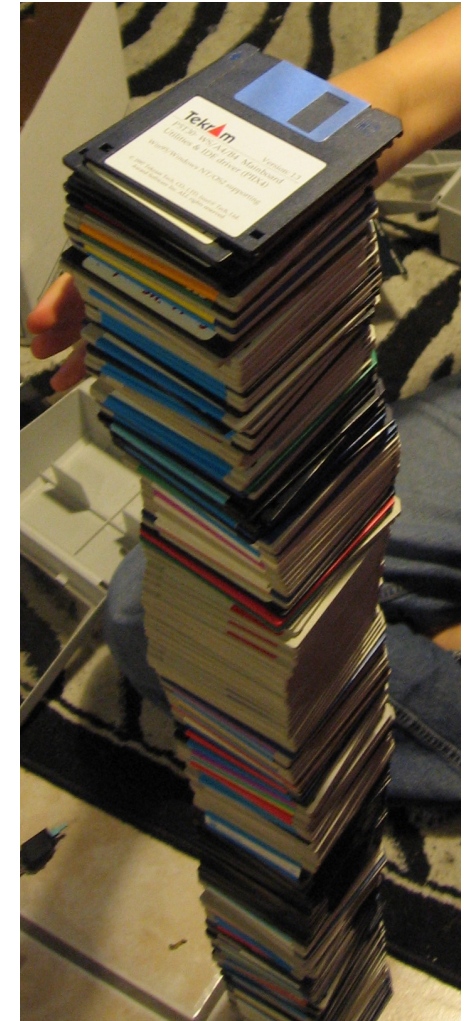




# The Death of Disk Panel Session

Erik Riedel, EMC  
HEC FSIO Workshop  
August 2011



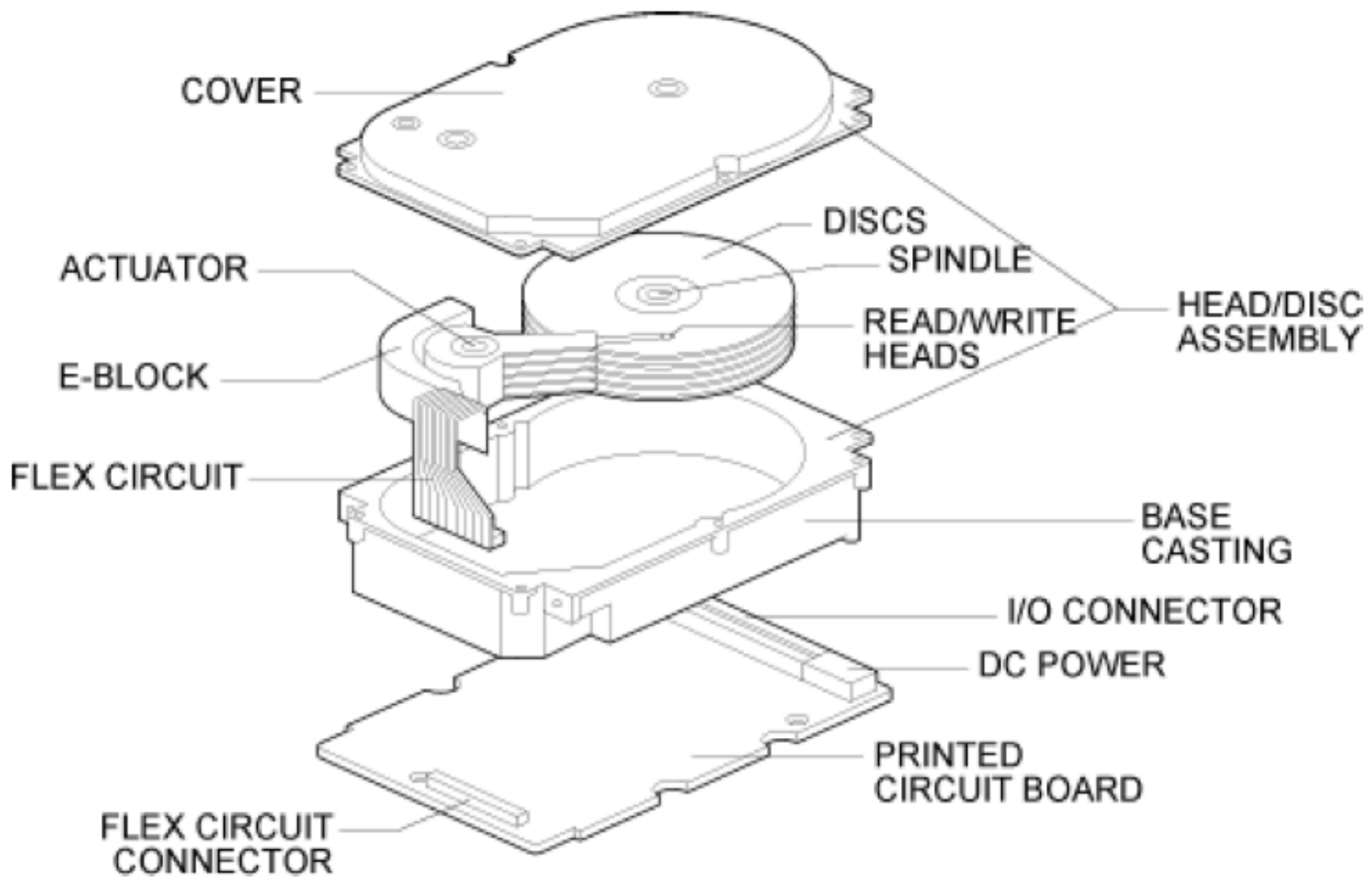
top picture "floppy disks for breakfast" by Blude via flickr/cc  
right picture by Austin Marshall via flickr/cc

