MINT Workshop, Oct. 2017
University of Alabama

Zooming in on Data Storage

Roger Wood
Western Digital Fellow
(retired)

[previously presented at the 25th ASME Annual Conference on Information Storage and Processing Systems (ISPS), June 20-21, 2016, Santa Clara Marriot, Santa Clara, California]

The Superb Hard Disk Drive

- Where to begin? How to tell the story?
- by History?, by Technology? by Personalities?
- by Size or Scale?

The Superb Hard Disk Drive

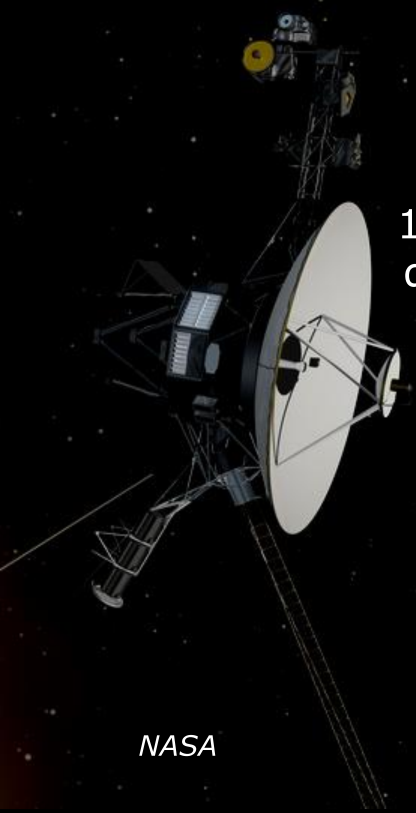
- Where to begin? How to tell the story?
- by History?, by Technology? by Personalities?
- by Size or Scale?: Yes → **ZOOM!** by factors of ten
- where to end: 1 Angstrom
- but where to start?

  - ***what is the largest distance over which humans have control and feedback?***

Voyager I

Exited Heliosphere and entered interstellar space Aug 25, 2012

← 38,000 mi/hr (16 km/s)



NASA

12 ft dish

Information Storage and Processing System

- **Tape: ½-inch x 100 ft**
  8-tracks, serpentine
  Capacity: 67 MBytes
- **Plated-wire memory**
  32 kWords x 18 bits

Voyager I records 48 seconds of data on tape once a week. The data is played back to earth every six months.

Launched Sept. 5, 1977 (41 years ago)

2.1×10^{13} m = 20 hours (0.002 light years)

Path loss = 320 dB

21 Watts
8.4 GHz
7.2 kbit/s →

← 20 kW
2.1 GHz
16 bit/s

NASA deep space network (several 210 ft dishes)

Mission ends in ~2025 when power from plutonium-238 thermoelectric generator will not be enough to run any scientific instruments

Galileo

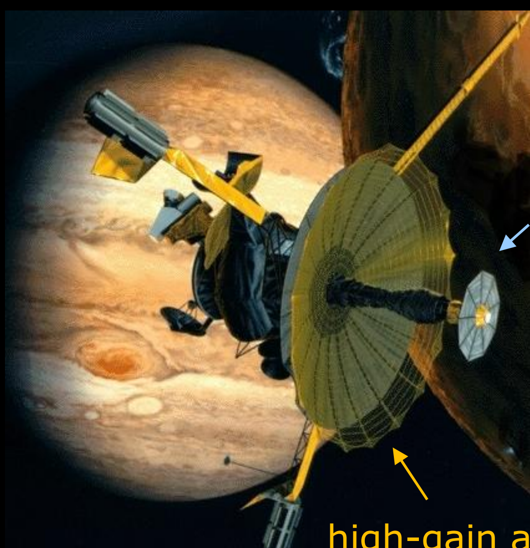
Fly-by data was stored on the tape-recorder and subsequently transmitted using the low-gain antenna at 10-40 bit/sec

Information Storage and Processing System
Tape Recorder: 900 MB

counter-rotating spools

Launched
Oct. 18, 1989
(29 years ago)

Galileo flyby of Io
(*Jupiter's moon in foreground*)



low-gain antenna

high-gain antenna
(12 ft dish)
failed to unfurl

NASA/Germany

0.8x10¹² m = 45 mins

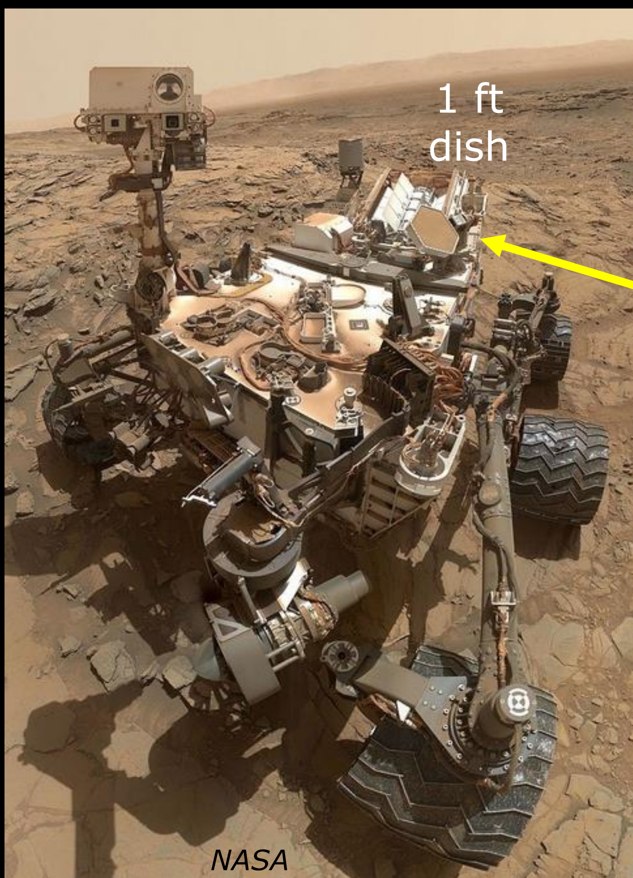
Tape recorder broke down twice:
(diagnosed and fixed remotely)
1. stiction on dummy head
2. radiation-damaged LED
(repaired with anneal cycles)

Sept. 21, 2003
final plunge into
Jupiter atmosphere
(electronics got fried)

Curiosity

Traveled ~18 km exploring Mt. Sharp since landing in Gale crater (*no Martians sighted yet*)

'Selfie' of Curiosity Mars Rover



Information Storage and Processing System

- **256 MBytes DRAM**
- **2 Gbytes Flash**
  (+ 16 GB on the cameras)

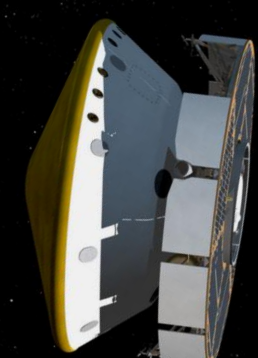
No HDD or Tape ☹
no magnetometer?

$2x10^{11}$ m = 10 minutes

Launched Nov. 26, 2011 (7 years ago)

Landed on Mars Aug. 6, 2012

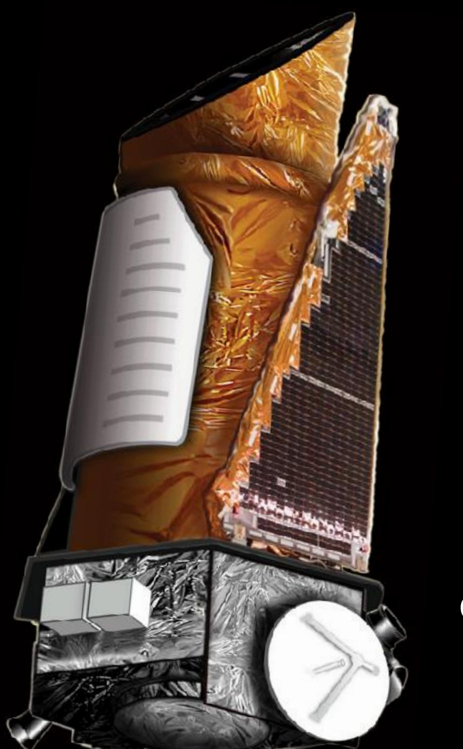
landing via heat-shield, parachute, & sky-crane



Spacecraft configuration during voyage

Kepler

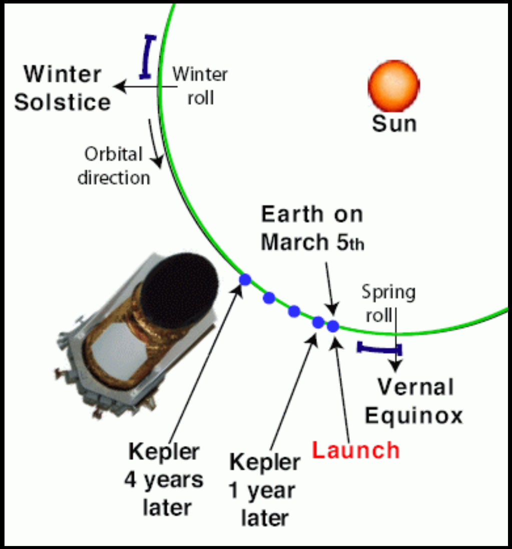
Kepler space telescope has found 3,743 exoplanets as of Mar 8th 2018
(no evidence of HDDs on other planets yet)



NASA

Information Storage and Processing System
Solid-State: 16 Gbyte
~60 days of data

No HDD or Tape ☹



earth-trailing orbit

gradually drifting further behind

Launched 9 years ago

March 7, 2009

3 ft dish

~10^{10} m = 30 seconds
(has since drifted to ~1 AU behind Earth)

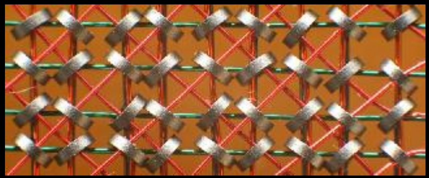
Apollo 11 Lunar Lander

Neil Armstrong and Buzz Aldrin stepped onto the moon on July 21st 1969
(and don't forget Michael Collins)

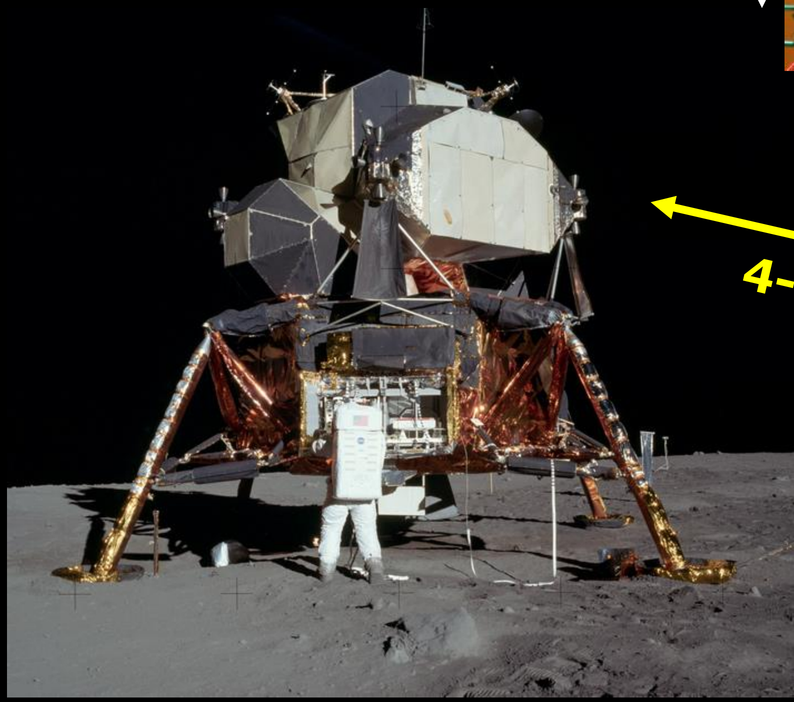
Information Storage and Processing System
4 kBytes core memory

No HDD or Tape but ferrite cores are still magnetic recording ☺

1 mm ↕



Launched
July 16, 1969
(49 years ago)



0.38×10^9 m = 1.3 seconds

4-day spaceflight each way

next humans on the Moon?
China by 2025?

returned
July 24, 1969,

Syncom 3

Information Storage and Processing System
Probably just a handful of bistables and latching relays?

