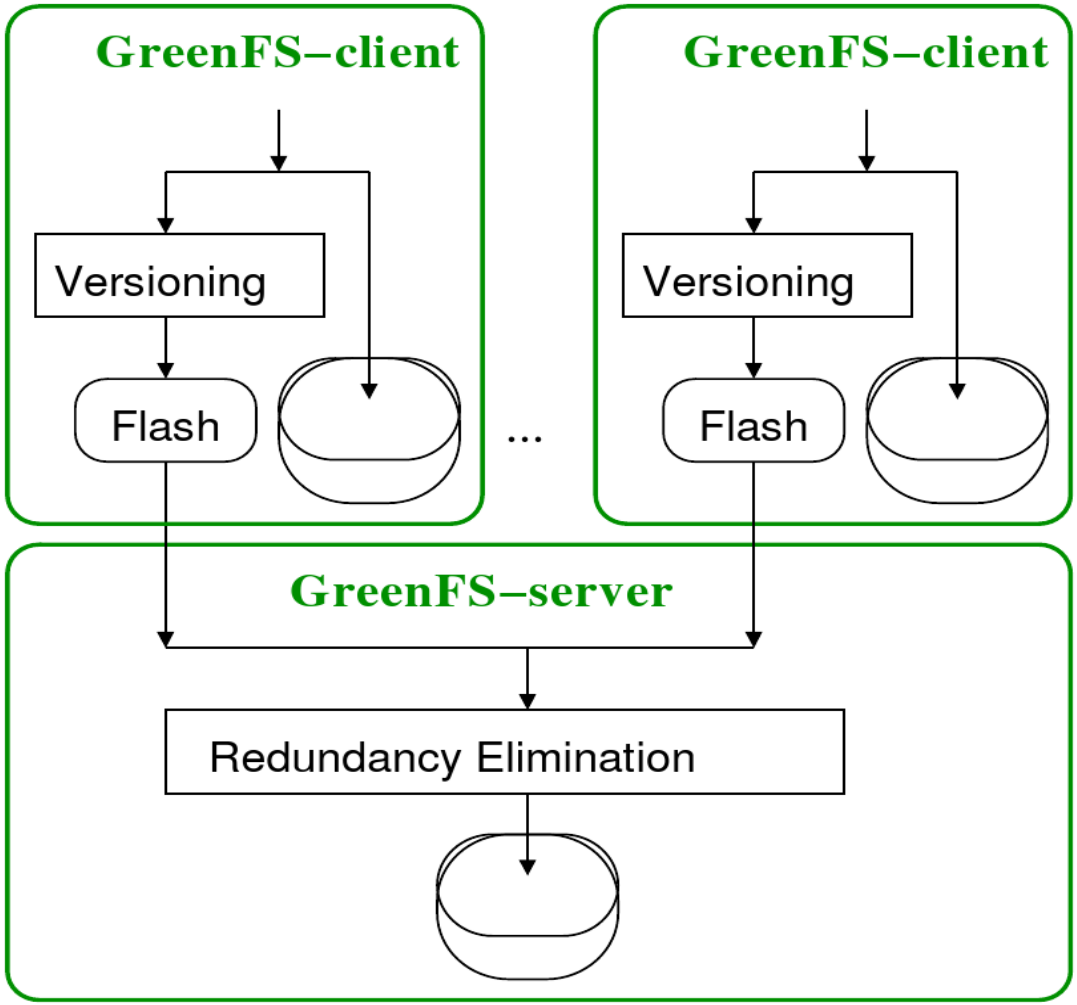

Figure copied from GreenFS paper.