# Conclusion

- About 80% of stored data will never be accessed again
- About 80% of the rest will be accessed predictably
- That leaves (maybe) 4% of stored data that potentially requires “quick” random access
- => Buy as much flash as you can afford, use disks for the rest

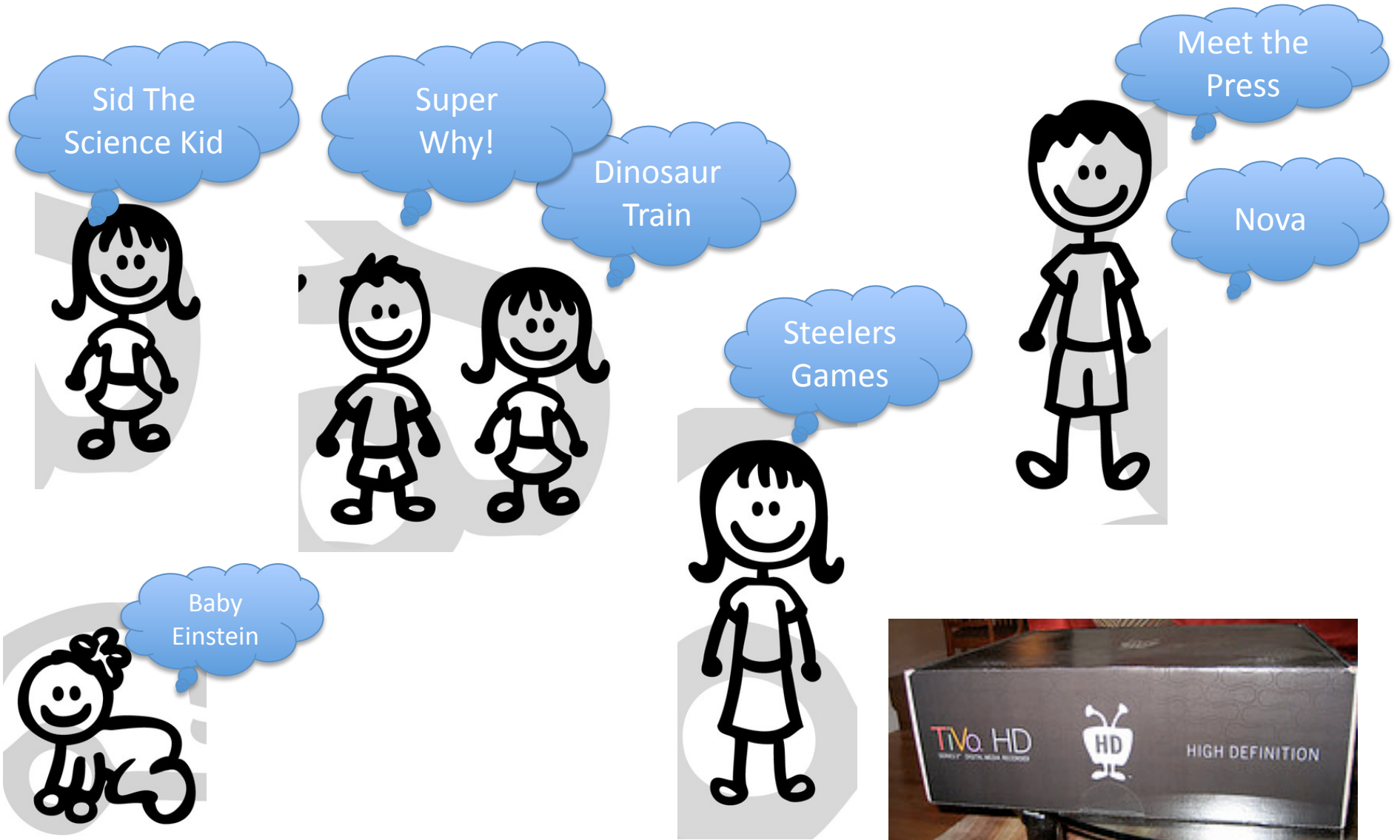
# Most Data Is Idle

- About 80% of stored data will never be accessed again
- Disk drives have long been designed around this key fact of the digital world
- Amortize a relatively small amount of expensive read/write electronics and fancy material science over a large and cheap magnetic media



