Flash-Based Hybrid Systems

汤　显

磁盘是应用最广泛的存储设备，但是它的性能提升空间很小，而近几年 flash 的性能不断提升，价格却在不断的降低，它固有的性能优势使得它大有取代磁盘的趋势，但是 flash 的读写代价不平衡及擦除次数的限制使得它的应用受到一定的限制，另外，尽管 flash 的容量在不断的变大，但还远不及磁盘，而且 flash 的价格也远高于磁盘，这使得 flash 在近几年还不可能完全取代磁盘，因此可将 flash 与磁盘进行组合，利用 flash 和磁盘各自的优点，克服其相应的缺点，从而提升整个混合式系统的性能。

现有的混合式系统，按照其所应用的环境，可分为网络服务器，移动计算和数据库三种。

在网络服务器环境中，读操作远远多于写操作，系统的性能很大程度上依赖于内存，带宽和读写延迟，要想提高系统的性能，就需要使用大容量的内存，但是大容量的内存价格昂贵并且耗电量巨大。同样大小的 RAM 和 flash 相比，flash 的价格要低很多，而且 flash 的耗电量要远远小于 RAM，而且采用 flash 之后，系统的性能并未下降，因此可采用 flash 作为二级缓存来减少 RAM 的空闲时耗电量以及提高系统整体的性价比。

在移动计算环境中，因为需要用电池供电，因此耗电量是主要考虑的因素。而 flash 耗电量少的这一特性使得它很适合用在移动计算环境中，而且使用 flash 作为磁盘的缓冲，可减少磁盘的读写延迟；可延长磁盘空闲的时间，从而节省能量。

在数据库环境中，既有使用 flash 作为缓存的应用，也有使用 flash 作为辅存的应用。利用 flash 作为缓存，主要是为了减少磁盘的读写延迟；利用 flash 作为辅存，主要是利用了 flash 的快速随机读的特性。

不管是那种应用，都是为了提高系统整体的性能，详细内容见下面的 ppt。

Flash-based Hybrid Systems

汤显

Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

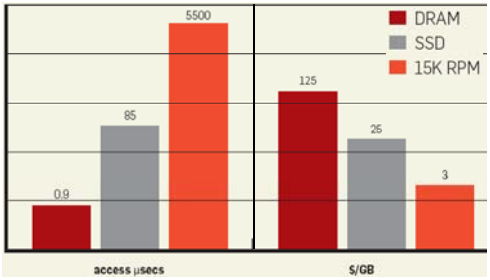
Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

Introduction

10TB	Cost	power
7,200RPM	$3,000	112W
15,000RPM	$22,000	473W

- DISK
  - cheap
    - < $1 per GB for 7,200RPM drives
    - $3 per GB for 15,000RPM drives
  - a latency improvement of a bit more than a factor of two
- FLASH
  - Cost, Capacity, performance, power consumption
  - Unbalanced r/w performance, lifetime, capacity, cost



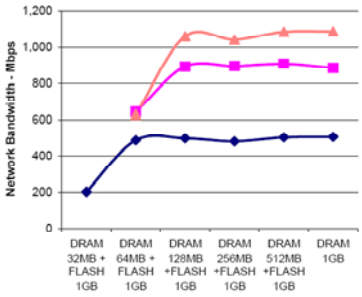
A. Leventhal. Flash Storage Memory. Communications of ACM, July 2008

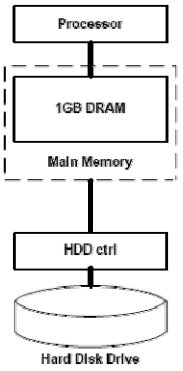
Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

Hybrid systems on web server

- Background:
  - Read >> write
  - Performance depends heavily on memory, I/O bandwidth, access latency
  - RAM is a large power consumer
  - RAM is expensive
- Goal:
  - Reduce RAM idle power consumption
  - Cheaper
- Solution:
  - Flash as secondary buffer

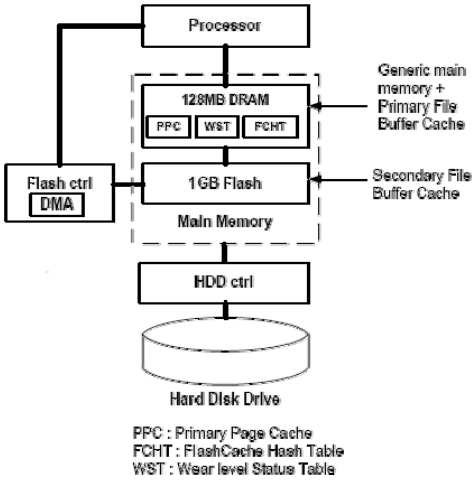




Taeho Kgil, Trevor N. Mudge: FlashCache: a NAND flash memory file cache for low power web servers. CASES 2006:103-112

Hybrid systems on web server

- Background:
  - Read >> write
  - Performance depends heavily on memory, I/O bandwidth, access latency
  - RAM is a large power consumer
  - RAM is expensive
- Goal:
  - Reduce RAM idle power consumption
  - Cheaper
- Solution:
  - Flash as secondary buffer



Taeho Kgil, Trevor N. Mudge: FlashCache: a NAND flash memory file cache for low power web servers. CASES 2006:103-112