First Geostationary Satellite

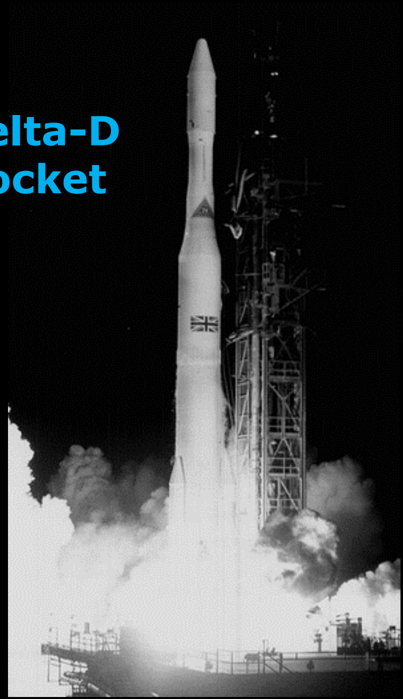
(orbits at the *equator* every 23.93 hours)

Telecast the 1964 Summer Olympics in Tokyo to the United States

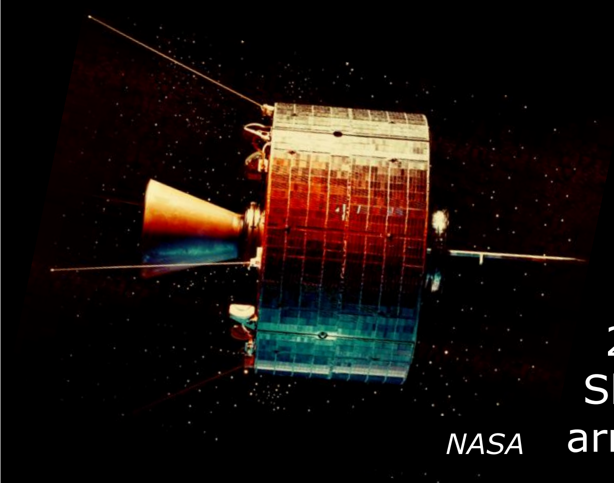
Geosynchonous orbit occurs at an altitude of 35,786 km (22,236 mi)

First popularized by Arthur C. Clarke in 1945
(SciFi: "2001, A Space Odyssey")

Delta-D rocket



(Delta-M Launch of British ***'Skynet'*** *satellite in 1969)*



NASA

0.42×10^8 m = 0.14 seconds

2 Watt TWT
Slotted-dipole array: 6dB gain

Launched 54 years ago

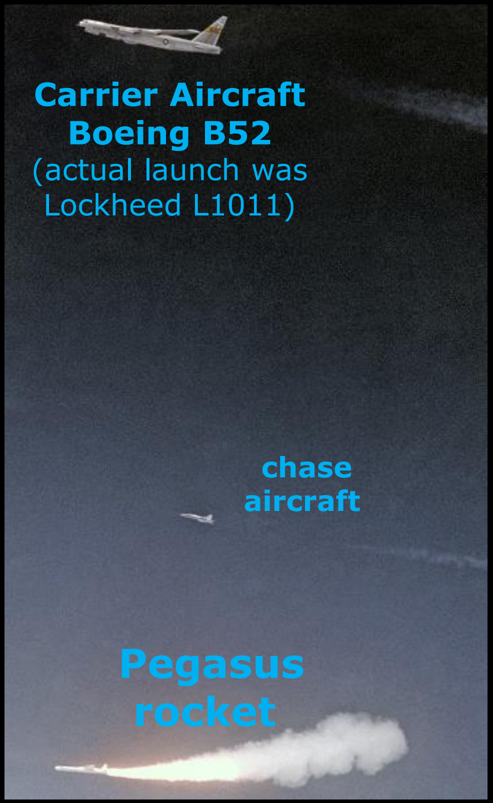
Aug. 19, 1964

MSTI-3

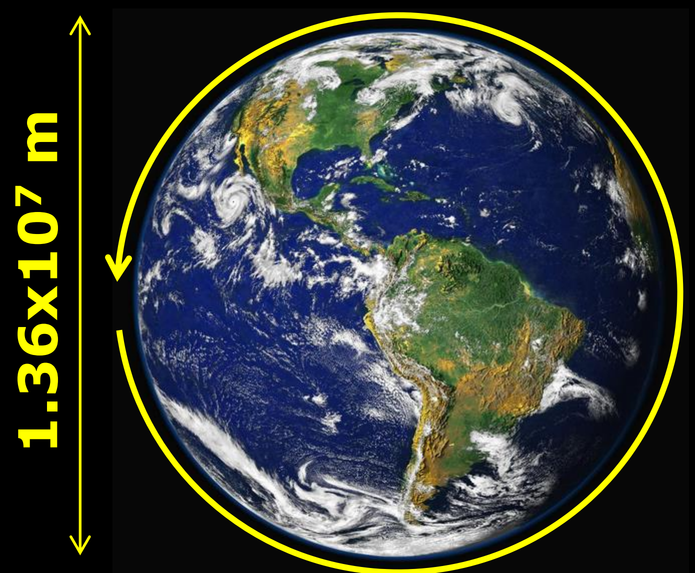
Only HDDs ever launched on a spacecraft
(except on Intnl. Space Station)

Information Storage and Processing System
2 x **Conner 500 MB HDDs**
(counter-rotating configuration)

Conner Peripherals CP-3540
Half-High 3.5" Form-Factor
4500 rpm, 6 disks, 2.5 MB/s



MSTI = Miniature Sensor Technology Integration (*for tracking thermal profile of ballistic missiles*)



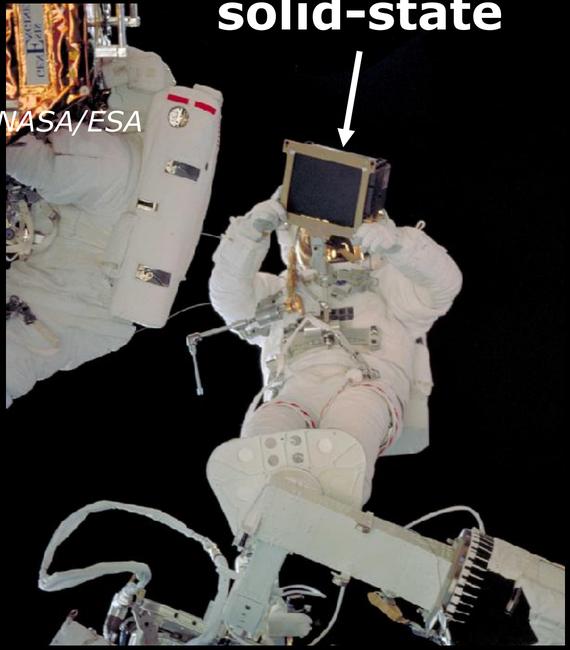
circular polar orbit 425 km altitude
(Orbit diameter = 12713+2*425 km = 1.36×10^7 m

US Air-Force

Launched from Vandenberg AFB
20 years ago
May 16, 1996
till Dec 1st 1998

Hubble

Feb 1997
tape-recorder replaced with solid-state



Information Storage and Processing System
***was* 3 Tape-recorders each 1.2 GB capacity**

Replaced with 3 x 12 GB solid-state radiation-hardened memories during 2 of the 5 service missions

launch from Shuttle cargo bay

Space-Shuttle Discovery Mission STS-31

Hubble orbits at altitude of 560 km = 0.56×10^6 m

"Pillars of Creation"

Eagle Nebula

Infamous for its fuzzy vision (1st service mission fixed it in Dec 1993)

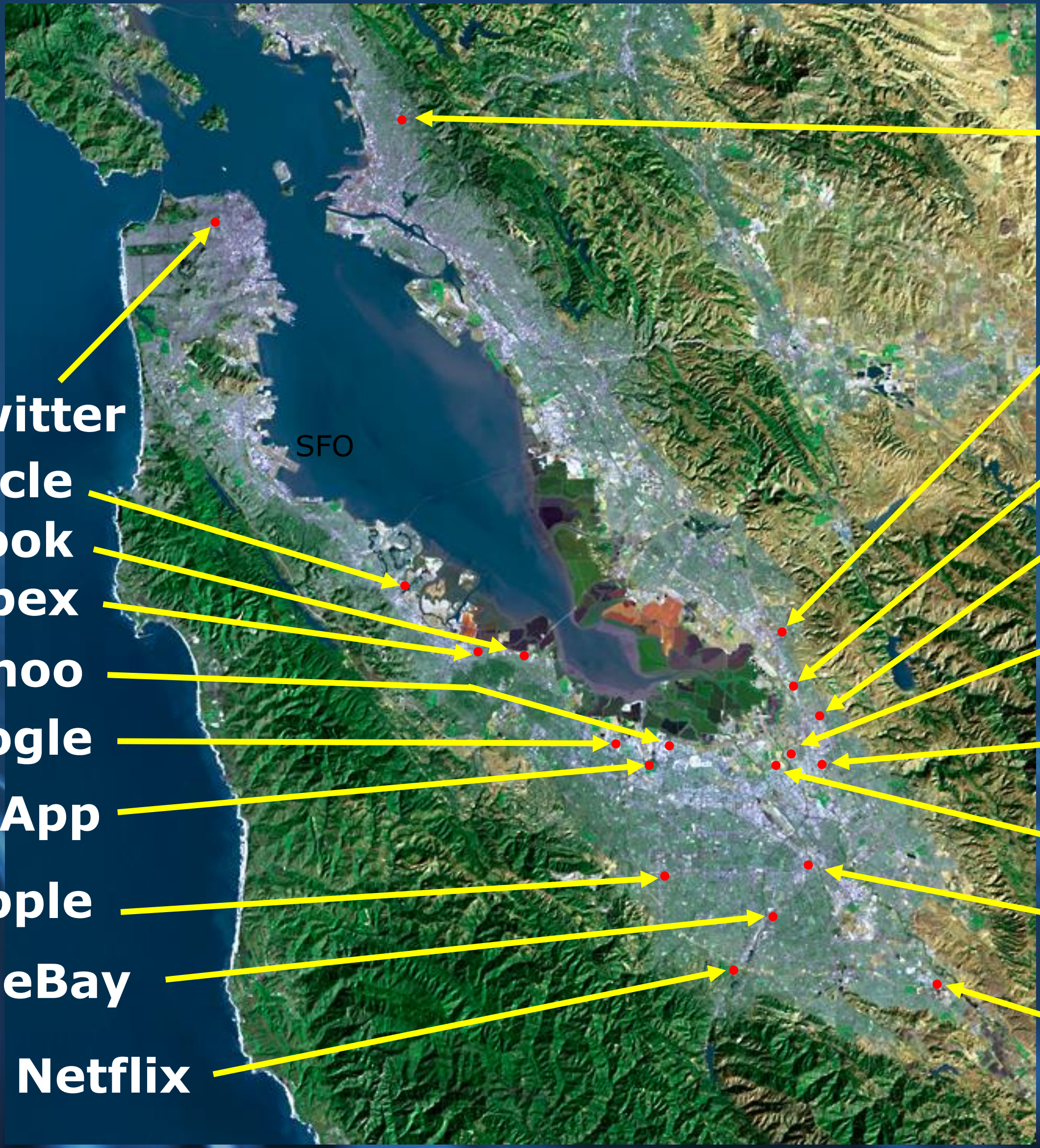
Launched on space-shuttle 26 years ago Apr. 24, 1990

(originally planned for 1983)

San Francisco Bay
(Storage Industry)

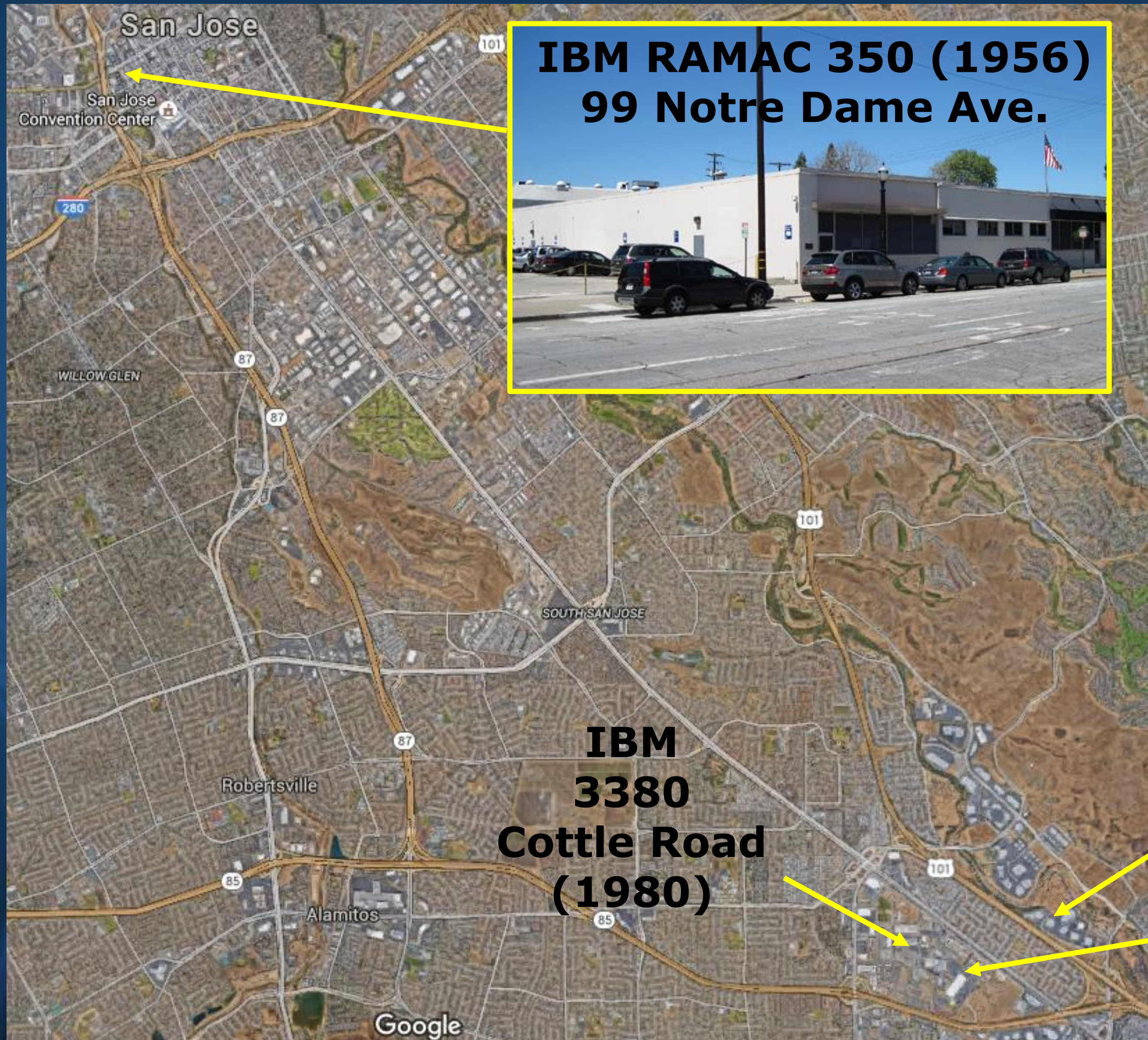
Scale
~100 km
~10^5 m

Finally, Down to Earth!



