# Exascale Memory: Requirements and Options

Rob Schreiber

HP Labs

May 2013

# It will be a very parallel machine

Exaflops at gigahertz

➔

a billion operations per clock

# How will we get that much parallelism?

- Big problems
- More than one problem
- Pipelines of problems
  - Preprocess, mesh gen
  - Solve, resolve, UQ, optimize
  - Postprocess and visualize

# And we will need a lot of memory

- Many problems on the machine at one time
- Performance costs memory
  - Demmel's 2.5D matrix multiply
  - Large ghost cell halos
  - Numerous other examples of this
  - Replication of data (nuclide cross sections, for example)
  - Checkpoints in memory, perhaps multiple versions
  - Message logs

# But we won't have a lot of power

Must be 25X better than today's best in flops/Watt

# What matters for memory (at exascale)

1. Cost per bit, density
2. Static and dynamic power
3. Access latency
4. Error probability, vulnerabilities
5. Durability, wearout
6. Data retention time
7. Volatility

# Density challenge

- Ten bucks a gigabyte (DRAM) today
  - A DRAM exabyte costs ten billion dollars
  - Two or three factors-of-two won't be enough
- Let's consider other memory technologies
- I don't mean block-oriented, sequentially accessed, slow-to-write storage technologies

# How do we get the low cost per bit?

- Feature shrink
- MLC technologies
- 3D technologies (not stacking, real 3D…)
- DRAM seems able to benefit only from feature shrink
  - And that is getting hard for DRAM

# Power challenges

- DRAM – it is barely doable
  - Overfetch must be corrected
  - Leakage, refresh

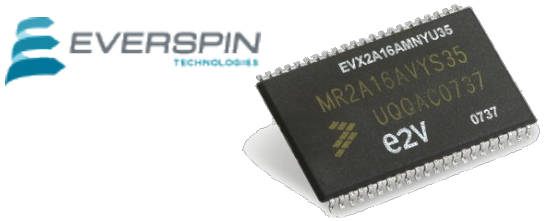
- Nonvolatile memory is usually energy-costly to write
  - But there may be a lot of infrequently written data in memory

# Flash

- Already a bigger business than DRAM
- Challenging HDDs in density, and way faster
- Way too slow to use it for memory
- And it can wear out

# New memory on the horizon

- Spin-Torque-Transfer RAM (STTRAM)
  - Grandis (54nm, acquired by Samsung)
- Phase-Change RAM (PCRAM)
  - Samsung (20nm, diode, up to 8Gb)
  - Micron and Nokia – In phones now
- Resistive RAM (ReRAM)
  - Panasonic (180nm process, 4-layer xpoint)
  - Unity Semi (64MB, acquired by Rambus)