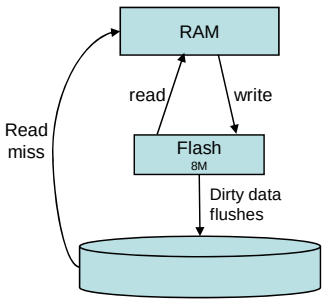
Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

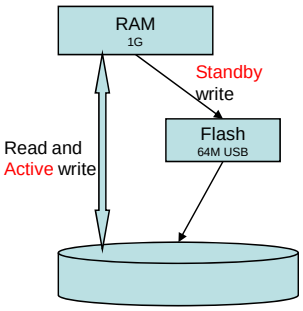
Hybrid systems on mobile computing (1)

- Background:
  - Limited energy
  - Disk is a big energy consumer
    - Active/standby = 5~10
    - Idle: spin-down; request: spin-up
    - OS incur periodic writes during idle periods
- Goal:
  - Reduce energy consumption
- Solution:
  - Flash as buffer

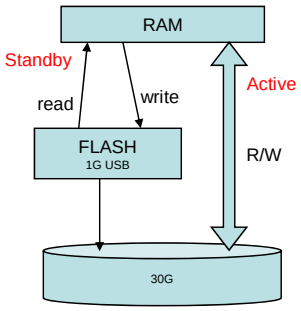
Hybrid systems on mobile computing (1)



- Flash as read and write buffer
  - reduce latency



- Flash as standby write-only buffer
  - Reduce the number of spin-up/down operation
  - Prolong the period of spin-down

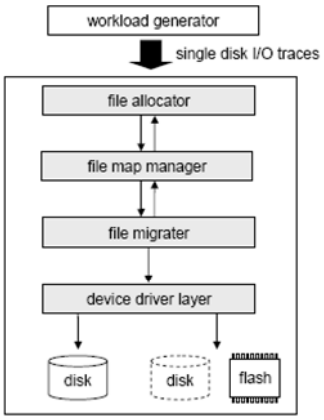


- Flash as standby buffer
  - Move slowly from/to disk
  - Widely scattered data
  - Flash full or disk active

[1]Brian Marsh, Fred Douglis, P. Krishnan: Flash Memory File Caching for Mobile Computers. HICSS 1994:451-461
[2]T. Bisson and S. Brandt, "Reducing energy consumption with a nonvolatile storage cache," IWSSPS, 2005
[3]Feng Chen, Song Jiang, Xiaodong Zhang: SmartSaver: turning flash drive into a disk energy saver for mobile computers. ISLPED 2006:412-417

Hybrid systems on mobile computing (2)

- Background:
  - Limited energy
  - Disk is a big energy consumer
  - Flash as cache
    - Frequently write reduce the lifetime
- Goal:
  - Energy efficient
- Solution:
  - Flash as secondary storage



Young-Jin Kim, Kwon-Taek Kwon, Jihong Kim: Energy-efficient file placement techniques for heterogeneous mobile storage systems. EMSOFT 2006:171-177

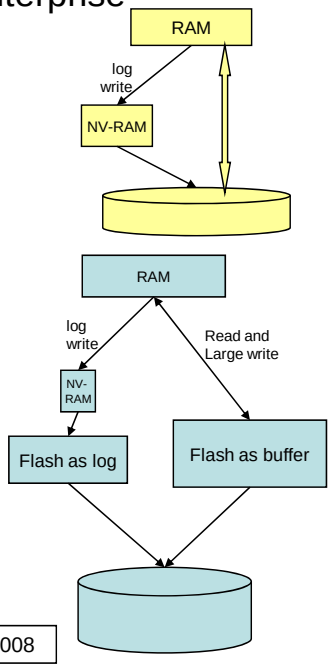
Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

Hybrid systems on database(1)

-- enterprise

- Background:
  - Use dedicated log device to improve write performance
  - NV-RAM as log device
    - write faster
    - Cost, size (2GB–4GB), battery (power, fail)
- Goal:
  - Cheap
  - High performance
  - Energy efficient
- Solution:
  - Smaller flash as log device, larger flash as read and write buffer



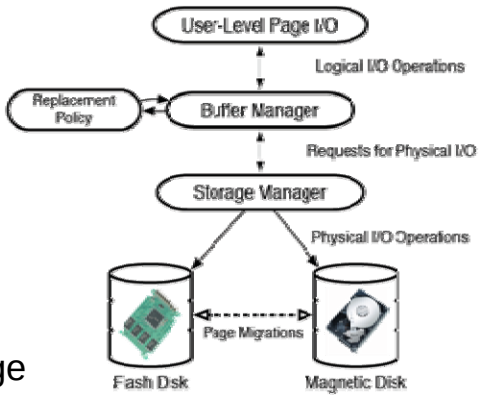
A. Leventhal. Flash Storage Memory. Communications of ACM, July 2008

Hybrid systems on database(2)

-- commercial

- Background:
  - Flash (MLC)
    - large capacity
    - faster at random reads
    - low random write speed
  - Magnetic disk
    - faster at random writes
- Goal:
  - high performance
- Solution:
  - Flash as secondary storage

	Random Read (ms)	Random Write (ms)
Flash Disk	0.4	48
Magnetic Disk	9.5	4.9



I. Koltsidas, and S. D. Viglas. Flashing Up the Storage Layer. VLDB 2008.

Outline

- Introduction
- Hybrid systems
  - Web server
  - Mobile computing
  - Database
- Conclusion

Conclusion

- Flash-based hybrid systems involves extensive applications
- Research on flash-based hybrid systems is an important issue