San Jose (HDD History)

Scale
~10 km
~10^4 m

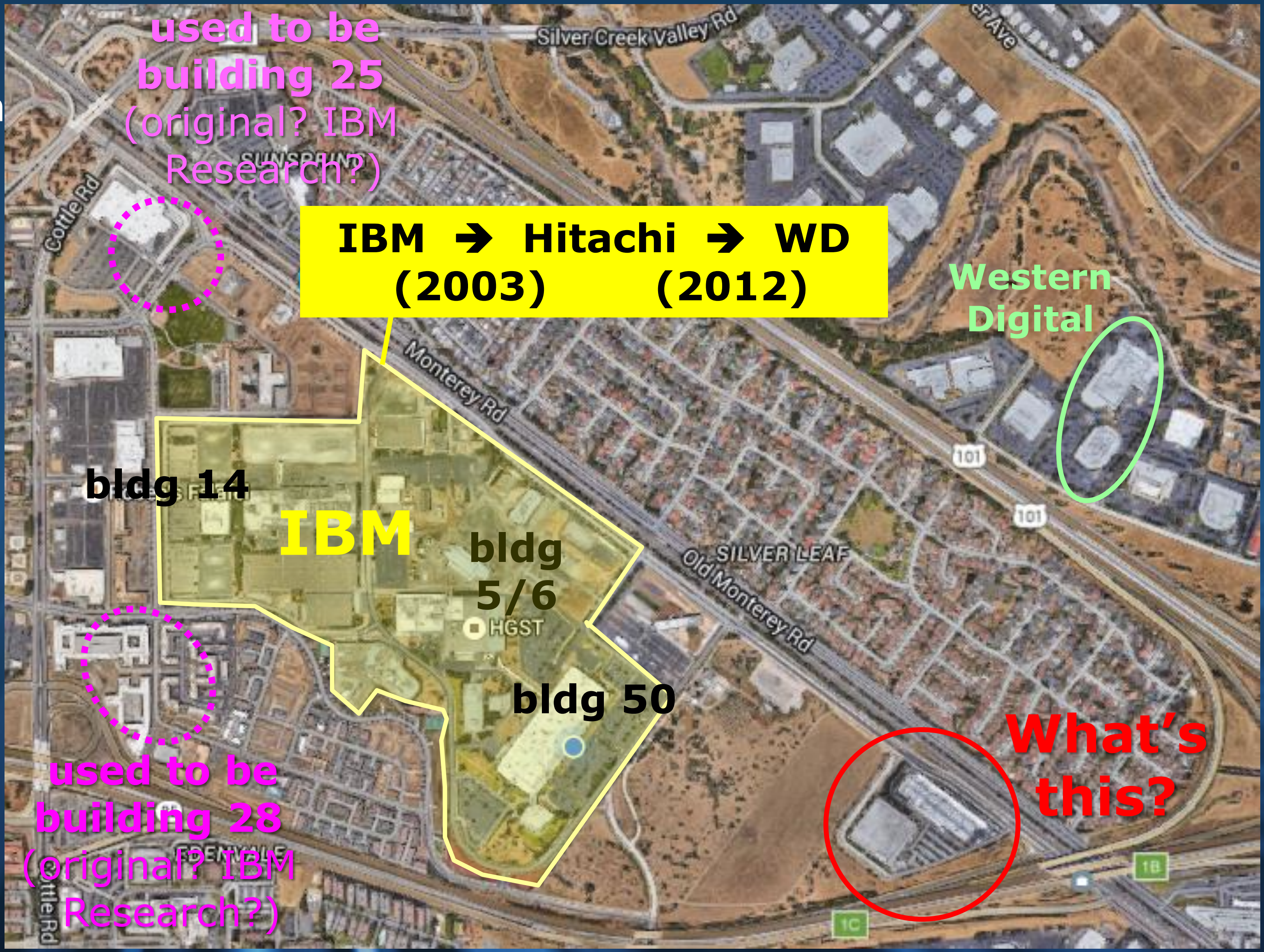


Western Digital

IBM/ HGST

Nostalgia: old IBM site

Scale
~1 km
~10^3 m



Data Center
9/11 Great Oaks Parkway

Scale ~100 m
~10^2 m
(soccer pitch)

Equinix SV1 facility

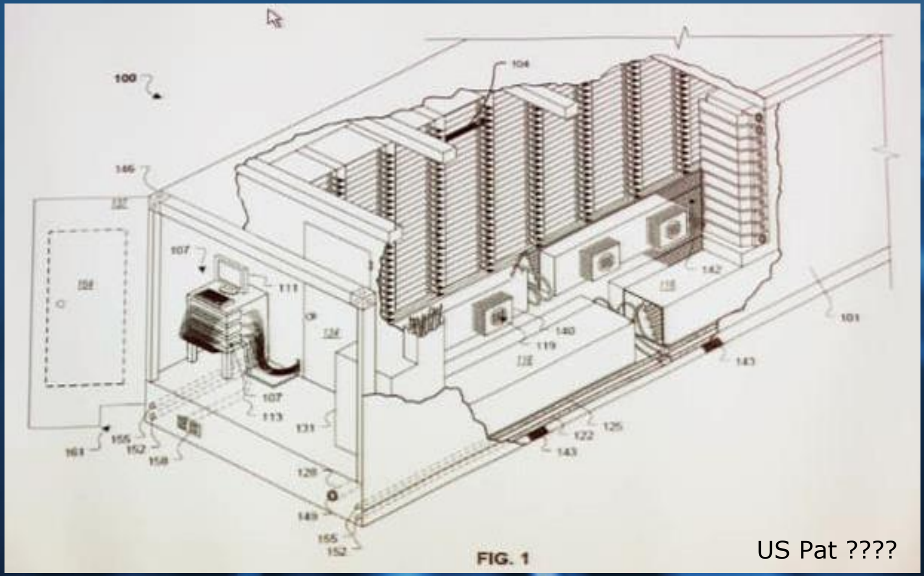
- **83,000 ft²** **air-conditioned**
- **10 MW back-up power**
- **Uptime 99.9999%**
- **Seismic Category IV** (hospitals/fire/ police, etc)

Google Container Storage

Scale ~10 m
(shipping container)

Modular storage based on standard shipping container
(40 x 9.5 x 8 feet)

https://www.youtube.com/watch?v=zRwPSFpLX8I



US Pat ????

"Standard 1AAA shipping containers packed with 1,160 servers each, with many containers in each data center."

cnet.com 2009

Big HDDs and Racks

IBM RAMAC 350 → IBM 3380 → WD SA7000

Scale ~1 meter
(person-sized)

IBM RAMAC
1956, < 5 MB
50 disks, 2 heads

24 years
500 x

IBM 3380
1980, 2.5 GB

35 years
2 million x

"Active Archive"
2015, 4.7 PB

Al Hoagland, Computer History Museum, Mountain View

1960 Olympic Winter Games
Squaw Valley (USA),

Finnish Data Processing Museum, Jyväskylä, Finland

2 actuators/spindle
2 spindles/box

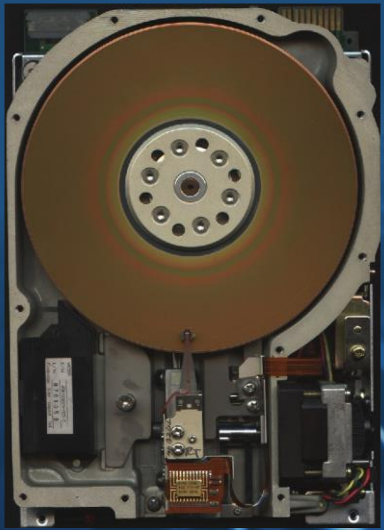
http://www.hgst.com/

98 actuators
98 spindles

‘Small’ HDDs

Scale ~10 cm
10^{-1} m
(hand-sized)

1st 5¼” HDD
Seagate ST506
1980, 5 MB



stepping motor
+ linear-actuator
170 ms ave. access

1st 3.5” HDD
Rodime RO352
1983, 12.75 MB

96 mm disk
(interesting history)
Rodime ➔ Littlewoods!

33 years
1 million x

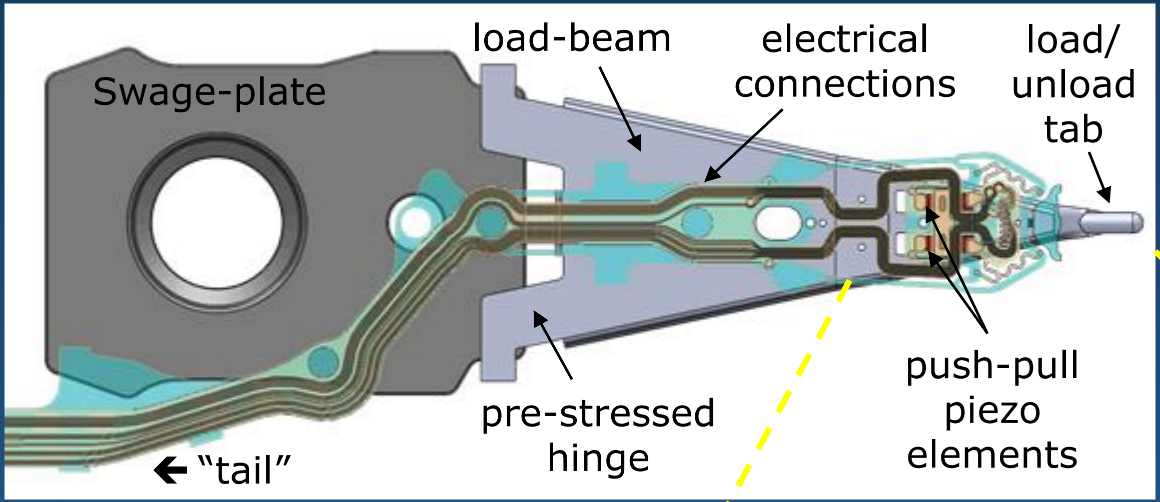
‘Modern’ 3.5” HDD:
UltraStar HE10
2016, 10 TB
7200 rpm, 1” high
7 thin-film PMR disks
14 spin-valve TMR heads
Helium sealed
Rotary Actuator
+ Micro-actuator
Fly-height actuator

95 mm disk
rotary actuator
12 ms ave. access

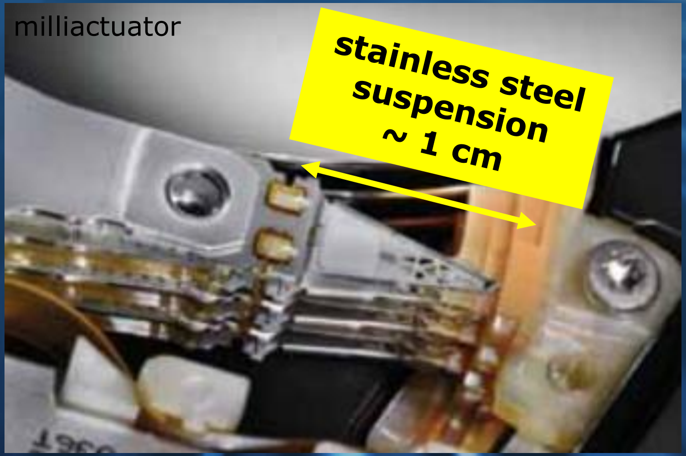
Suspension and Microactuator
(HGA, Head/Gimbal Assembly)

Scale
~ 1 cm
(finger-nail)

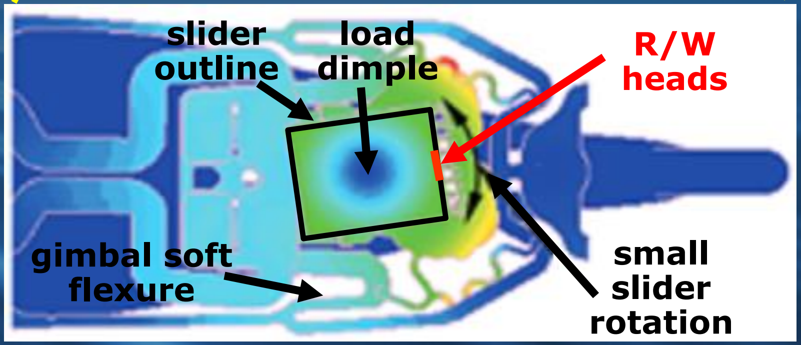
Suspension carries control signals and ~3 Gb/s data over lines with carefully controlled impedance



Suspension loads the slider against disk and includes gimbal to allow slider pitch & roll



Heads shown parked on loading ramp



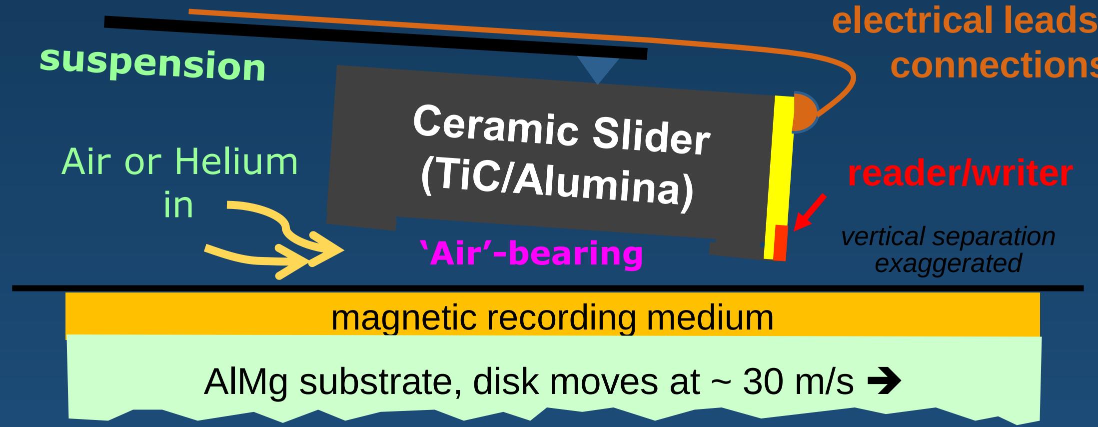
HGST Microactuator, White Paper, Oct 2015

tiny stroke but very high bandwidth (Main Resonance ~40 kHz)