David Anderson, James Dykes, Erik Riedel "SCSI vs. ATA - More than an interface" *2nd Conference on File and Storage Technology (FAST)*. San Francisco, CA. April 2003. [www.cs.cmu.edu/~riedel](http://www.cs.cmu.edu/~riedel)

# Consumer Example (At My House)



# Most Data Access Is Predictable

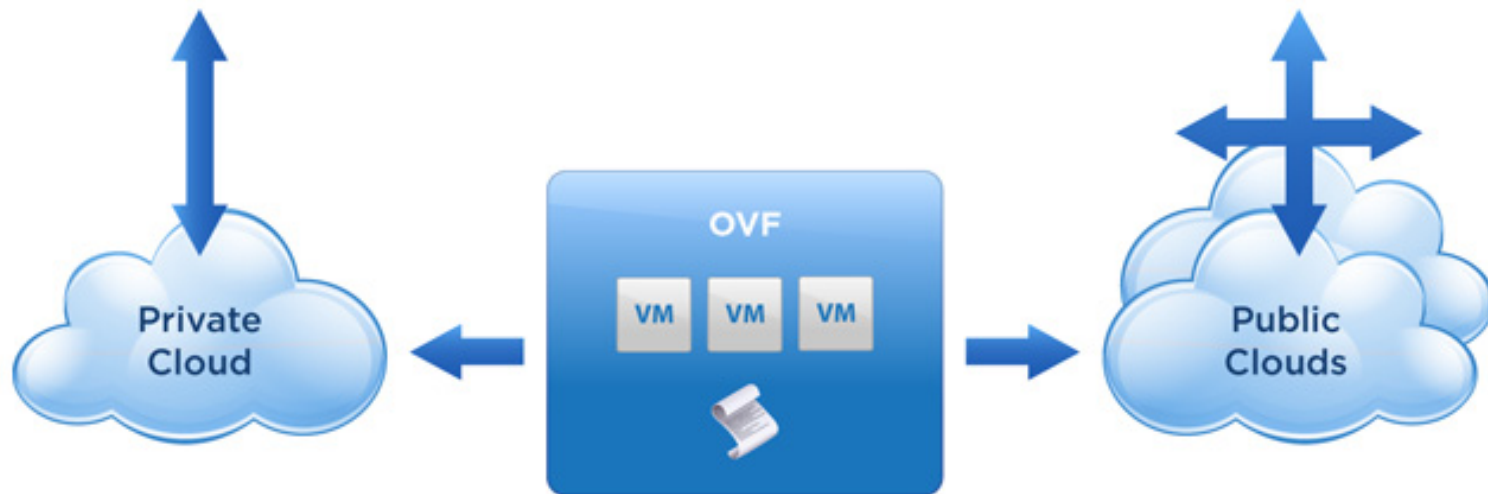
- Caching
  - Prefetching
  - Tiering
  - Staging
  - Hierarchical Storage Mgmt
- 
- all these tools have been known for years
  - just need to open our toolbox, sharpen some of them to apply to today's infrastructure/apps