GRANDIS

RENESAS

TOSHIBA

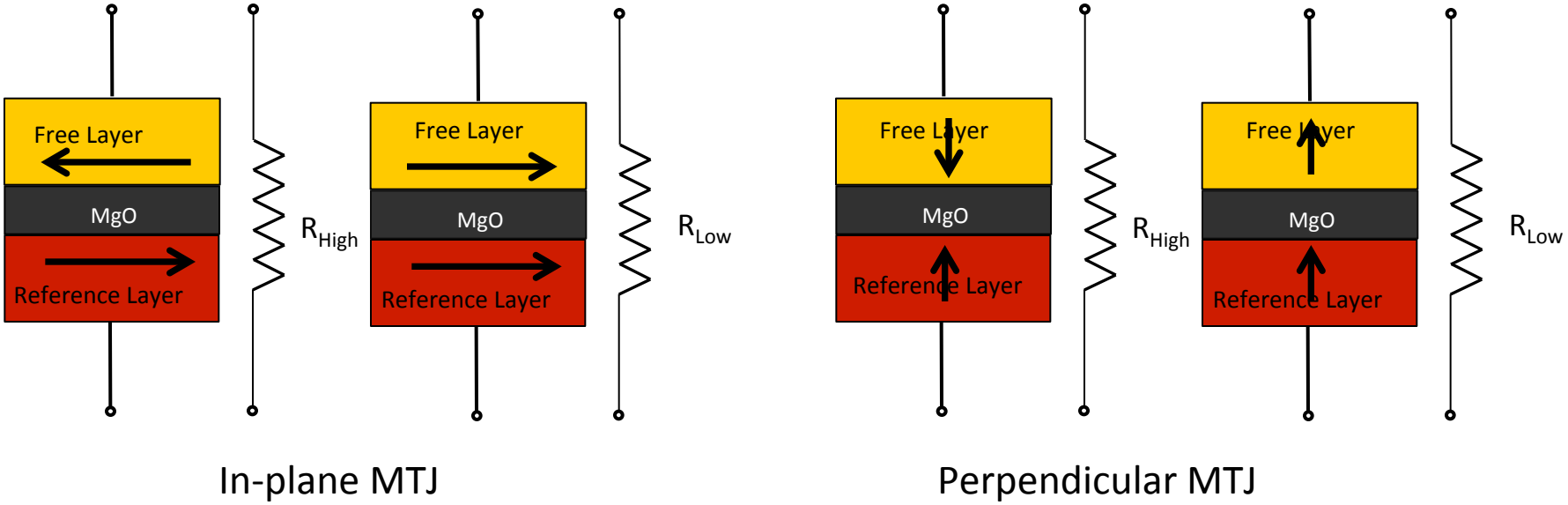

numonyx™

hynix

Panasonic
ideas for life

Rambus

# Spin transfer torque (STTRAM)




In-plane MTJ

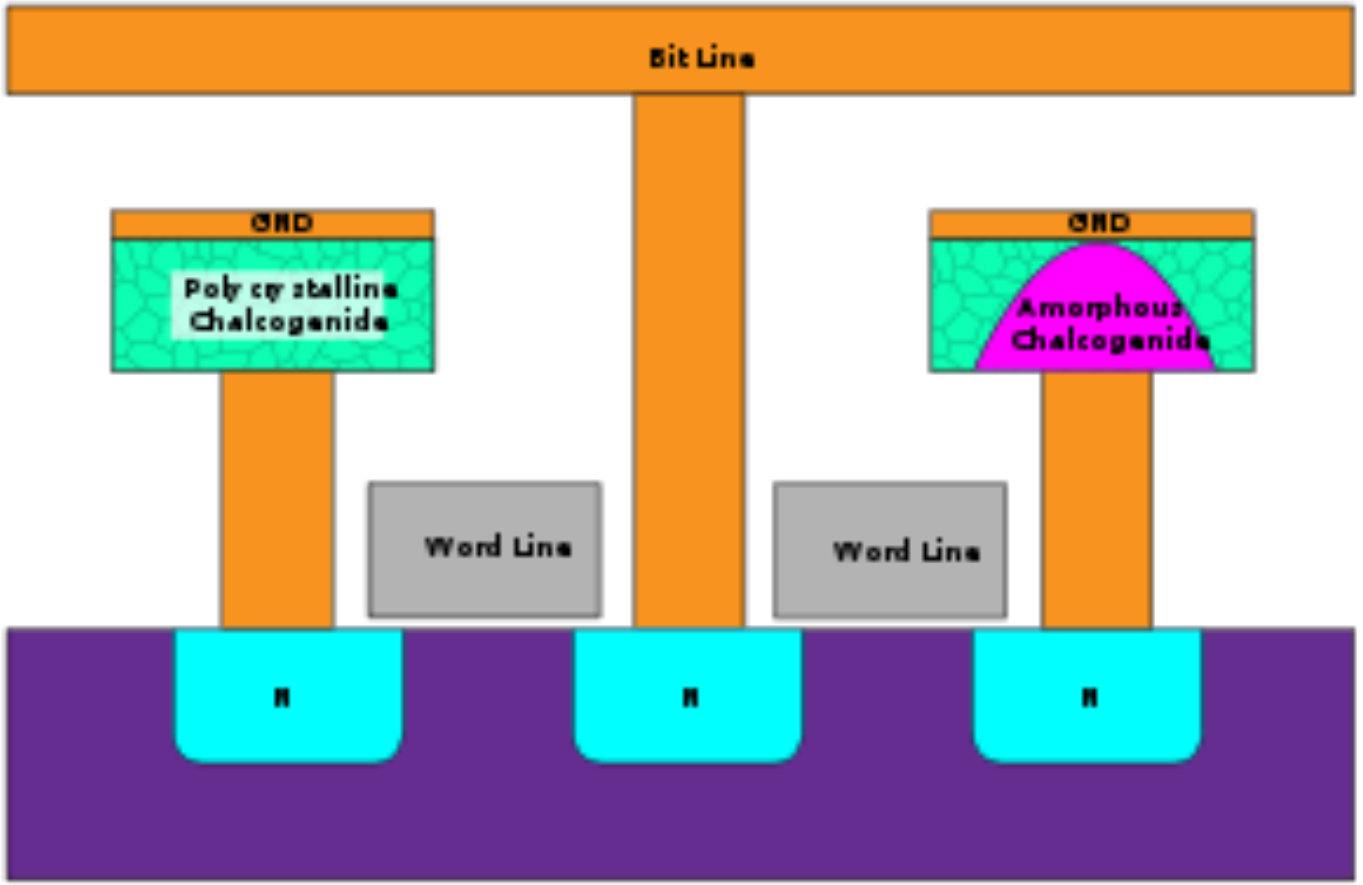
Perpendicular MTJ

Fast enough for use in cache

