USB Flash Drives as an Energy Efficient Storage Alternative

Olga Mordvinova, Julian Martin Kunkel,
Christian Baun, Thomas Ludwig and Marcel Kunze
University of Heidelberg
Karlsruhe Institute of Technology
University of Hamburg c/o DKRZ

E2GC2 | Banff | October 13, 2009

Agenda

- Motivation
- Energy Efficiency
- Our Approach
- Performance Measurements
- Evaluated Scenarios
- Conclusion

Motivation

- Increasing energy costs and the trend to green solutions are generating growing interest in eco-friendly computing
- In the storage area, flash storage technology satisfies low-energy requirements
- Falling prices for flash make replacement of conventional HDDs by SSDs economic for industry, but they are still too expensive for the low-cost server market
- USB flash drives are today the cheapest available flash storage

Energy Efficiency

Performance per Joule

- Energy efficiency (P) is defined as amount of work per joule. Work depends from the actual workload
- For I/O subsystem it is considered as amount of data or metadata (T) accessed per joule (E):

$$P = \frac{T}{E}$$

- It can also be considered as sustained throughput of a device per watt

Performance per Joule

- Due to throughput and energy consumption variance for flash storage, we assume:

$$P_{SSD} = \frac{T_{read} + T_{write}}{2E}$$

$$P_{USBFlashDrive} = \frac{1}{2}\left(\frac{T_{read}}{E_{read}} + \frac{T_{write}}{E_{write}}\right)$$

Performance per Joule - Hard Disk Drives

Model	RPM	Form Factor	Capacity [GB]	Throughput (sustained) [MB/s]	Power Consumption (transfer) [W]	Power Consumption (idle) [W]	Performance per Joule [MB/J]
Western Digital RE2 WD4000YR	7200	3.5”	400	65	10.8	8.9	6.1
Seagate ST3450856SS Cheetah 15K.6	15000	3.5”	450	140	17.3	12.4	8.1
Western Digital WD1001FALS	7200	3.5”	1000	80	8.4	7.8	9.5
Samsung HD103UI EcoGreen	5400	3.5”	1000	65	6.2	5.0	10.5
Seagate ST9250421AS Momentus	7200	2.5”	250	60	2.1	0.7	28.6
Hitachi Travelstar 5K500	5400	2.5”	500	50	1.9	0.7	26.3

Performance per Joule - Solid State Drives

Model	Type	Form Factor	Capacity [GB]	Throughput (read) [MB/s]	Throughput (write) [MB/s]	Power Consumption (transfer) [W]	Power Consumption (idle) [W]	Performance per Joule [MB/J]
Samsung MCBQE32G5-MPP	SLC	2.5”	32	55	40	0.2	0.1	237.5
Samsung MCCOE64G5-MPP	SLC	2.5”	64	90	80	0.8	0.2	106.3
Mtron MSP-SAA7535032	SLC	2.5”	32	115	110	2.4	1.6	46.9
Crucial CT64GBFAA0	MLC	2.5”	32	125	55	2.1	1.6	42.9
Hama 00090853	SLC	2.5”	32	60	30	1.8	0.8	25.0

Performance per Joule - USB Flash Drives

Model	Type	Capacity	Throughput		Power Consumption			Performance
			(read)	(write)	(read)	(write)	(idle)	per Joule
		[GB]	[MB/s]	[MB/s]	[W]	[W]	[W]	[MB/J]
Samsung Flash Drive	MLC	8	18	16	0.22	0.38	0.20	57.6
SanDisk Cruzer Mini	MLC	1	13	8	0.13	0.15	0.08	75.0
Super Talent STU1GSMBL	MLC	1	14	5	0.07	0.08	0.06	126.7
CmMemory Core	MLC	1	12	9	0.07	0.08	0.05	150.0
SanDisk Cruzer Mini	MLC	0.5	16	5	0.13	0.13	0.08	80.8

Energy Costs per Year

- Energy costs per year (C_Y) can be calculated:

$$C_Y = E * 24 * 365 * 0.18 \left[\text{kW} * \frac{\text{hours}}{\text{day}} * \frac{\text{days}}{\text{year}} * \frac{€}{\text{kWh}} \right]$$

- This assumes €0.18 per kWh

Energy Costs per Year

Model	Capacity [GB]	Throughput (sustained) [MB/s]	Power Consumption (transfer) [W]	Power Consumption (idle) [W]	Performance per Joule [MB/J]	C_Y idle €	C_Y read/write €
Western Digital RE2 WD4000YR	400	65	10.8	8.9	6.1	14	17
Samsung MCBQE32G5-MPP	32	55/40	0.2	0.1	237.5	0.16	0.32
Samsung Flash Drive	8	18/16	0.2/0.4	0.2	57.6	0.32	0.34/0.59