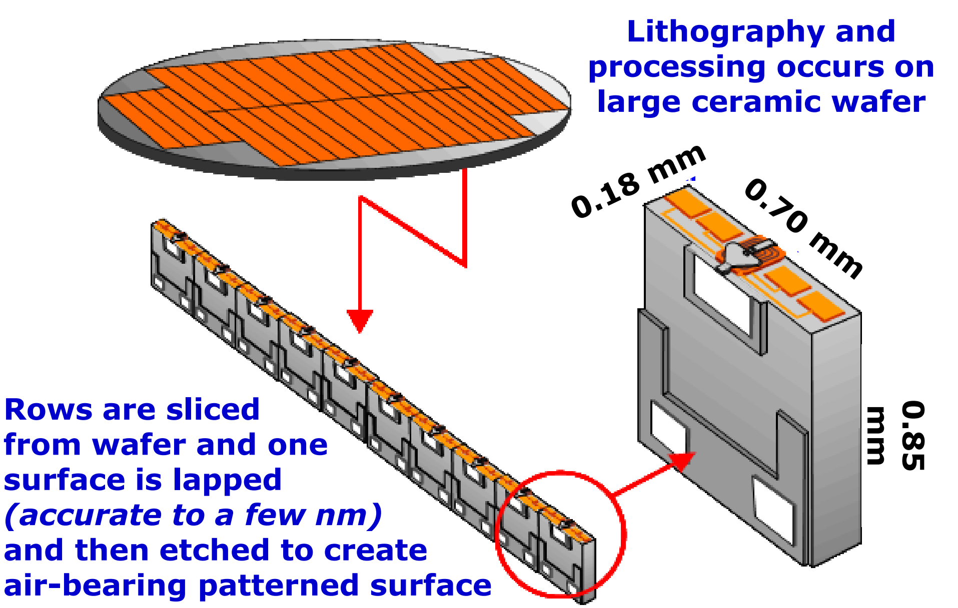
Slider

Scale ~ 1 mm
~10^{-3} m
(coarse sand)

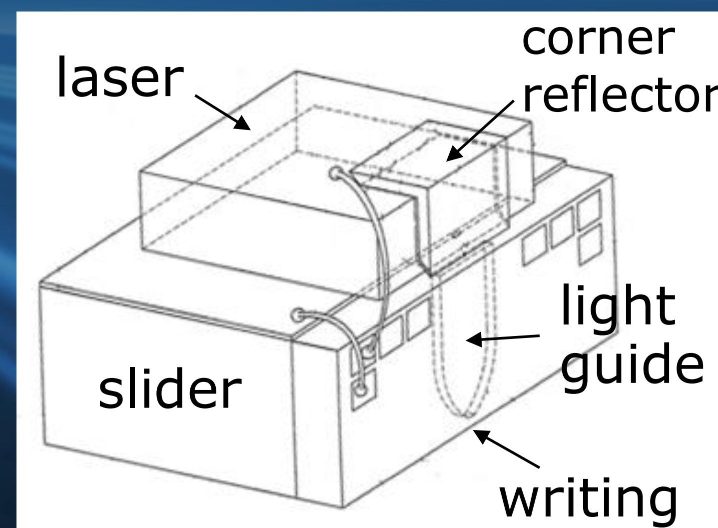
micro-femto-slider dimensions
0.85 x 0.7 x 0.18 mm
(Smaller than ball in ball-point pen)



~120,000 sliders/wafer on an 8" wafer



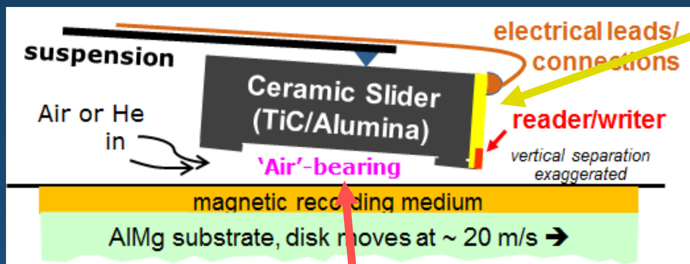
HAMR
heat-assisted magnetic recording
laser mounted on slider



Air-bearing, Head Connections

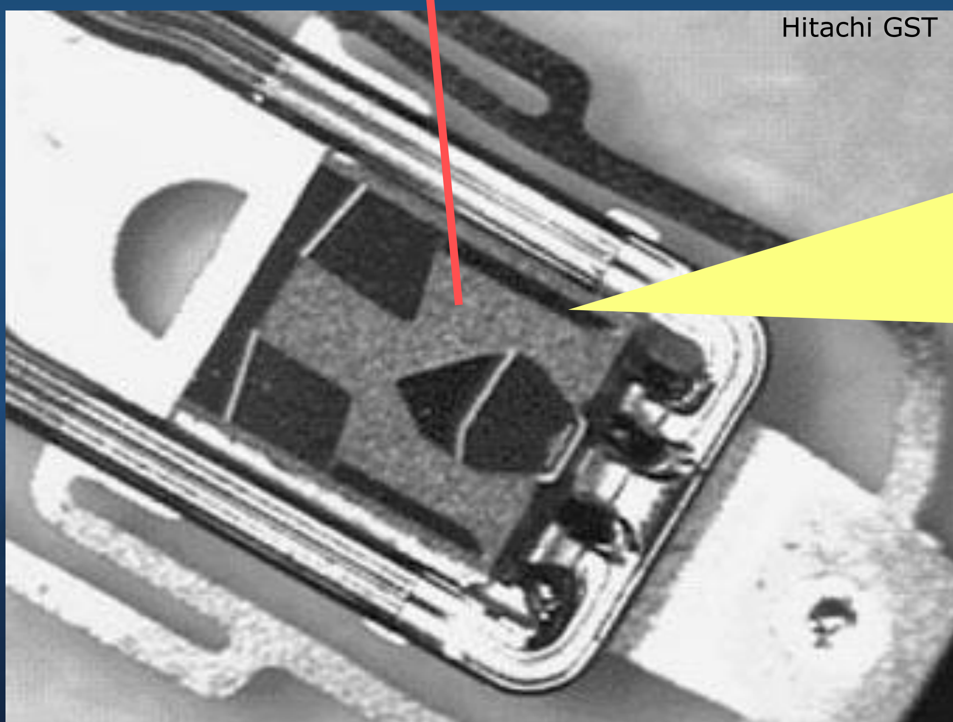
Scale ~ 100 μm
~10^{-4} m

human hair



Back of Slider (Wafer Surface)

Bonding pads and copper and NiFe leads are all on a sub 100 μm scale and occupy almost all the area on the back of the slider



'Air'-bearing

Designed to fly reliably with tiny, controlled separation but highly compliant with disk surface.

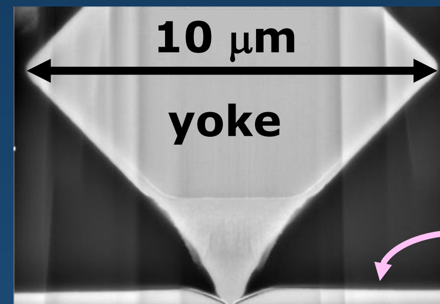
Design must operate over full actuator stroke (skew/ velocity), and tolerate a wide range of ambient pressure, temperature, and humidity and possibly different atmospheres (Air/He) and rpm (during servo-write)

Head Magnetic Structures

Scale ~ 10 μm
~10^{-5} m

(red blood cell
~8 μm)

write-head magnetic yoke

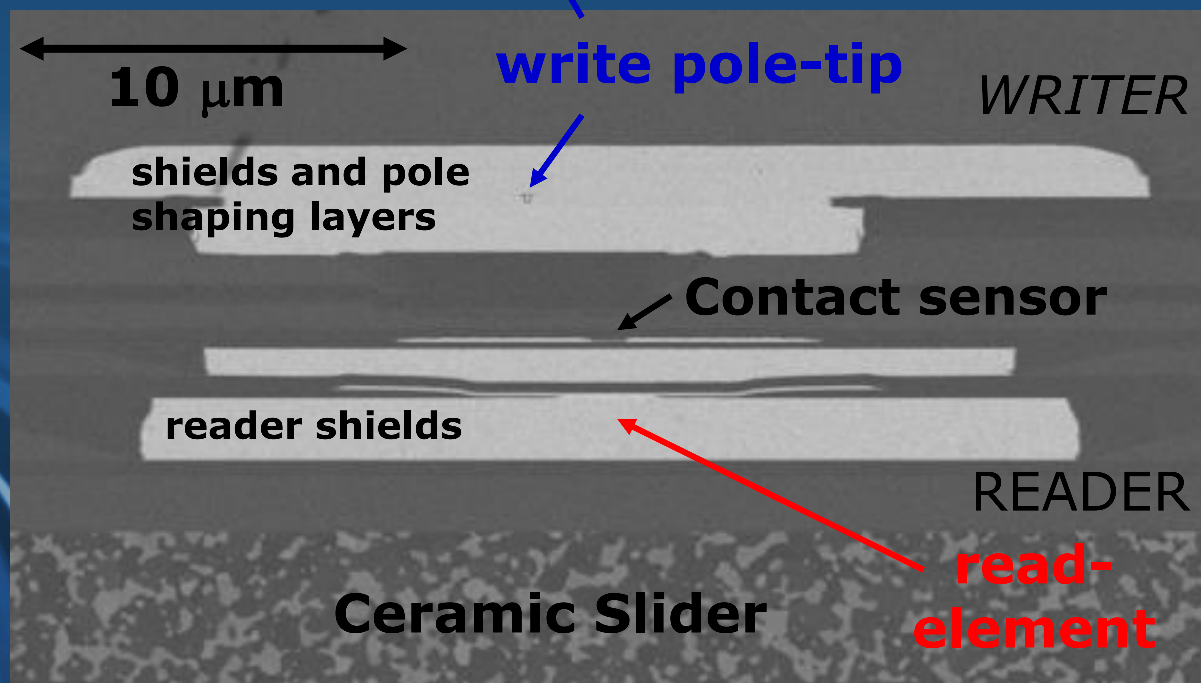


perspective from back of slider

side-shield

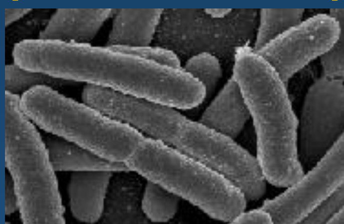
air-bearing surface ------

perspective from air-bearing

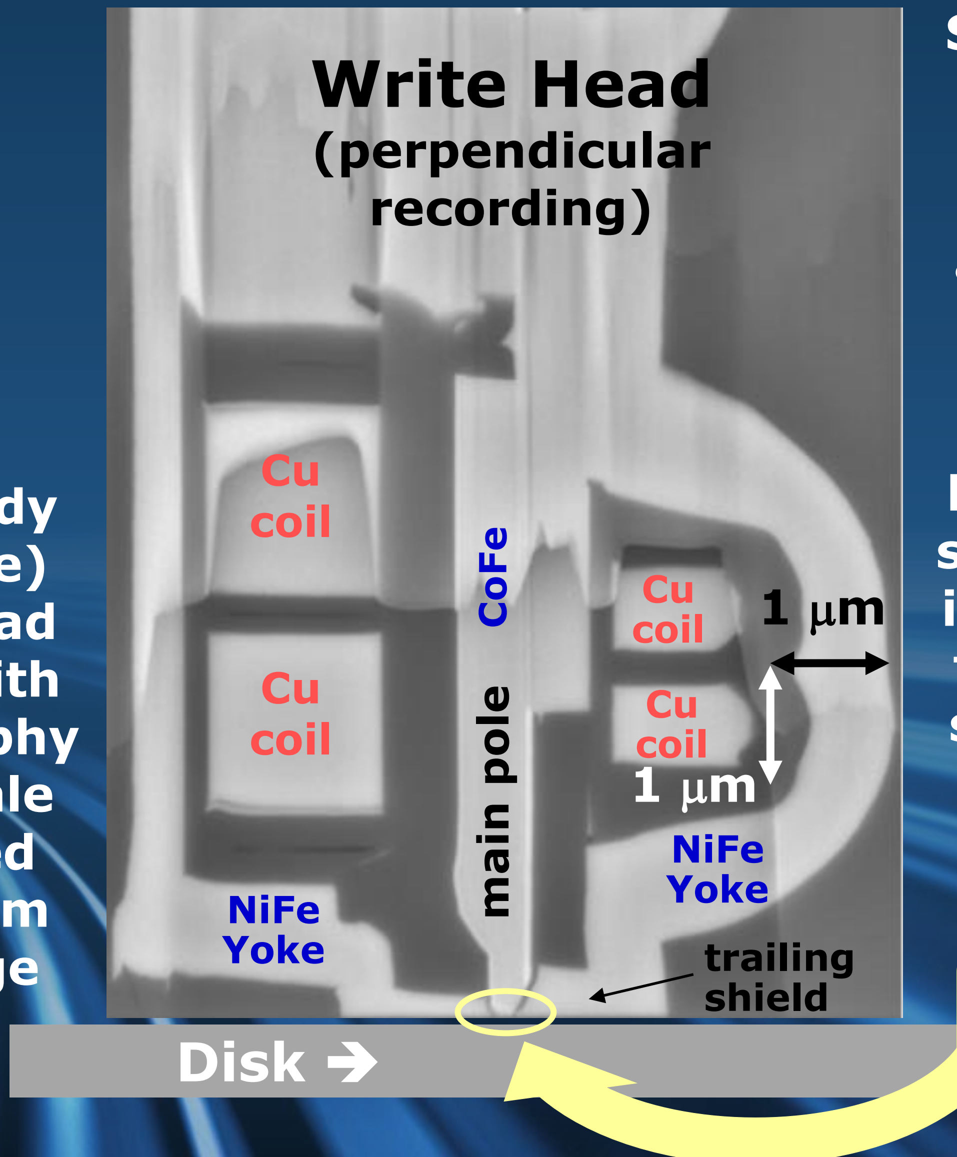


Write Head

Scale ~ 1 μm
~10^{-6} m
(bacteria-size)