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 Backup Operation

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 Protection

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

Power, Reliability & User Experience

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

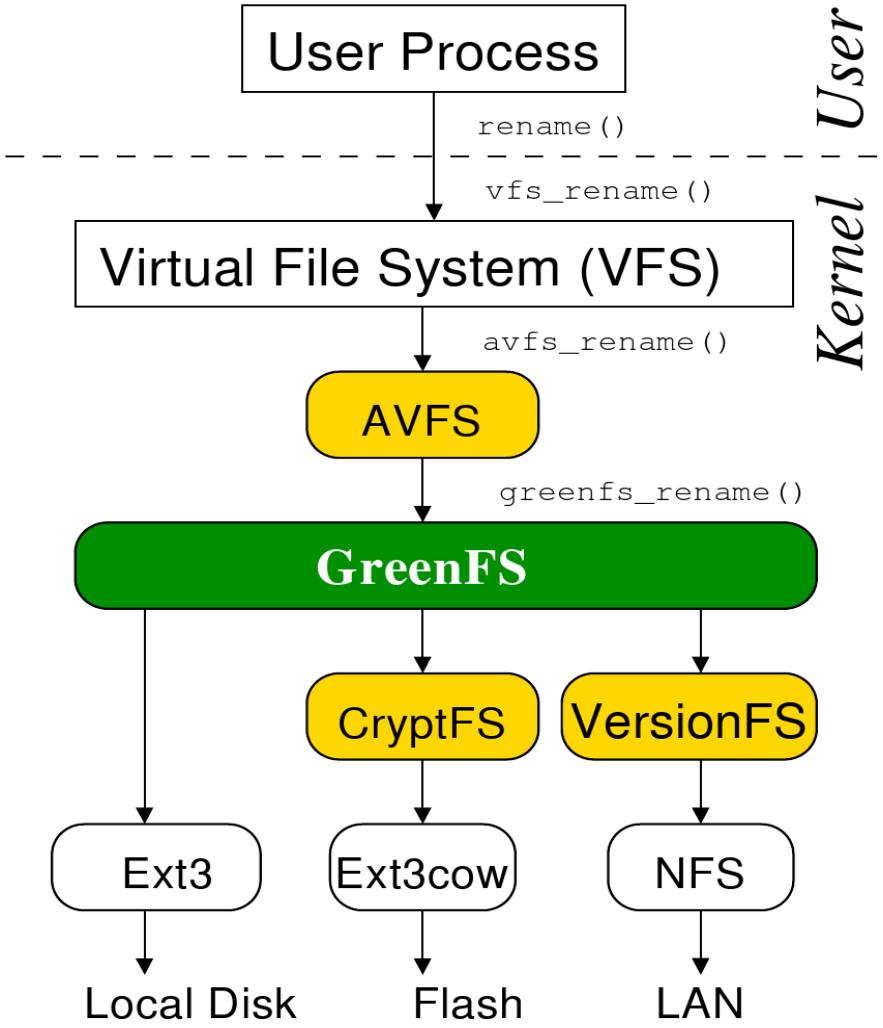

Figure copied from GreenFS paper.