- Compared to C_Y of HDD storage system:
  - SSD more energy efficient by factor 65
  - USB more energy efficient by factor 34

Our Approach

Our Approach

- Analyze a commodity server and its energy efficiency with different storage systems
- Our aims were to find out:
  - How reasonable is it to replace the HDD by USB flash drives?
  - Which scenario/workload is appropriate for this replacement?
- To answer these questions we measured:
  - I/O performance - sequential and random
  - Metadata performance
  - Energy efficiency - idle and during I/O

Test Environment

- The commodity server had following components:
  - Main board with Intel P35 chip set
  - CPU Intel Core 2 Duo E6750 2.66 GHz FSB1333
  - 2 GB RAM DDR2 2048 MB Kit PC800 CL5
  - 380 W ATX power supply
  - Operating system: Linux Ubuntu 8.04 with Kernel 2.6.24
- To measure energy consumption of the entire system we used the energy cost meter EKM 2000 from Olympia

Evaluated Storage Drives

	Hard Disk Drive	Solid-State Drive	CompactFlash Drive	USB Flash Drive
Model	RE2 WD4000YR SATA	Samsung MCBQE32G5MPP-03A PATA UDMA/66 SLC	SanDisk 8 GB Extreme Ducati Edition	Samsung K9HCG08U1M-PCB00 NAND (512 KB + 16 KB, MLC)
Capacity	400 GB	32 GB	8 GB	8 GB
Purchase Cost	€ 100	€ 370	€ 90	€ 8
Cost per MB	€ 0.25	€ 11.56	€ 11.25	€ 1

Evaluated Configuration

- We selected file systems that are commonly used on hard disk and on flash storage: ext2, ext3, XFS, and VFAT
- For comparison, we also analyzed raw read/write performance of the devices
- To improve capacity and availability of USB flash storage, we tested common RAID configurations: RAID 0, RAID 1, and RAID 5

Performance Measurements

Read Access Time: `h2benchw` (ms)

Access Time	Hard Disk Drive	Solid State Drive	CompactFlash Drive	USB Flash Drive
Minimal	2.98	0.15	0.20	0.48
Average	13.02	0.20	0.64	1.28
Maximal	25.48	1.26	2.35	2.00

- Because the seek time on flash does not depend on the physical location of data, its read performance is almost constant and deterministic
- The HDD access time is over 10 times slower than the USB flash drive and over 65 times slower than the SSD

Sequential Read: `dd` (4 KB block size, MB/s)

File System	Hard Disk Drive	Solid State Drive	Compact-Flash Drive	USB Flash Drive	2 USB Flash Drives (RAID 1)	4 USB Flash Drives (RAID 0)	4 USB Flash Drives (RAID 5)
Device	63.8	58.5	29.2	18.1	18.1	48.4	48.2
ext2	62.7	57.7	29.9	18.2	18.2	48.1	48.0
ext3	63.1	57.8	29.6	18.2	18.2	48.1	48.1
XFS	64.1	57.8	30.1	18.2	18.2	47.4	47.7
VFAT	31.8	57.3	29.1	17.2	21.8	46.7	46.2

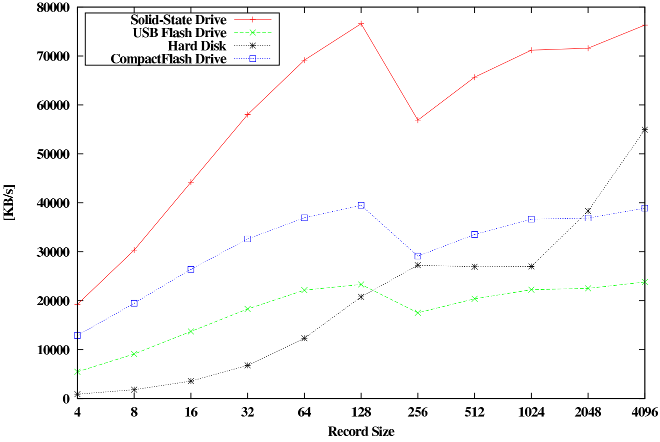
- In this test, HDD shows the best performance, except with VFAT, followed by SSD, CompactFlash, and USB flash
- SSD is only 8% slower than HDD
- USB flash drives in RAID 0 and RAID 5 performs 17% slower than SSD and 25% slower than HDD, except VFAT

