DARPA Workshop: Future of Storage

David S. H. Rosenthal

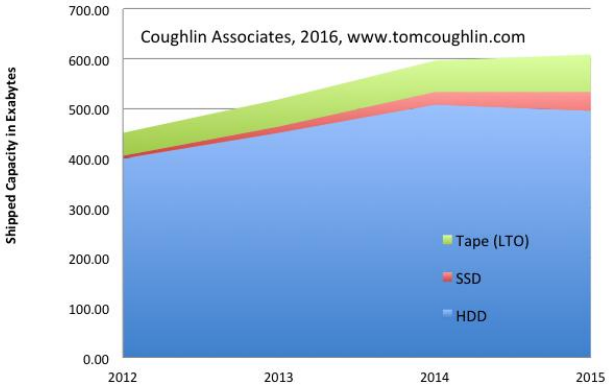
LOCKSS Program
Stanford University Libraries
http://www.lockss.org/
http://blog.dshr.org/

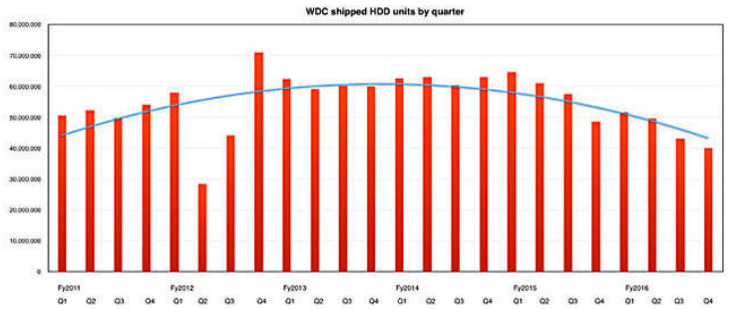
Main Topics

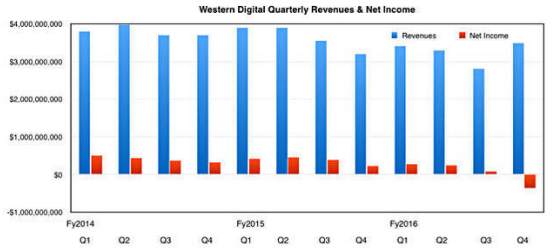
- Medium term (1-2 decade) bulk storage
- Alternative media
- Aggressive data compression

Hard disk vs. Flash

- It takes longer than it takes:
- Disk 60yrs, dominates bulk market
- Flash 30yrs, not yet in bulk market
- Peak disk is here
- Bulk only growth mkt, less volume
- Disk is volume business
- Less volume = less money
- Harder to finance improvement

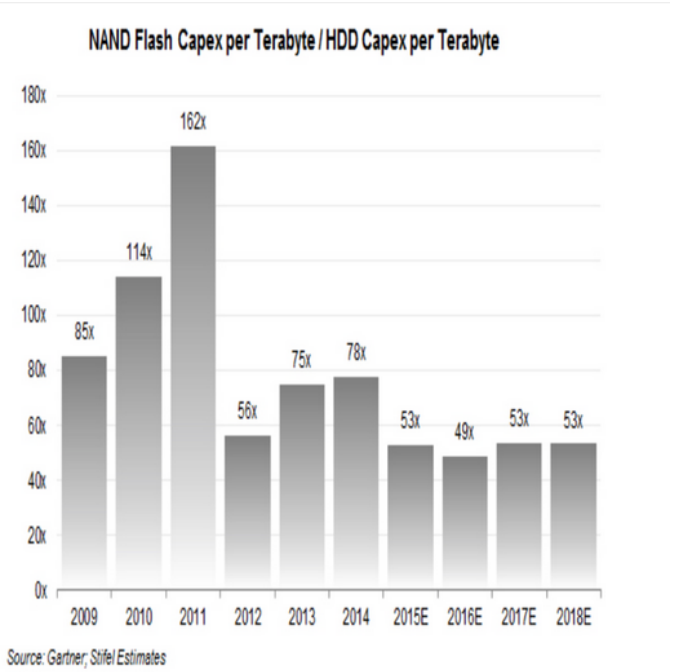






Flash displace disk?

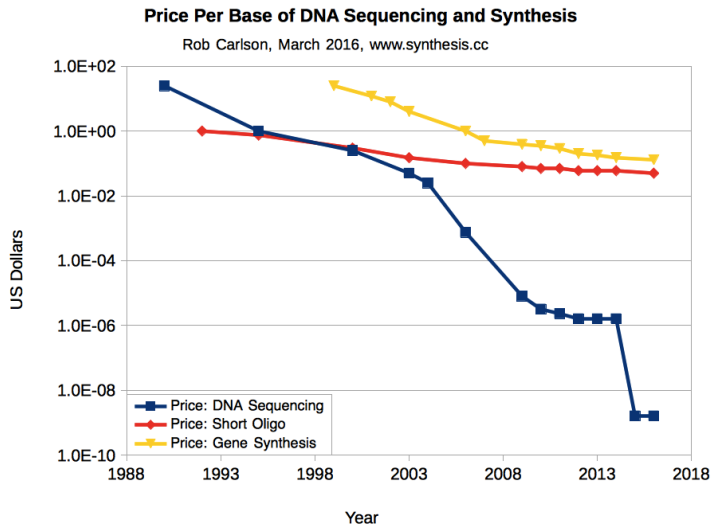
- More fabs, more wafers:
  - Expensive but low interest rate
  - Strategic investment
- More bytes per wafer:
  - 3D shipping
  - QLC in development
  - QLC maybe OK for archival



Cell Type	SLC (1-bit)	MLC (2-bit)	TLC (3-bit)	QLC (4-bit)
5Xnm	11,000	10,000	2,500	800
3Xnm	10,000	5,000	1,250	350
2Xnm	7,500	3,000	750	150
1Xnm	5,000	1,500	500	70

DNA storage

- Eventually may be great archive medium:
  - Dense, stable, lots of copies easy
- Huge problems to solve first:
  - Write 1E6-1E8 times too costly, read cheap but slow
  - High up-front, low running cost = bad economics



Aggressive Compression

- DARPA customers lots of surveillance video:
  - Feature extract – “Kilroy was here” - huge compression
- Forensic use – need to reconstruct video:
  - Only a small proportion, so reconstruct can be costly
- Now possible (not useful for archives):
  - Save features plus small amount more data
  - Very high compression
  - Very expensive multi-dimensional search to reconstruct
  - Looking for the image that would have features + data