Much of the body (magnetic yoke) of the write head is fabricated with optical lithography at the 1 μm scale and with plated films in the 1 μm thickness range



Small yoke and 2-turn plated coil minimizes inductance and allows high data-rates >3 Gb/s

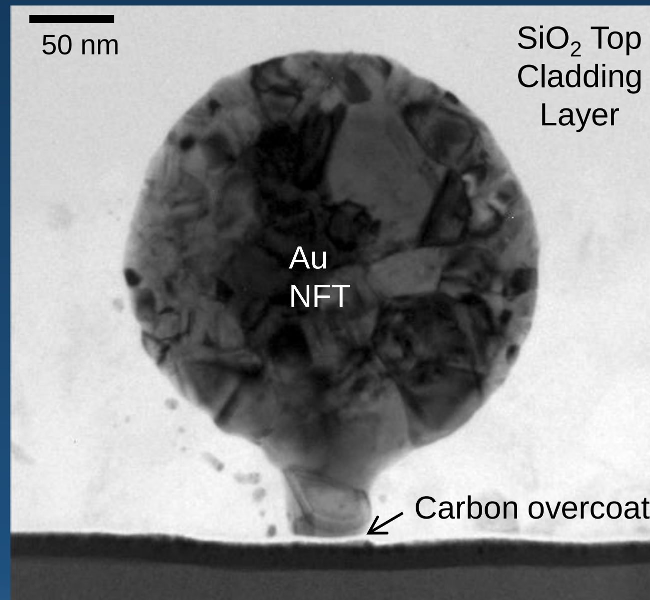
Write head uses high-permeability structures tapered in 3D to focus flux from the 1-10 μm scale down to the 10-100 nm scale of the pole-tip

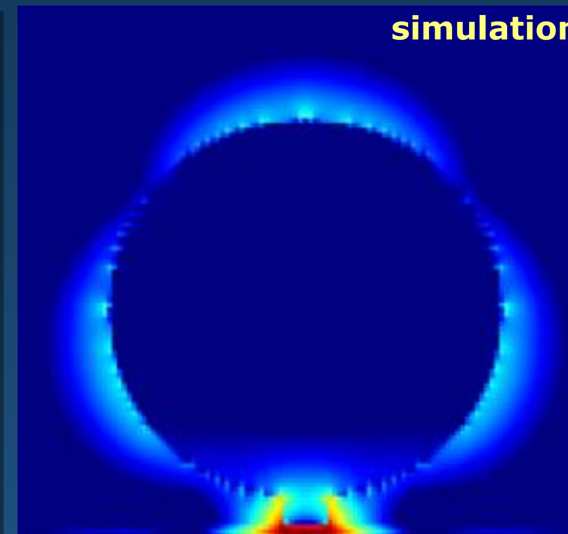
This is where 'rubber meets the road'

NFT and STO (new technologies)

Scale ~ 100 nm
~10^{-7} m
(virus-size)

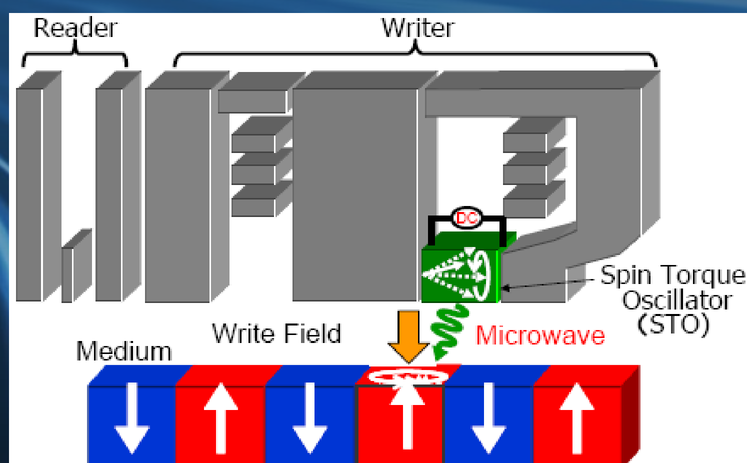
HAMR Near-Field Transducer
(plasmon resonance at optical frequencies provides localized heating of medium to assist writing)



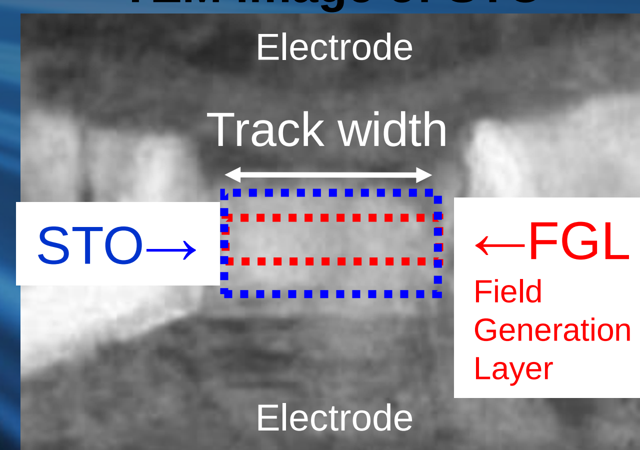


Challener et al, Nature Photonics 3, pp. 220 - 224 (2009)

MAMR Spin-Torque Oscillator
(ferromagnetic resonance at micro-wave frequencies provides localized excitation to assist writing)



TEM image of STO

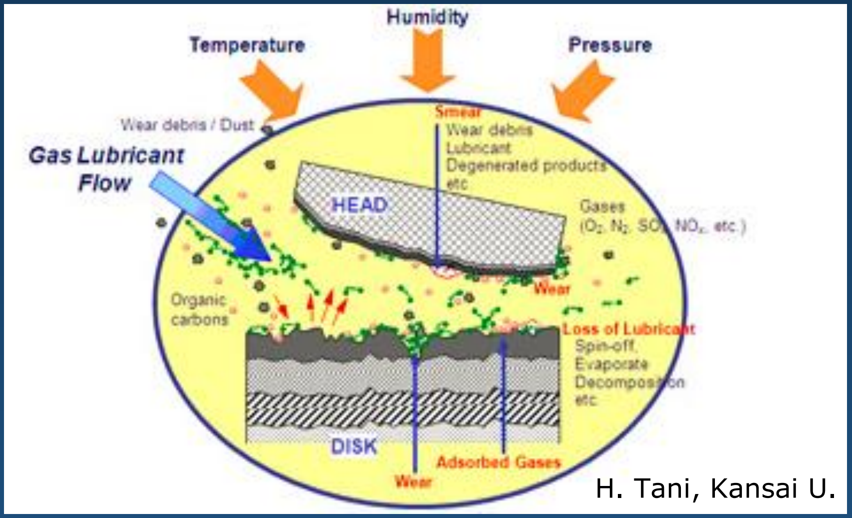


Matsubara et al, MMM 2010

Thermal Flying Control (TFC)

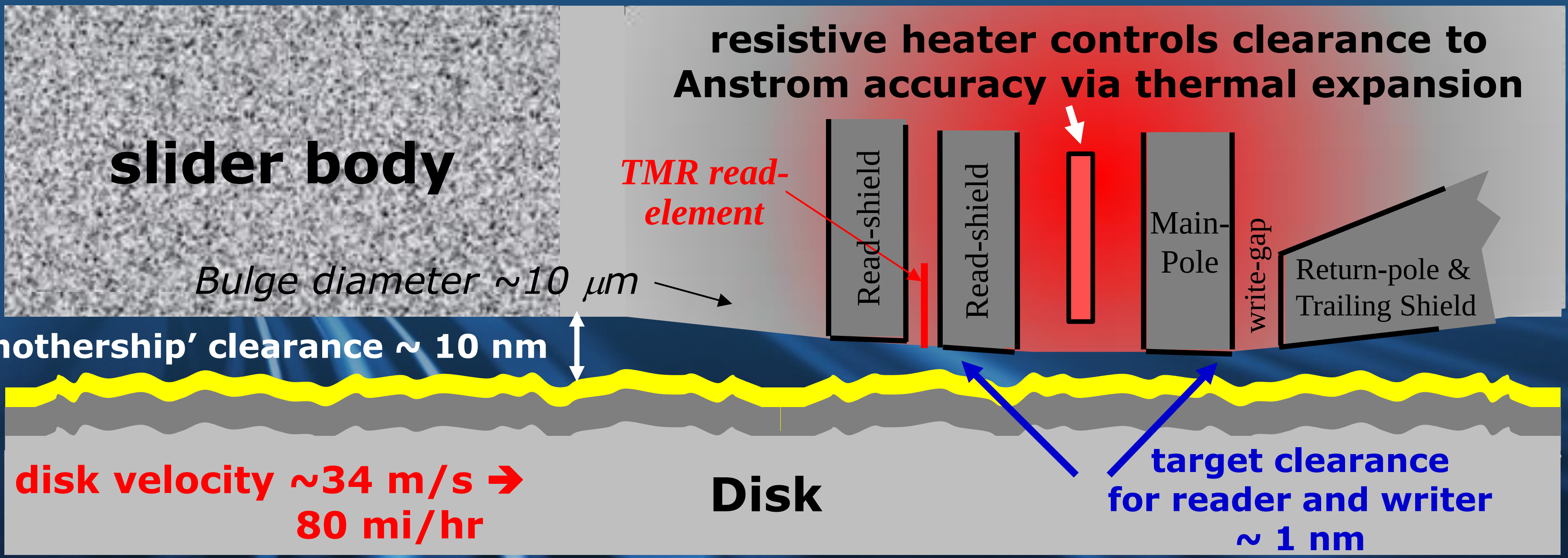
Scale ~ 10 nm
10^{-8} m
(prion-size)

many factors affect fly-height →



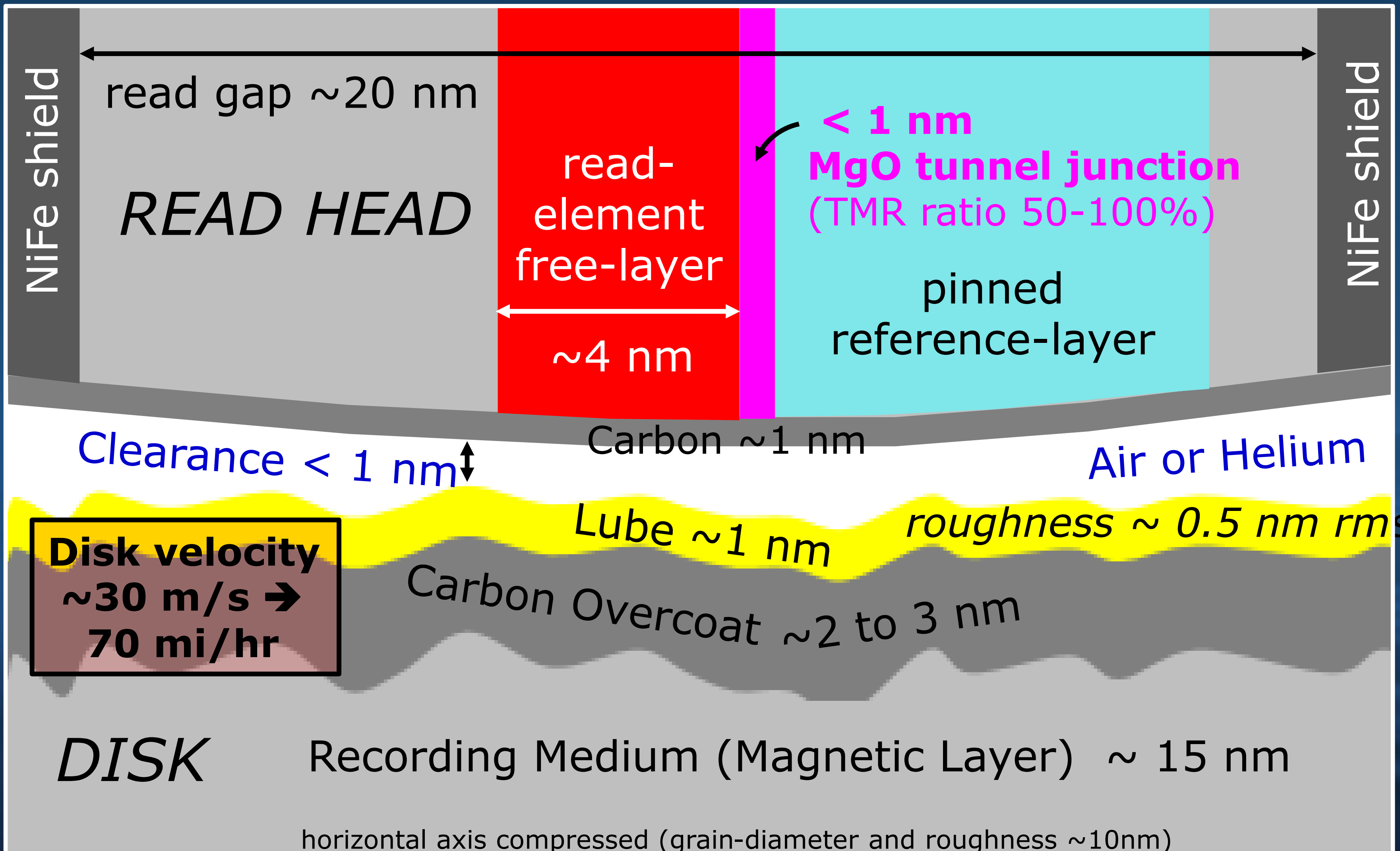
http://www2.ipcku.kansai-u.ac.jp/~hrstani/intro_e.htm