Sequential Write: `dd` (4 KB block size, MB/s)

File System	Hard Disk Drive	Solid State Drive	Compact-Flash Drive	USB Flash Drive	2 USB Flash Drives (RAID 1)	4 USB Flash Drives (RAID 0)	4 USB Flash Drives (RAID 5)
Device	64.3	39.0	30.4	16.1	13.8	46.1	4.5
ext2	61.0	31.4	30.2	11.4	9.7	42.5	3.3
ext3	58.9	25.4	25.1	3.7	3.8	32.8	3.3
XFS	65.8	36.2	25.4	14.4	12.8	41.8	2.5
VFAT	61.0	38.3	26.7	12.0	13.2	30.7	3.6

- Hard disk shows the best performance, followed by SSD, CompactFlash, and USB flash drive
- SSD speed reaches only half of HDD speed
- USB flash drives in RAID 0 perform 18% better than SSD, whereas RAID 5 performance is disappointing

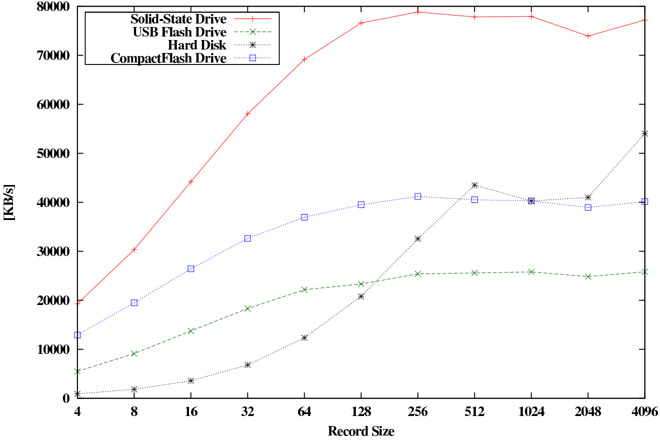
Random Read: IOzone (block size 4-4094 KB, KB/s)



Random Read

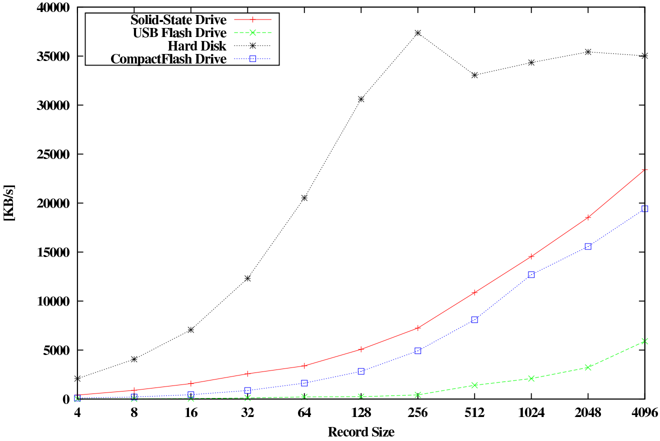
- With block sizes smaller than 128 KB, all flash drives outperform HDD
- The flash performance decrease for block size 128 KB can be explained by the kernel read-ahead technique. This is activated during random read when the data size is bigger than the size of the read-ahead window:
  - The read-ahead window in a current Linux kernel is 128 KB
  - We always read data to fit a read-ahead window, even if the required data size is smaller
  - By increasing a read-ahead window random read performance can be improved

Random Read with Read-Ahead Window of 4096 KB



- Improved throughput for all devices with block size bigger than 128 KB

Random Write: IOzone (block size 4-4094 KB, KB/s)



Random Write: `IOzone` (block size 4-4094 KB, KB/s)

- For random writes, HDD outperforms all flash devices
- Of flash devices, SSD is fastest and USB flash drive slowest
- For block size 256 KB, SSD is factor 5 worse than HDD
- Write performance of USB flash drive is disappointing

Metadata Performance: `fileop` (27 000 files, s)

- Measured file operations: `mkdir`, `rmdir`, `create`, `read`, `write`, `close`, `stat`, `access`, `chmod`, `readdir`, `link`, `unlink`, `delete`