No soft errors (an issue for SRAM)

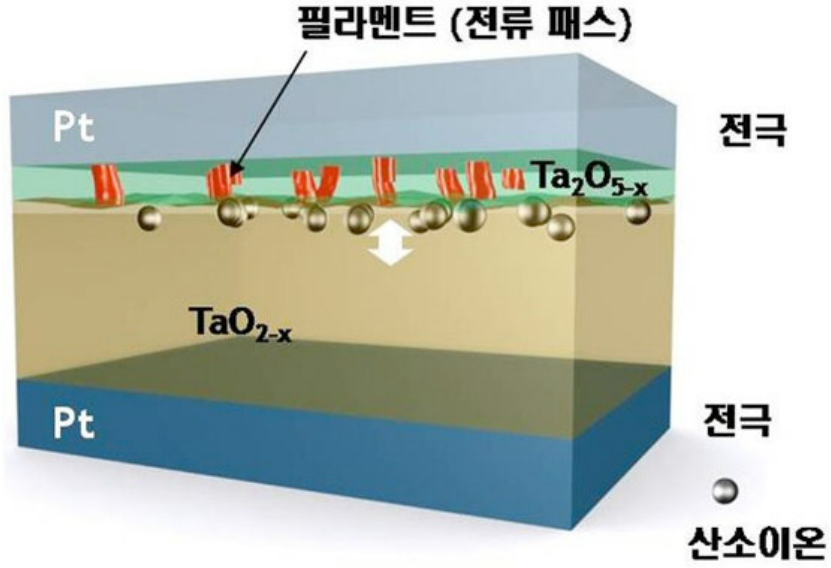
May not have the density we need for main memory

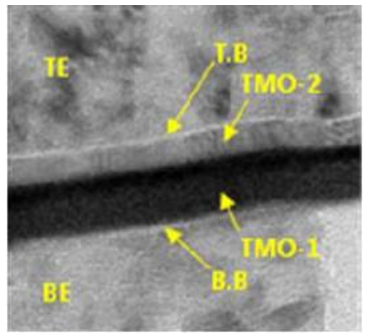
# PCRAM




- Shipping today
- MLC (limited by resistance drift)
- Slow, expensive writes
- Wearout issue