# New Tools In the “Cloud”

VMware vCloud API

The First Cloud API Submitted to Open Industry Standards

APIs: Programmatic Access to Resources



Open Virtualization Format (OVF)

The First Industry Standard for Cloud Workloads

Marketing buzz – IaaS – Infrastructure as a Service

# New Tools In the “Cloud” (2)



The image shows a screenshot of the VMware Cloud Foundry website. At the top right, the VMware logo is visible. Below it, there are links for "Blog", "FAQ", "Community", and "Forums". On the left side, there is a diagonal banner that says "BETA". The main heading is "CLOUD FOUNDRY" in large blue letters, with the tagline "No Obstacles: Deploy and Scale Your Applications in Seconds" underneath. To the left of the heading is a logo consisting of a blue cloud shape made of small squares. Below the heading, there is a paragraph of text: "The industry's first open platform as a service. Run your Spring, Rails and Node.js applications. Deploy from your IDE or command line." At the bottom, there is a blue banner with a white button that says "Free Sign Up for Cloud Foundry". To the right of the button, there is text: "Request an invite for a CloudFoundry.com account and get notified when your Micro Cloud is available for download".

vmware®

[Blog](#) [FAQ](#) [Community](#) [Forums](#)

**BETA**



## CLOUD FOUNDRY™

No Obstacles: Deploy and Scale Your Applications in Seconds

The industry's first open platform as a service. Run your Spring, Rails and Node.js applications. Deploy from your IDE or command line.

**Free Sign Up for  
Cloud Foundry**

Request an invite for a CloudFoundry.com account and get notified when your Micro Cloud is available for download

Marketing buzz – PaaS – Platform as a Service



# New Tools in the “Cloud” (3)

- Key takeaways
  - both IaaS and PaaS are “closed loop” infrastructures
  - apps cannot be deployed except at the “direction” of the system
  - logging and monitoring are constant
    - need to get high utilization rates (\$\$)
    - need to send out bills (\$\$)
    - want high rates of “multi-tenancy” to be efficient (\$\$)
  - this leads to a significant level of “predictability”

# Get Predictability Into Storage

- Key challenge is how to translate what “the system” knows about apps and behaviors and “SLAs” into guidance for our system-level tools (caching, prefetching, tiering, etc.)
- Secondary challenge is avoiding “surprises”
  - where performance or availability or durability don’t meet the SLAs (“quality of service”)
- Good news is that the new infrastructures have some powerful new ways to help us