File System	Hard Disk Drive	Solid State Drive	Compact-Flash Drive	USB Flash Drive	2 USB Flash Drives (RAID 1)	4 USB Flash Drives (RAID 0)	4 USB Flash Drives (RAID 5)
ext2	9.0	39.2	145.5	160.3	163.9	196.8	230.0
ext3(1)	1.7	2.3	3.5	30.9	25.5	24.4	30.5
ext3(2)	16.2	61.2	88.6	714.0	840.0	594.8	1110.9
ext3(3)	1.8	2.1	3.5	29.6	25.3	24.3	26.1
XFS	303.7	76.0	105.3	820.3	1058.1	851.2	2004.3
VFAT	8.9	75.5	84.0	11109.5	aborted	aborted	aborted

(1) `ordered`: data is forced to the FS before metadata is committed to the journal
(2) `journal`: data is committed into the journal before being written into the FS
(3) `write back`: data written into the FS after metadata is committed to the journal

Metadata Performance: `fileop` (27 000 files, s)

- HDD shows the best metadata performance
- Tested flash memory is slow due to lack of a cache buffer, whereas HDD has a cache (usually 8-32 MB)
- Considering FS ext3, journaling method `journal` is inappropriate for flash
- Due to intensive journaling activities performance of XFS on USB flash drives is disappointing
- Evaluation of FS VFAT on USB was aborted due to the unacceptable run time

Evaluated Scenarios

Evaluated Scenarios

- Performance measurments suggest that read-mostly and random-I/O workloads are appropriate usage scenarios for USB flash storage
- We tested several scenarios: mail server, database, web server, data server
- In following two examples are presented

Data Server Scenario with Sequential Workload

- Sequential I/O-bound workload with tool `dd` on FS ext3 and *E* of the entire system was measured
- Despite good performance per joule of flash storage subsystem, the performance per joule of entire server with flash is worse

	E_{read} [W]	E_{write} [W]	Read [MB/s]	Write [MB/s]	$P_{seqRead}$ [KB/J]	$P_{seqWrite}$ [KB/J]
Hard Disk Drive	104	97	63	59	621	622
4 USB Flash Drives (RAID 0)	84	84	48	33	586	400

Web Server Scenario

- We consider the web server scenario to be appropriate for flash deployment
- Flash energy efficiency makes it suitable for 24/7 deployment
- Web server runs with intermittent and limited I/O:
    - Its primary activity is often to read static content and deliver it to clients
    - Write access is only required for content updates, database updates, or access logging

Web Server Scenario

- Apache 2 web server was used:
  - 240.000 HTML pages and images with an average size of 8.7 KB
  - Total volume of the content 2.0 GB
  - Server did not access a database
- Benchmark `http_load` measured the time needed to fetch 100 000 files from the server using 40 concurrent requests. For this, a second machine was connected by Gb Ethernet
- Separate experiment measured logging overhead, where the logfile was placed on `tmpfs`

Web Server Scenario: `http_load` (s)

File System	Hard Disk Drive	Solid State Drive	USB Flash Drive	2 USB Flash Drives (RAID 1)	4 USB Flash Drives (RAID 0)	4 USB Flash Drives (RAID 5)
ext3	403.0	43.4	132.2	137.6	72.3	49.8
ext3 (log on tmpfs)	403.0	38.3	120.4	127.1	43.3	36.9

- Here, SSD is the fastest. A single USB drive is factor 3 better than on HDD, RAID 0 factor 5.6, and RAID 5 factor 8
- Using RAID 0 instead of one flash dive the performance increased by 45%
- Write access of the logfile has a significant impact on flash performance, especially in RAID 0 configuration

Energy Metrics for `http_load`

Metric	Hard Disk Drive	Solid State Drive	USB Flash Drive	4 USB Flash Drives (RAID 0)
System Power [W]	95	100	88	100
Time to perform 100 000 fetches [s]	403.0	43.4	132	72.3
Request/Time [1/s]	248	2304	756	2310
Request/Energy [1/J]	2.7	23.3	8.6	23
Total Energy [kJ]	38.3	4.3	11.6	7.2
Energy Costs [cent]	0.180	0.022	0.054	0.036

- Server with four USB drives or SSD provides factor 8.5 more requests per joule than system with HDD
- Energy costs for the web server with HDD are factor 3 higher than with single UBB flash drive, factor 5 than four USB flash in RAID 0, and factor 8 than SSD

Conclusion

- USB flash storage is appropriate for read-mostly and random-I/O workloads
- USB flash storage drawbacks are low write performance and higher purchase cost in comparison to HDD
- For read-mostly workload, energy efficiency for entire system with USB flash drives is better than for system using HDD
- USB flash drives are now an option for scenarios without high requirements to write performance and capacity

Thank you for your attention!