'touchdown' or head-disk contact disturbs all 3 dimensions and can be sensed from servo-position amplitude, and timing (and/or from an explicit contact sensor)



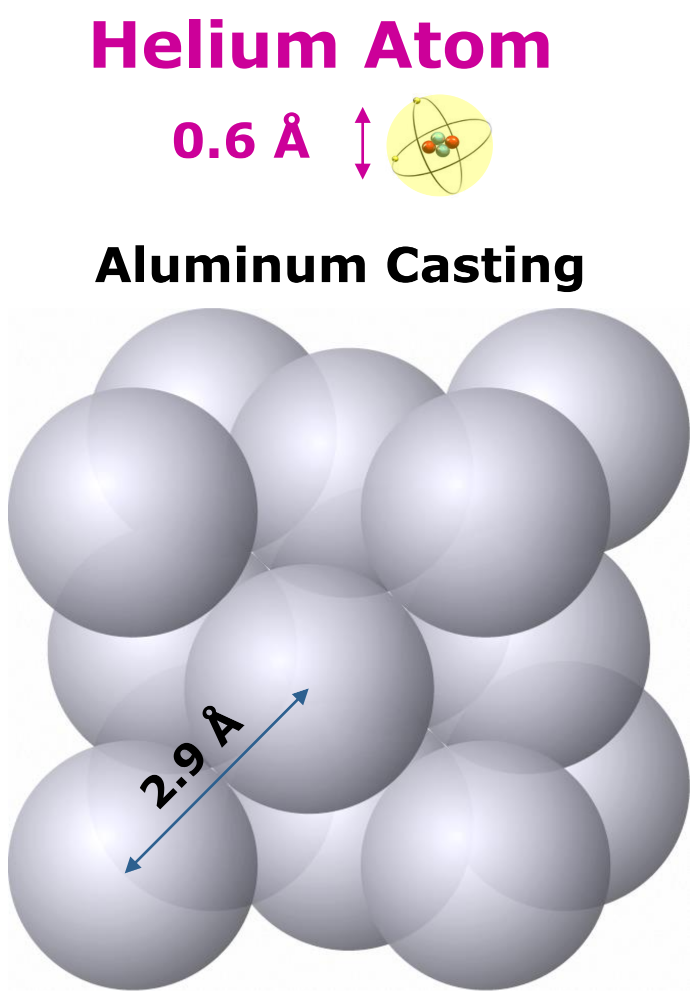
Head-Disk Interface and Tunnel-Junction Read Sensor

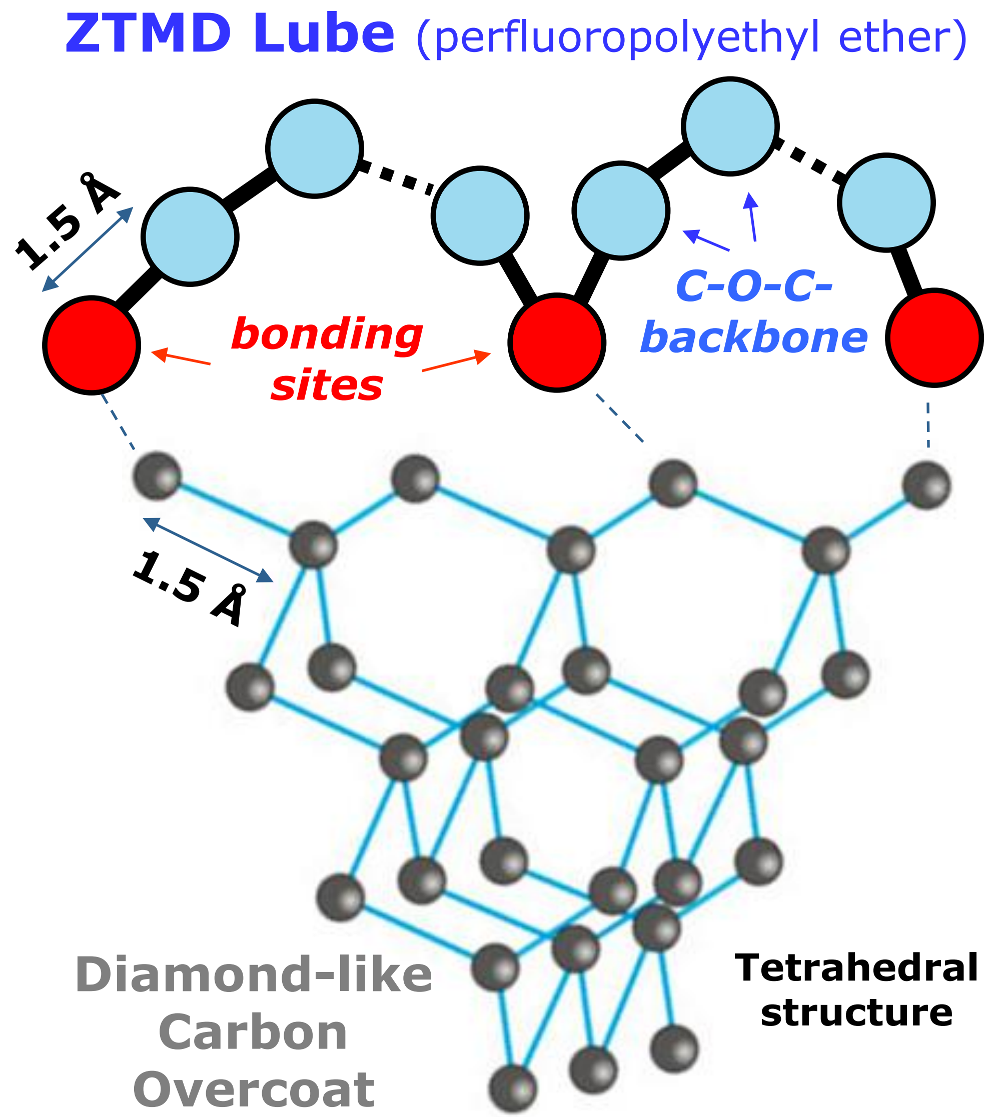
Scale ~ 1 nm



Atomic Level

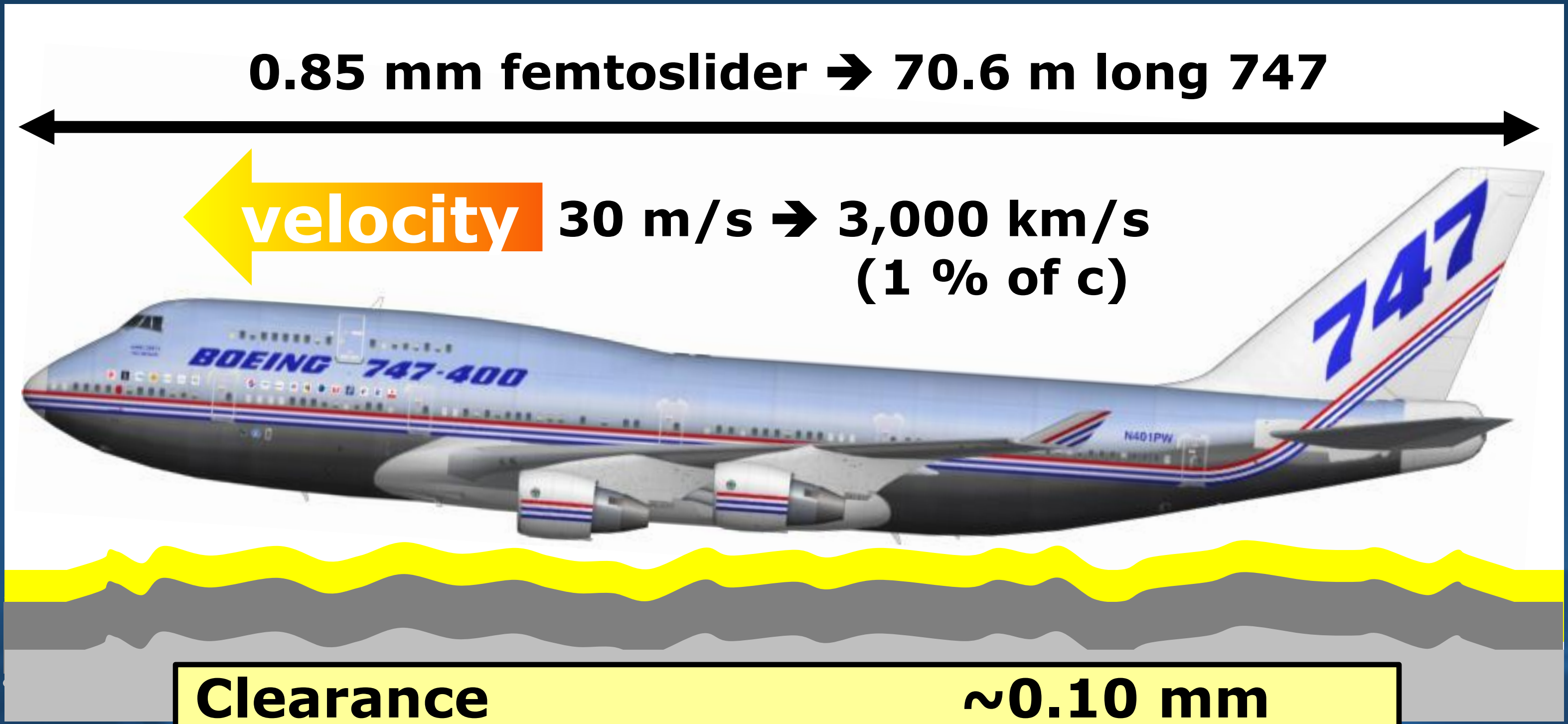
Scale ~1 Å





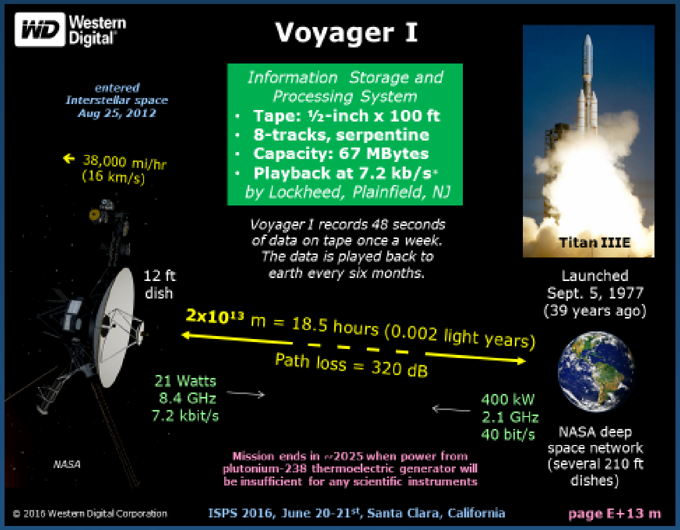
Scaled to a Boeing 747

0.85 mm femtoslider scaled to length of a Boeing 747



Clearance	~0.10 mm
Disk roughness (rms)	~0.05 mm
Disk lubricant	~0.10 mm
Disk Carbon overcoat	~0.20 mm
Media recording layer	~1.50 mm
Disk thickness	~100 m
Disk diameter	~10 km

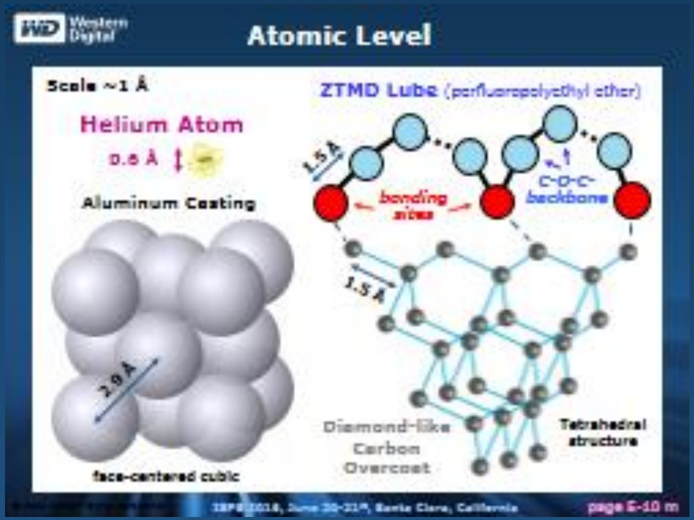
Magnetic Recording and HDD



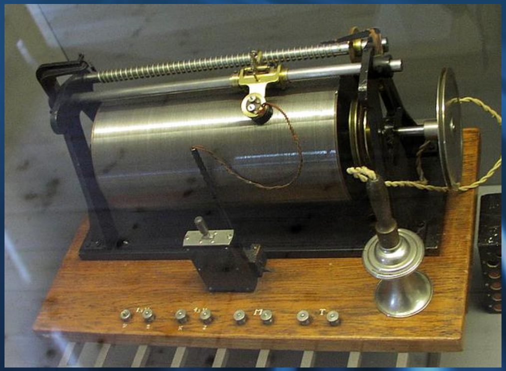
10^{+13} m

23 decades

10^{-10} m

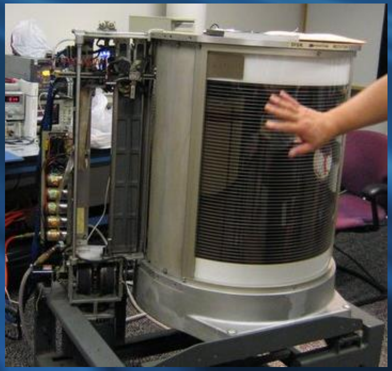


1898
(Wire recorder, V. Poulson)