# ReRAM




Samsung, HP-Hynix, Sandisk, Toshiba

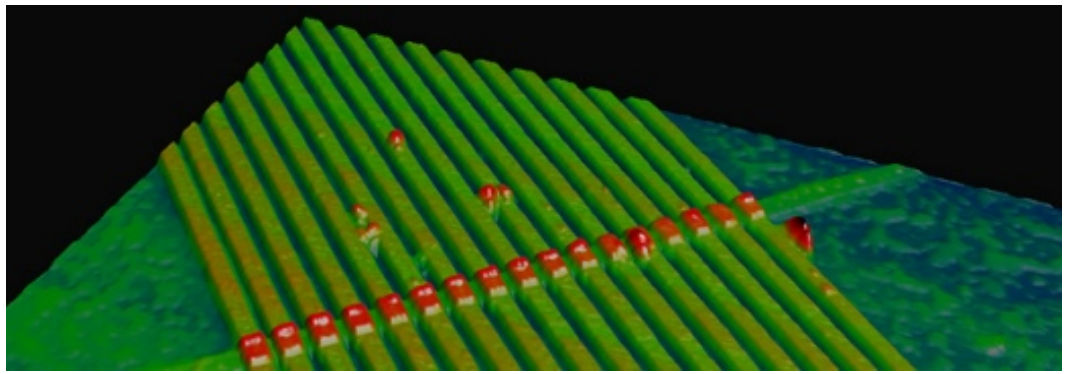
32Gb test chip (Sandisk/Toshiba. 24 nm. ISSCC 2013)