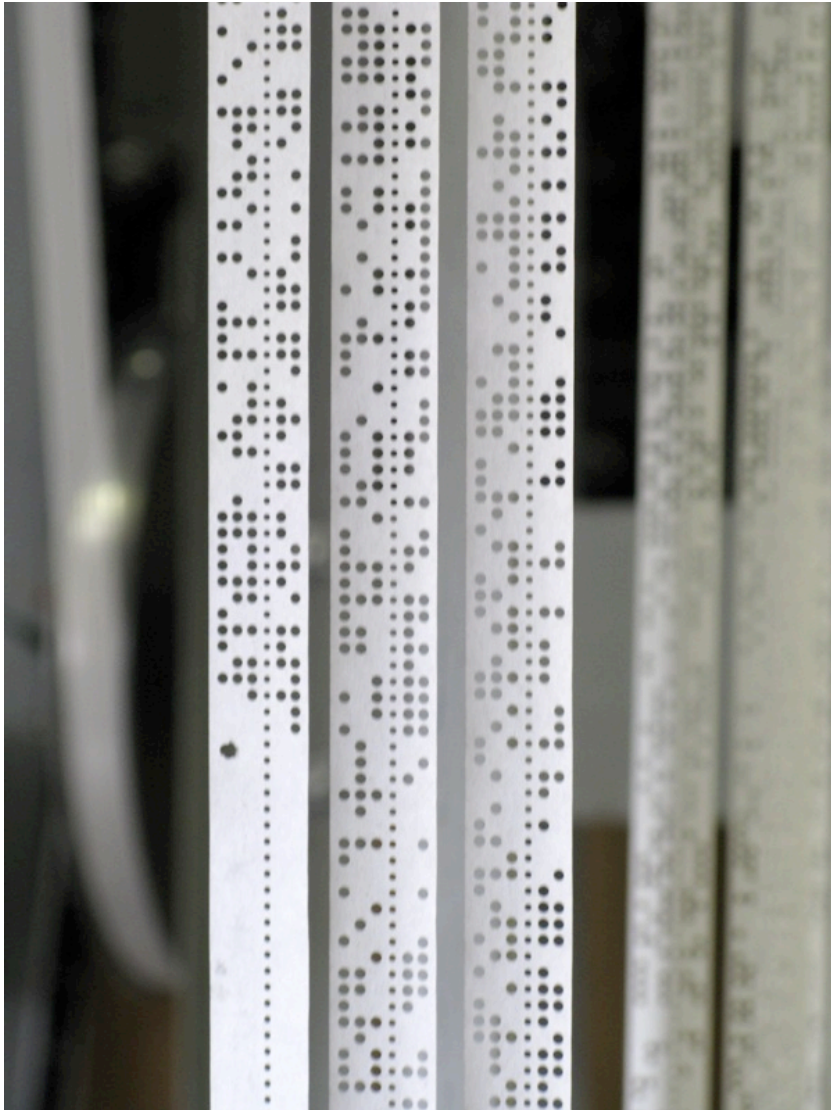
# One Example New Tool – Stunning



- “The amount of time the virtual machine is stunned is dependent on the amount of memory to be written to disk for such an operation, and the speed and responsiveness of the datastore's backing storage.”  
– VMware KnowledgeBase

[http://kb.vmware.com/selfservice/microsites/search.do?language=en\\_US&cmd=displayKC&externalId=1013163](http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1013163)

# What About Tape?



# What About Tape?

- Tapes are not a commodity technology
- 2011 total worldwide market for tape cartridges is about 8m units (just under \$1b annual revenue)
- Compare to the HDD business at 650m units in 2010 (close to \$40b annual revenue)
- 80 disk drives are manufactured for each tape cartridge; robots are complicated
- Fits particular application segments very well, but is not a general-purpose solution

<http://www.storagenewsletter.com/news/tapes/sccg-ww-tape-market-lto-1q11>

<http://techreport.com/discussions.x/20890>

# Conclusion

- About 80% of stored data will never be accessed again
- About 80% of the rest will be accessed predictably
- That leaves (maybe) 4% of stored data that potentially requires “quick” random access

# Pragmatic Issues

- Power
  - if the data is predictably idle, then don't spin it
- Wearout
  - look at the data occasionally (once/month, once/yr); such access for “scrubbing” is very predictable
- Backup
  - backup 1) is not an “application” and 2) is predictable
- Replication
  - estimates run to 75% of stored data is copies/replicas, only 25% unique bytes; replication is predictable

# Summary – How Much Data

- 1.2 million PB estimated in 2010
- 25% unique => leaves 300,000 PB
- 80% idle => leaves 60,000 PB
- 80% predictable => leaves 12,000 PB
- at \$1/GB for flash, that requires \$12b
- is that affordable?
- (remember the world bought ~\$40b of HDD in 2010)