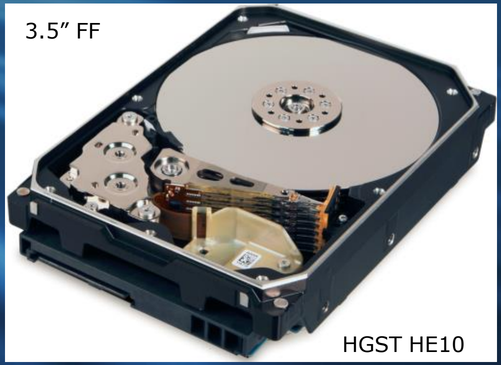
6 decades (time)

1956
(5 MB, Ramac, IBM)



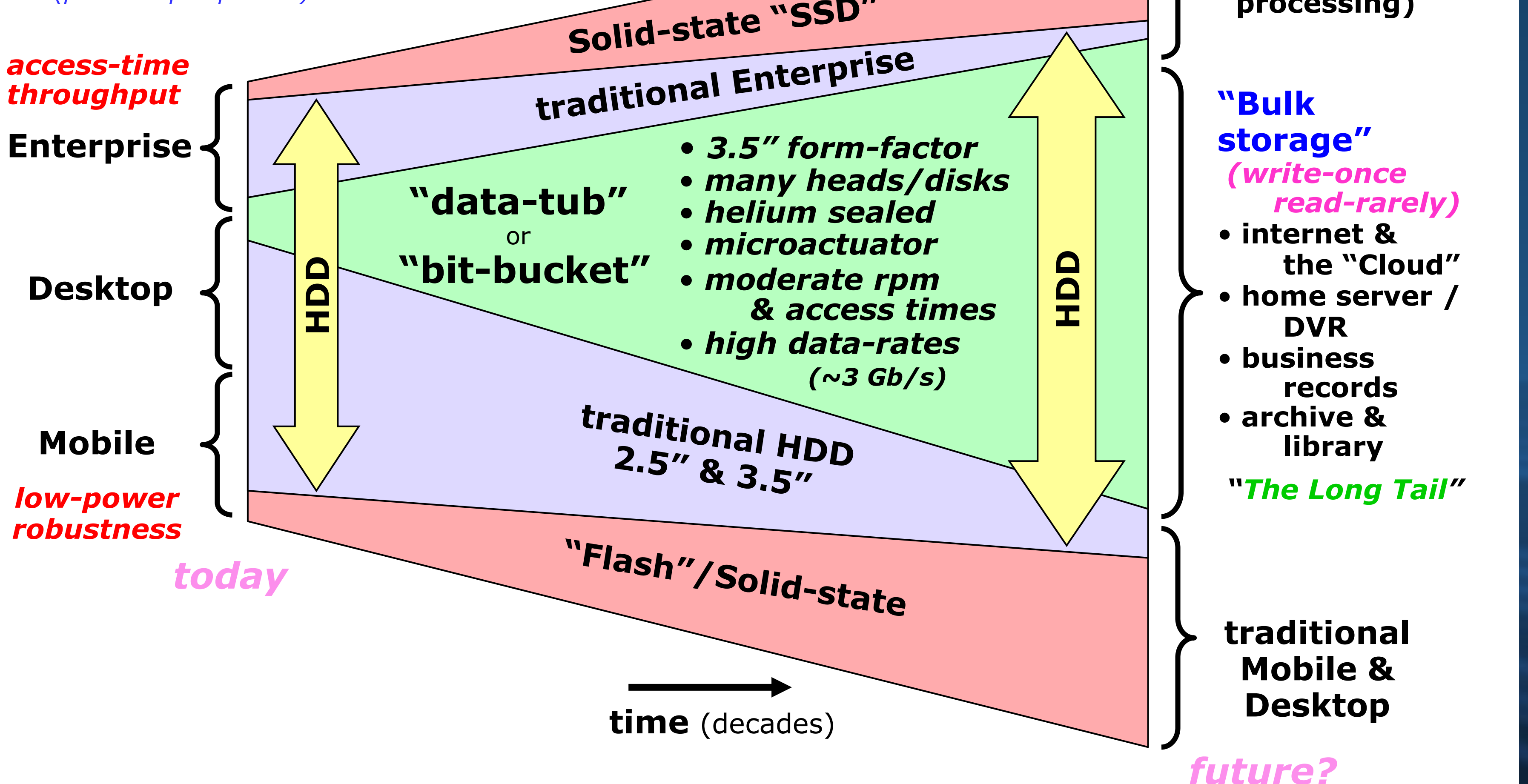
6 decades (time)

2016
(10TB, HGST/WDC)



The Future: Market Evolution

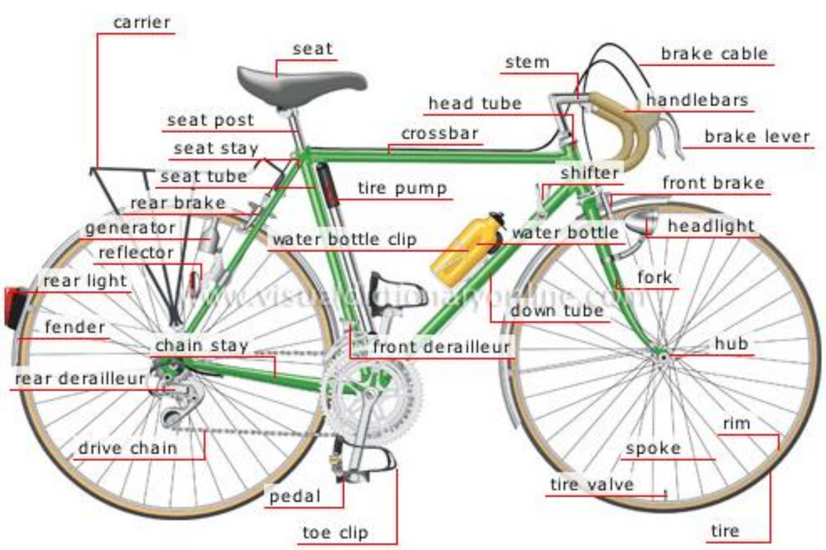
Conceptual picture showing storage capacities shifting towards solid-state and to "bulk" storage on HDD
(personal perspective)



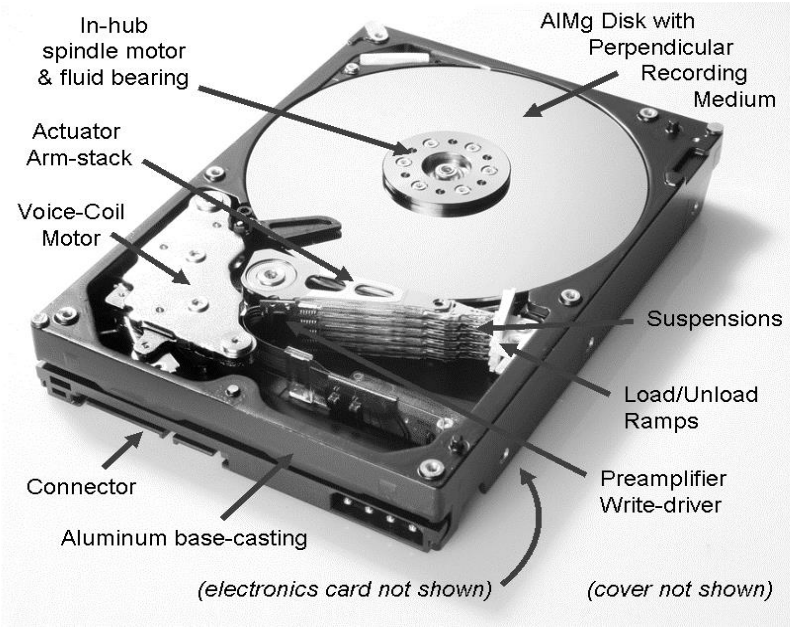
R. Wood, J. MMM 321 (2009) pp. 555–561

Perfect Inventions
(that superb hard disk drive)

Bicycle



Hard Disk Drive



- two large similar-diameter wheels
- front-wheel pivots with handlebars
- rear wheel driven through sprocket & chain from rotating pedals
- operated in seated position

- fluid-bearing spindle with multiple disks
- rotary actuator carrying multiple heads
- slider with self-generated air-bearing
- thermal μ-actuator for magnetic spacing
- perpendicular recording mode

Hard Disk drives will be here for many decades to come

G. Tarnopolsky, 4th Intnl. Conf. IGNOIE-SOIM, Jan. 23–25, 2007, Sendai, Japan, 2007

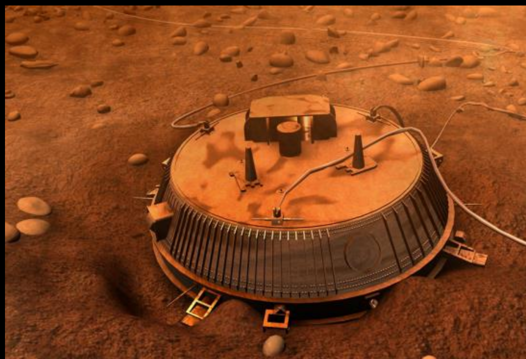
Acknowledgements

John Contreras
Qing Dai
Mark Haertling
Yoshihiro Ikeda
Wen Jiang
Ian McFadyen
Bob Reinhart
Mike Salo
Marilee Schultz
Barry Stipe
Hiroyasu Tsuchida
Monica Vargas
and many others …

"Powers of Ten" a short film by Charles and Ray Eames 1977

Western Digital®

Cassini-Huygens



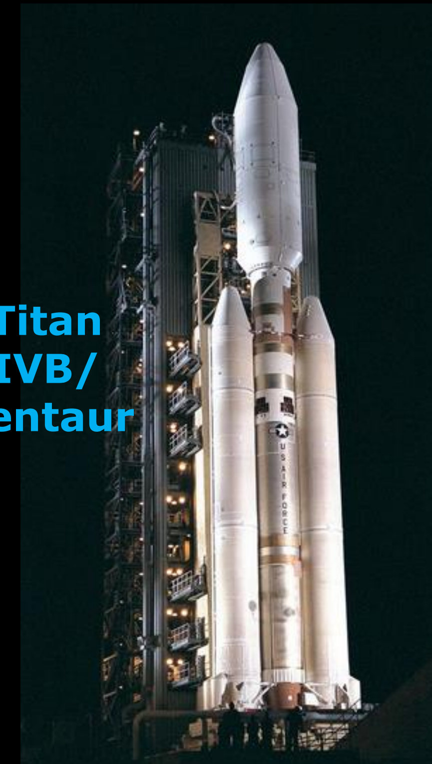
Huygens on surface of Titan

Information Storage and Processing System

- **Solid-State: DRAM** (by Oki)
- **Capacity: 300 MBytes**

ECC: Hamming (39,32,4)

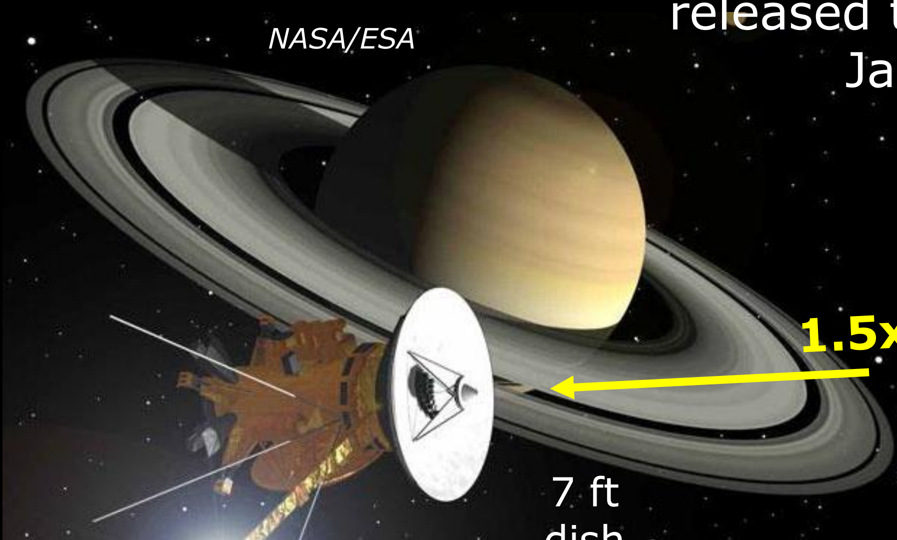
280 bit-flips/day (cosmic rays)



Titan IVB/ Centaur

Cassini was placed in orbit around Saturn July 2004

The Huygens lander was released to land on Titan in January 2005



1.5×10^{12} m = 1.4 hours

Launched 19 years ago

Oct. 15, 1997

California

Scale
~1000 km
~10^6 m

San Francisco

Stanford University

Vandenberg AFB
USA West Coast space launch site

Hollywood
World #1 in movie, TV, music production

San Diego
one of largest naval bases in world

Just Read the Instructions
Space-X landing barge

State Capital
Sacramento
(gold discovered 1848)

University of California
Berkeley

Palo Alto to San Jose
"Silicon Valley"
(also birthplace of video tape recorders and hard disk drives)

Edwards AFB
Shuttle landing site

Cal Tech
NASA Jet Propulsion Laboratory

Greater
Los Angeles
US second city
(pop. ~19 million)