Fast (tens of nsecs) for both read and write

Good data retention and reliability

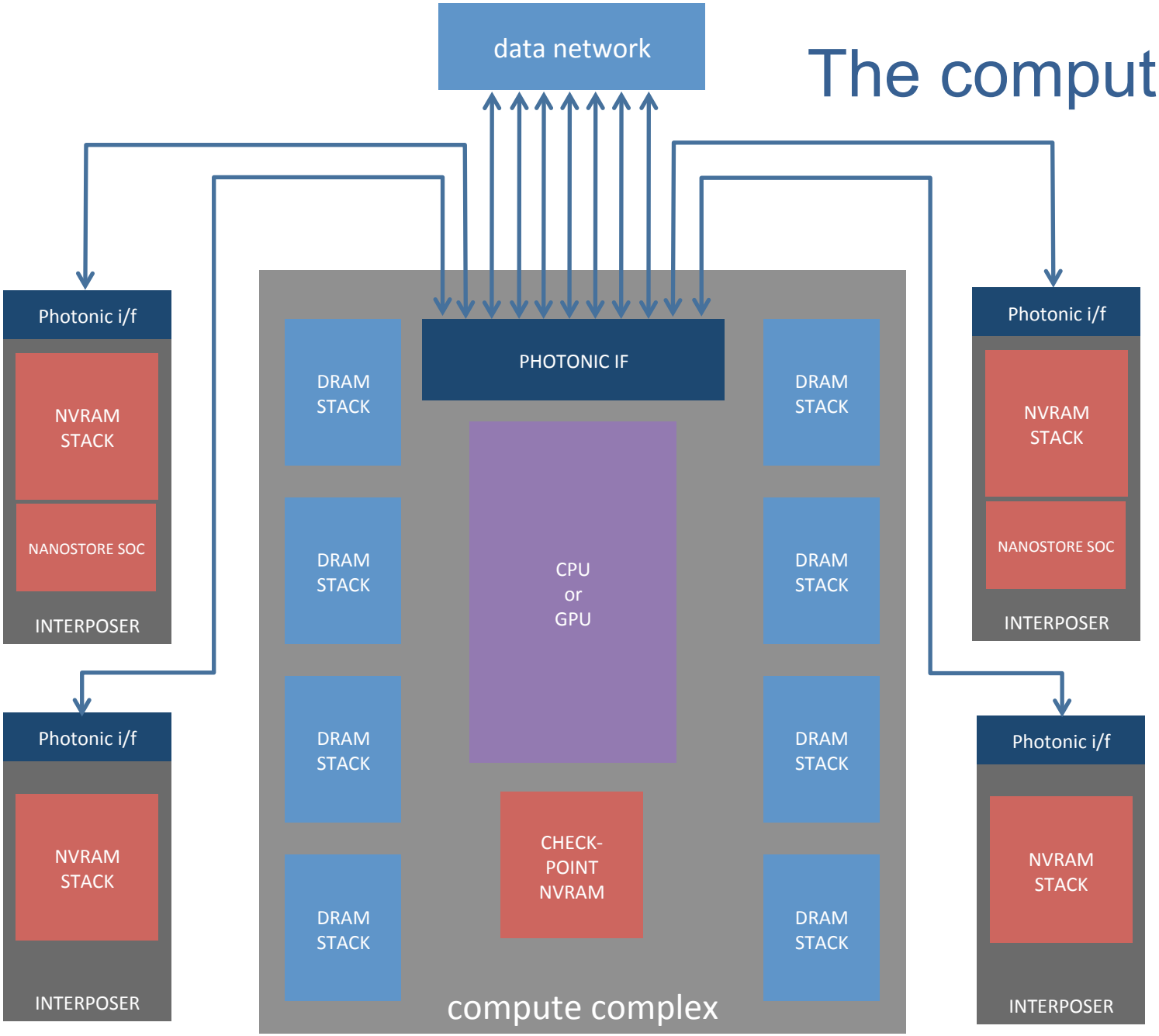
3D -- 2 to 4 layers

MLC possible



| | SRAM | DRAM | eDRAM | NAND Flash | PCRAM (SLC) | STTRAM | ReRAM (1T1R) | ReRAM (Xpoint) |
|---|---|---|---|---|---|---|---|---|
| Data Retention | N | N | N | Y | Y | Y | Y | Y |
| Cell Size/bit ($F^2$) | 50-200 | 4-6 | 19-26 | 2-5 | 4-10 | 8-40 | 6-20 | 1- 4 |
| Read Time (ns) | < 1 | 30 | 5 | $10^4$ | 10-50 | 10 | 5-10 | 50 |
| Write Time (ns) | < 1 | 50 | 5 | $10^5$ | 100-300 | 5-20 | 5-10 | 10-100 |
| Number of Rewrites | $10^{16}$ | $10^{16}$ | $10^{16}$ | $10^4$-$10^5$ | $10^8$-$10^{12}$ | $10^{15}$ | $10^8$-$10^{12}$ | $10^6$-$10^{10}$ |
| Read Power | Low | Low | Low | High | Low | Low | Low | Medium |
| Write Power | Low | Low | Low | High | High | Medium | Medium | Medium |
| Power (other than R/W) | Leakage Current | Refresh Power | Refresh Power | None | None | None | None | Sneak Leakage |

# The compute node




12-14Tflops

Memory BW
>1Tbyte/s

Network BW
400GB/s

Power
<200W

# Summing up

- I know the memory system will be a key to success in exascale computing
- I don't yet know what it should be, but
- It's clear that DRAM/NVRAM hybrids are very interesting
- Memory is a tough business. Twofold variability of demand! Slow growth. Narrow margins.
- Physics and engineering can produce breakthroughs
- HDDs are a "sitting duck" target

# Looking back: 1974

- Nixon out
- IRA bombs the Tower of London
- 55 mph speed limit in USA
- Patty Hearst kidnapped
- Inflation > 10 %
- Hair
- Rob (and Bobby Schnabel, Franklin Luk, Tom Howell, Andy Sherman) summer at ANL-MCS

## 30 years of parallel computing…

*Were they visionaries or lucky gamblers?*

*Moore's Law*

*Amdahl*

*Signal was there to see, despite considerable noise*