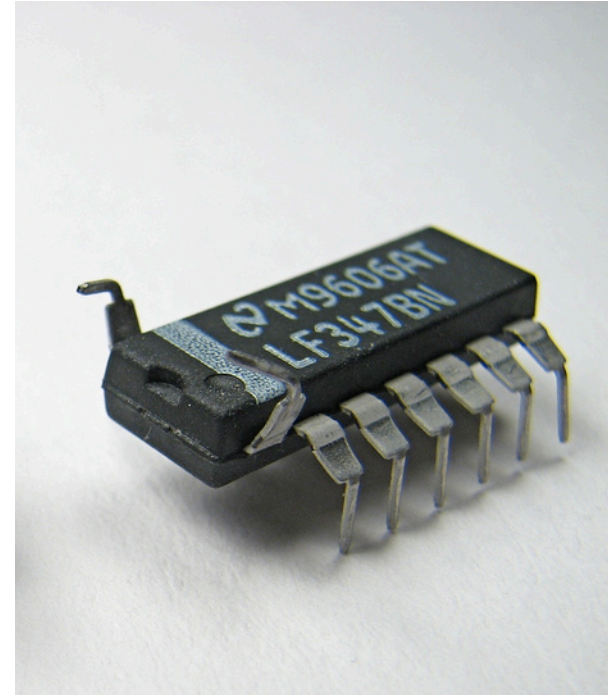
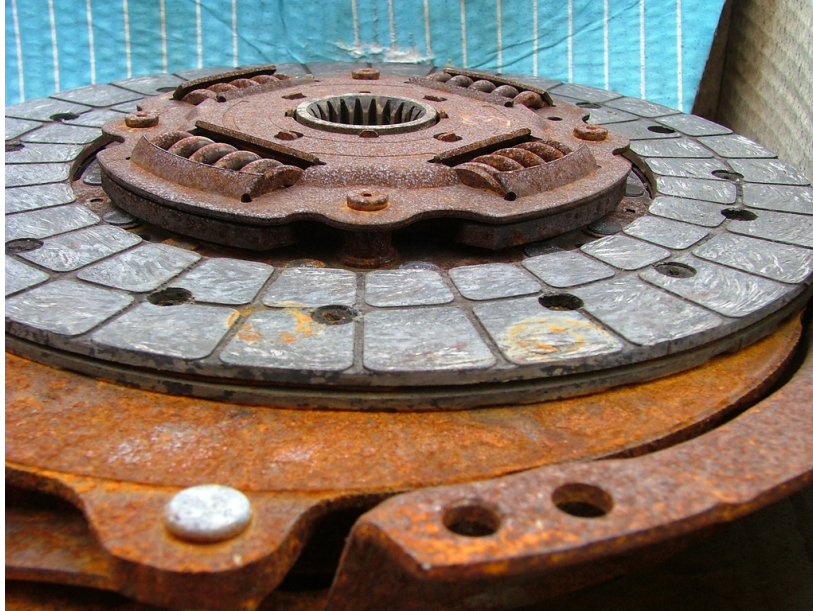
[www.zdnet.com/blog/service-oriented/size-of-the-data-universe-12-zettabytes-and-growing-fast/4750](http://www.zdnet.com/blog/service-oriented/size-of-the-data-universe-12-zettabytes-and-growing-fast/4750)  
[www.computerworld.com/s/article/9180943/NAND\\_flash\\_memory\\_pricing\\_to\\_plummet\\_to\\_1\\_per\\_GB](http://www.computerworld.com/s/article/9180943/NAND_flash_memory_pricing_to_plummet_to_1_per_GB)  
recently – “the price of flash has not been dropping as fast as the suppliers predicted”, August 2011



# Conclusion

- About 80% of stored data will never be accessed again
- About 80% of the rest will be accessed predictably
- That leaves (maybe) 4% of stored data that potentially requires “quick” random access
- => Buy as much flash as you can afford, use disks for the rest

n.b. 73.6% of all statistics are made up, do the calculations for your own environments, your mileage may vary



- In the battle of rust vs. silicon, both will survive

rust picture by Jos Faber via flickr/cc

silicon picture from "Chip bug vs chip bug" by Windell Oskay via flickr/cc

# ATMOS

MANAGING BIG DATA IN THE CLOUD

WATCH VIDEO

 Announcing New EMC Atmos 2.0 »



**OVERVIEW**

OPTIONS

RESOURCES



**BIG DATA-READY**

Use flexible, cloud-based architecture for big data.

<http://www.emc.com/storage/atmos/atmos.htm>

a brief word from my sponsors