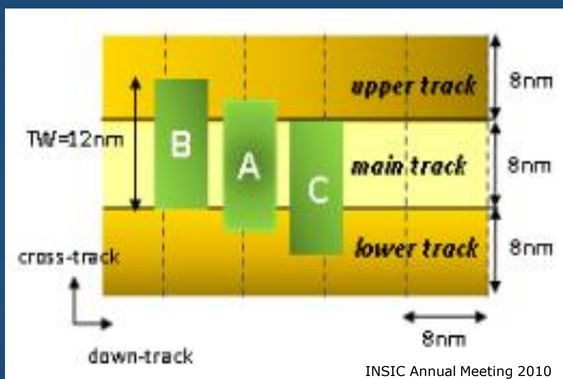
Track-Following and Signal-Processing

Scale
~ 1 cm *(still)*

TDMR

(2D Magnetic Recording)
Employs multiple readers across one or more tracks

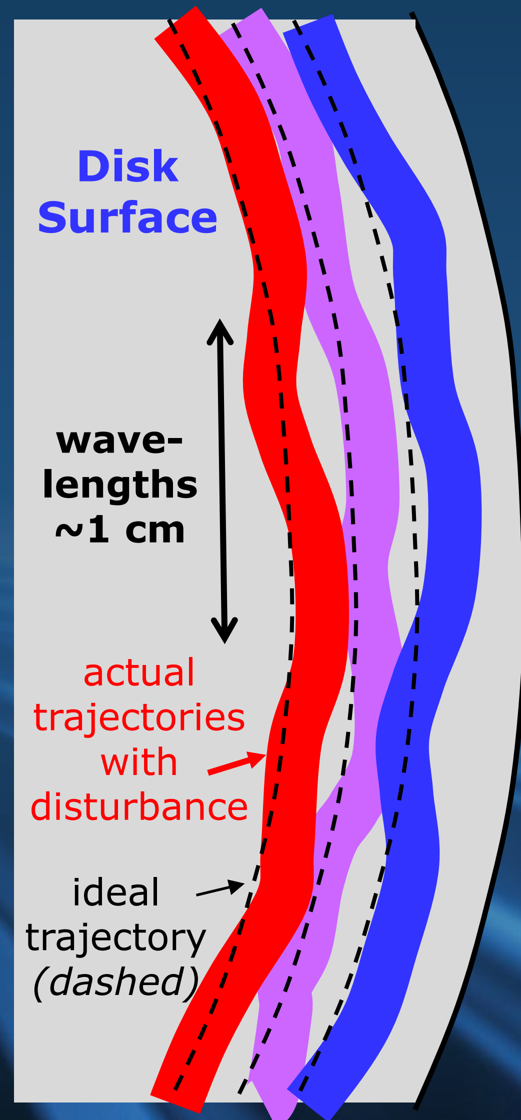
IEEE Trans. Magn., vol. MAG-45, pp. 917-923



Electronic Tracking

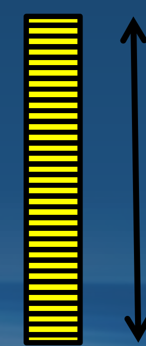
2 or 3 waveforms combine in one equalizer to optimize detection of center track

Accurate adaption of equalizer and detector can take several 4 kB sectors ~1 cm



Very long data-blocks

created with techniques such as distributed sector or soft track-ECC



64 kB block-length ~ 1 cm

Performance no longer dominated by weakest link

Decoding done iteratively: likelihood-information gets passed back & forth across entire span of the code.

US Pat. 9059737 & IEEE Trans. Magn. 9401704

Servo Sectors & Data Sectors

Western Digital

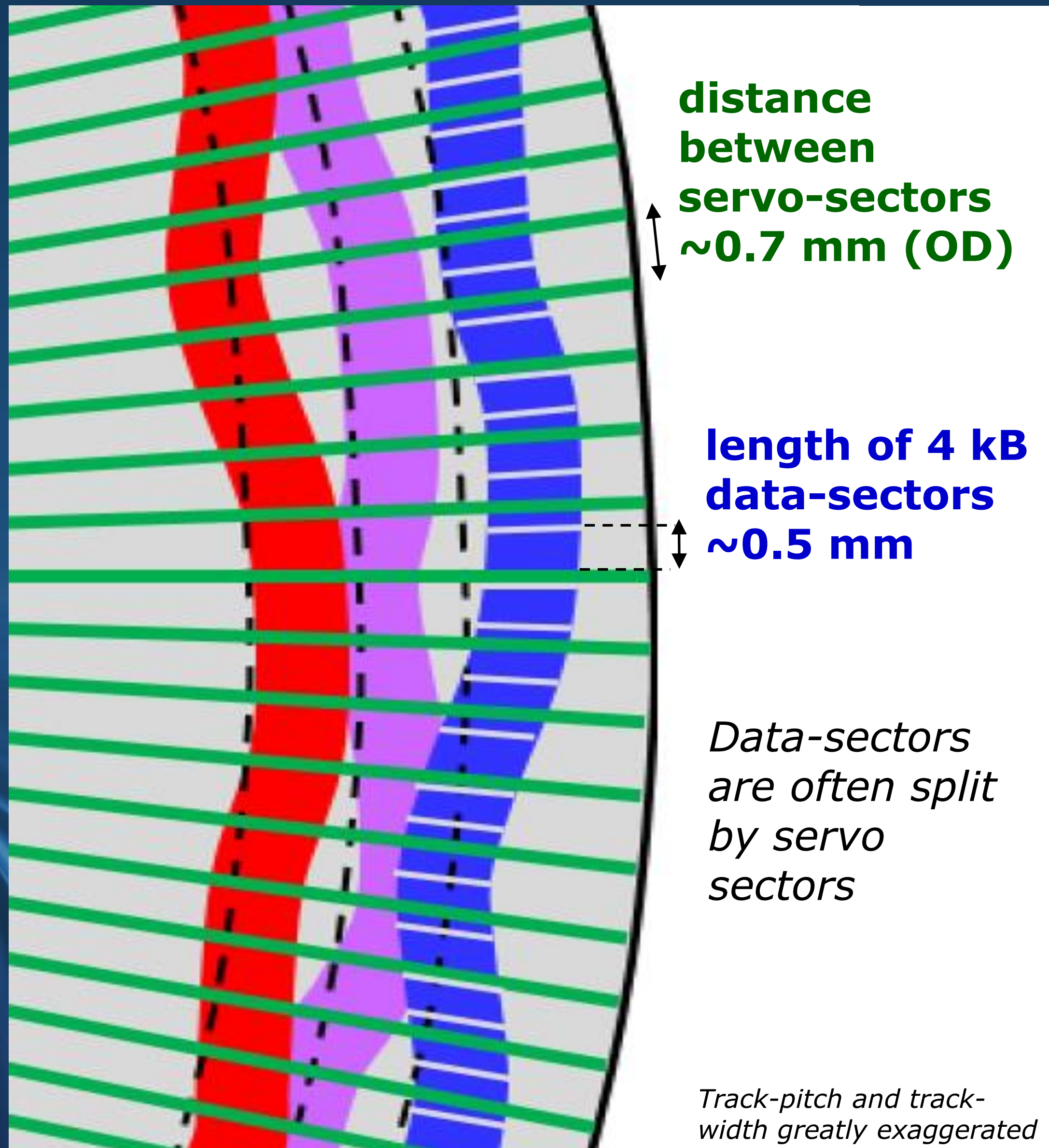
Scale ~ 1 mm (still)

SERVO

Example:

360 sectors
7200 rpm
95 mm disk

43 kHz sampling-rate (23 μs)



DATA

Example:

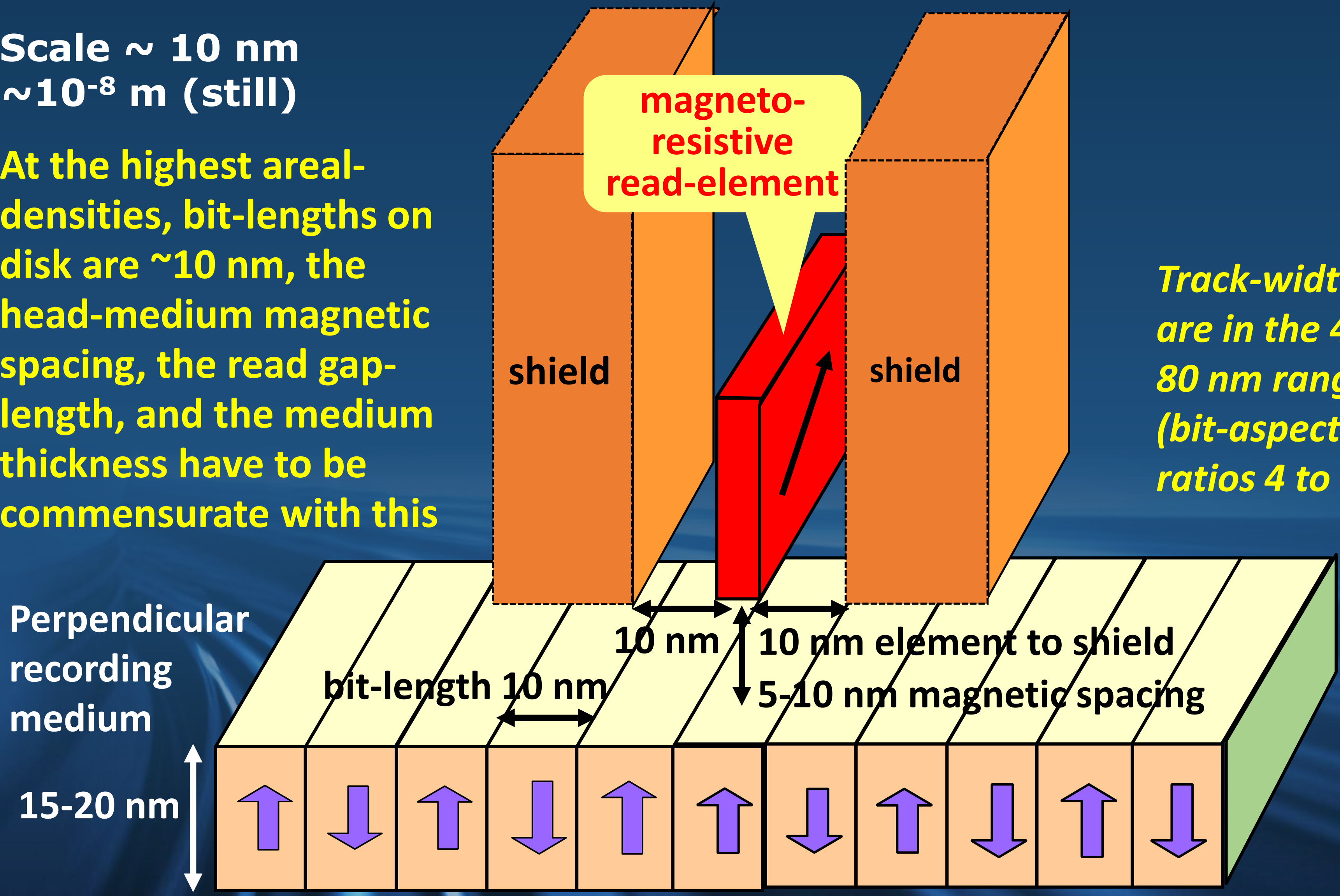
4 kB block
7200 rpm
95 mm disk

34 m/s x 2100 kbpi = 3 Gbit/s

Bit-length, magnetic-spacing, read-gap

Scale ~ 10 nm
~10^{-8} m (still)

At the highest areal-densities, bit-lengths on disk are ~10 nm, the head-medium magnetic spacing, the read gap-length, and the medium thickness have to be commensurate with this



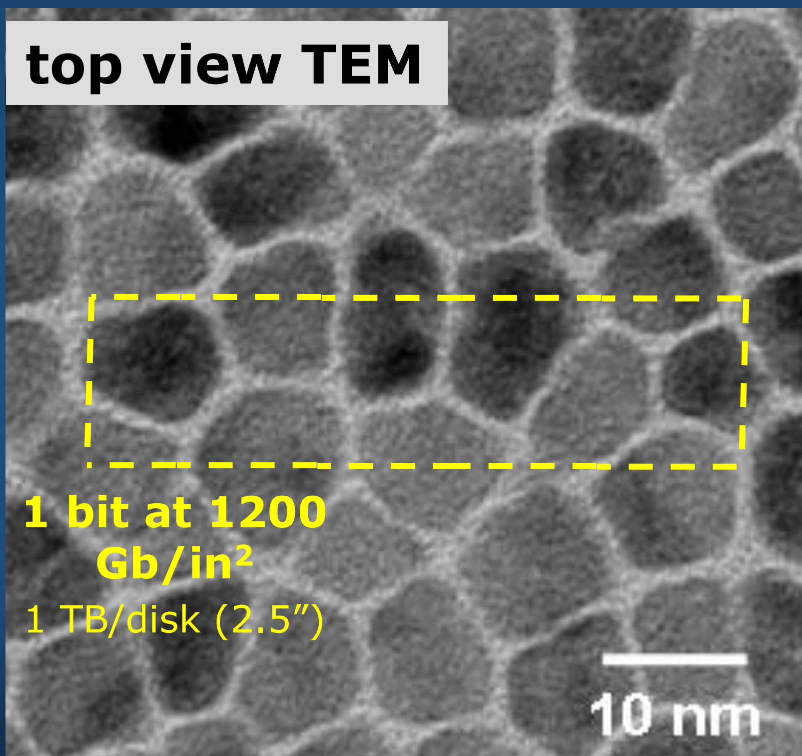
Track-widths are in the 40 – 80 nm range (bit-aspect ratios 4 to 6)

Magnetic Recording Medium

Scale ~ 10 nm (still)

Recording-layer properties:
high coercivity (switching field), small grains (~8nm), thin (10-20nm), high squareness, thermally stable

Example Disk Structure



CoCrPt alloy with Oxide segregant

- Lubricant
- Protective-overcoat
- Granular columnar recording layer (perpendicular) CoPtCr/Oxide
- Exchange-break or interlayer
  - high-pressure Ruthenium
  - low-pressure Ruthenium
- NiW Seed layer
- Soft Magnetic Underlayer (thick)