Application Launch Benchmark

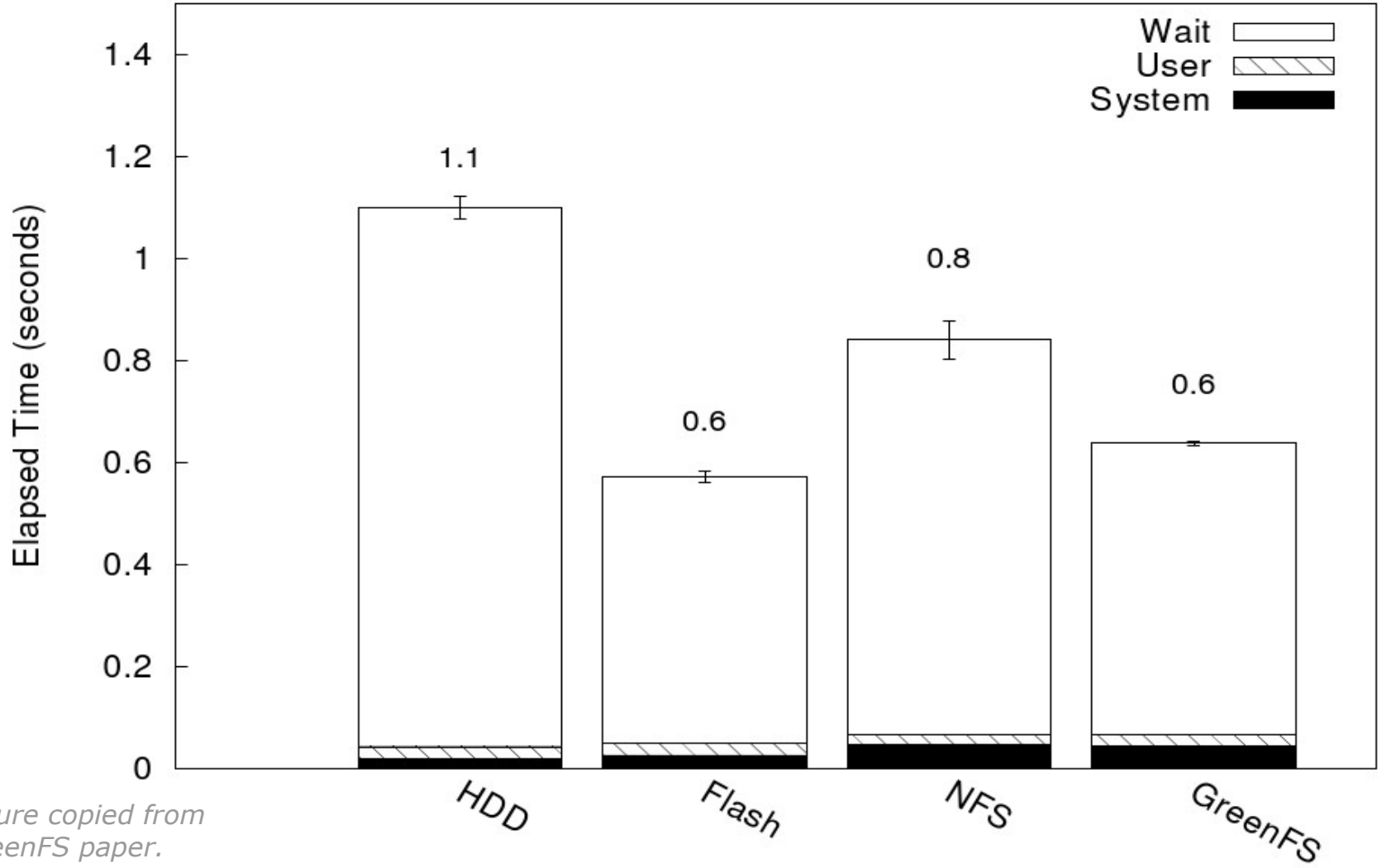

Figure copied from GreenFS paper.

OpenSSH Compile Benchmark

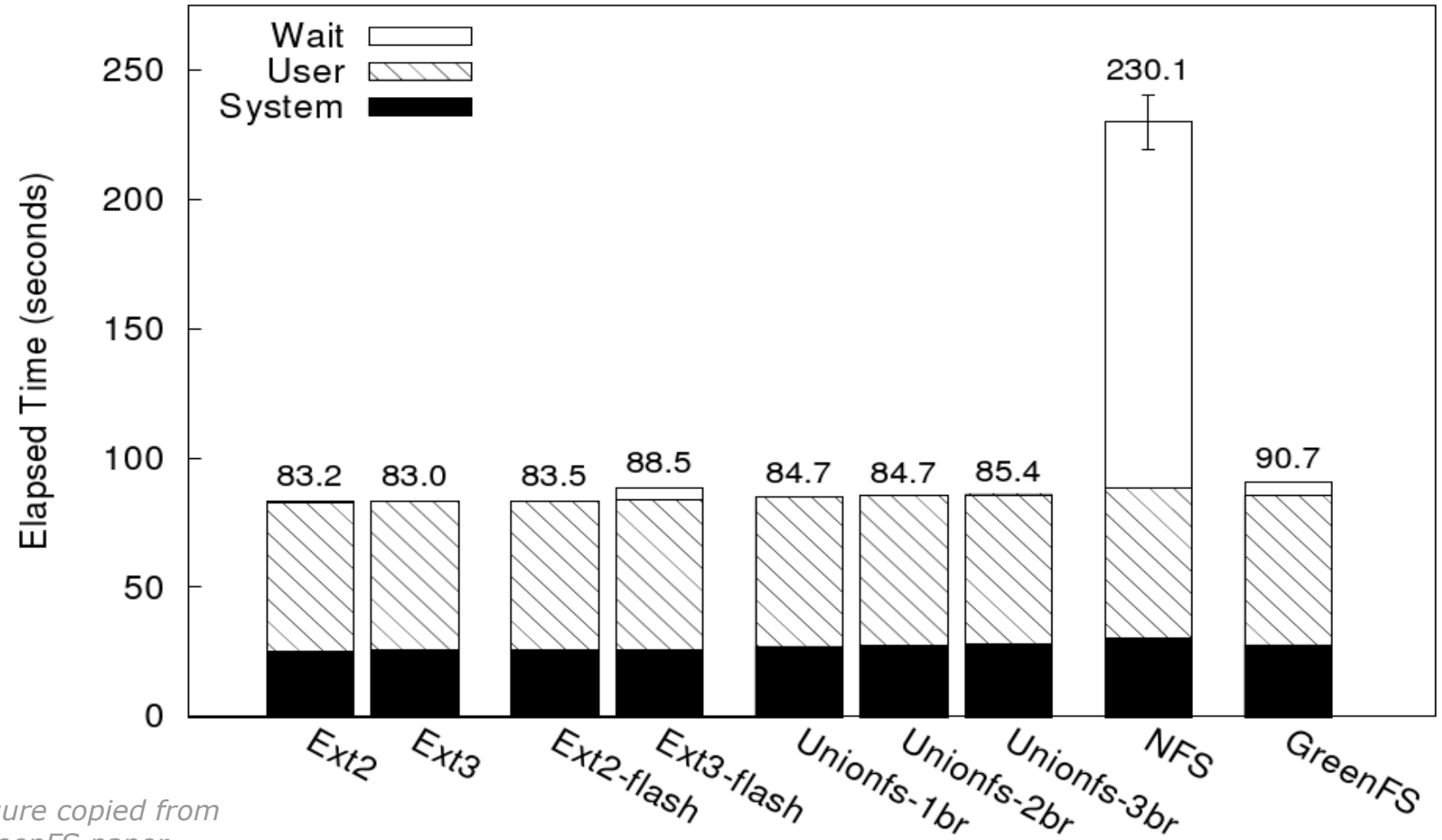

Figure copied from GreenFS paper.

Power Efficiency

Estimation of power savings:

System	Original Power (W)	GreenFS Power (W)	Savings (%)
BLADE	90	N/A	N/A
SERVER	113	101	12
DESKTOP	54.1	46.8	14
NOTEBOOK	20.1	19.6	2.5
NOTEBOOK'	13.9	13.4	3.6

Figure copied from GreenFS paper.

Shock Exposure

Benchmark:

- Notebook carried around
- Disk is active

	elevator	stairs
trip time (sec)	85	58
network available (%)	94	100
disk is shock protected APS (%)	29	63
disk is shock protected GreenFS (%)	100	100
disk stop operations APS	4	6
disk stop operations GreenFS	0	0

(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

- Nikolai Joukov and Josef Sipek, **"GreenFS: Making Enterprise Computers Greener by Protecting Them Better"**, EuroSys 2008, Glasgow
- http://www.heise.de/newsticker/Flache-Notebookplatten-mit-500-GByte--/meldung/106339/from/rss09

Power Trade Off: Disk or Network

Energy consumption: $U = U_{up} + P \times \frac{S}{B}$

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

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