Microsoft

Optics for Cloud: new approaches to data centre technology

EPSRC CDT in Applied Photonics Annual Conference, Heriot-Watt University, 7 June 2019

Tom Empson
Azure Storage Research Laboratory

Dublin

Microsoft

Data is doubling every two years

2013

```
1
0
0
1
1
0
1
0
1
0
```

2020

```
1 0 1 1 0 1 0 1 0 0
0 1 1 0 1 0 0 0 0 1
0 0 0 1 0 0 1 1 0 1
1 0 1 1 0 1 0 1 1 0
1 0 1 0 0 1 0 1 0 0
0 1 0 0 1 0 1 0 0 0
1 0 1 0 0 1 1 1 1 1
0 1 0 1 1 0 0 0 1 0
1 0 1 1 0 1 1 1 1 1
0 1 0 0 1 0 1 0 0 1
```

Slide credit: Mark Russinovich

Source: IDC 2014

Microsoft

1 ZB is one billion TB: we’re facing a data tsunami

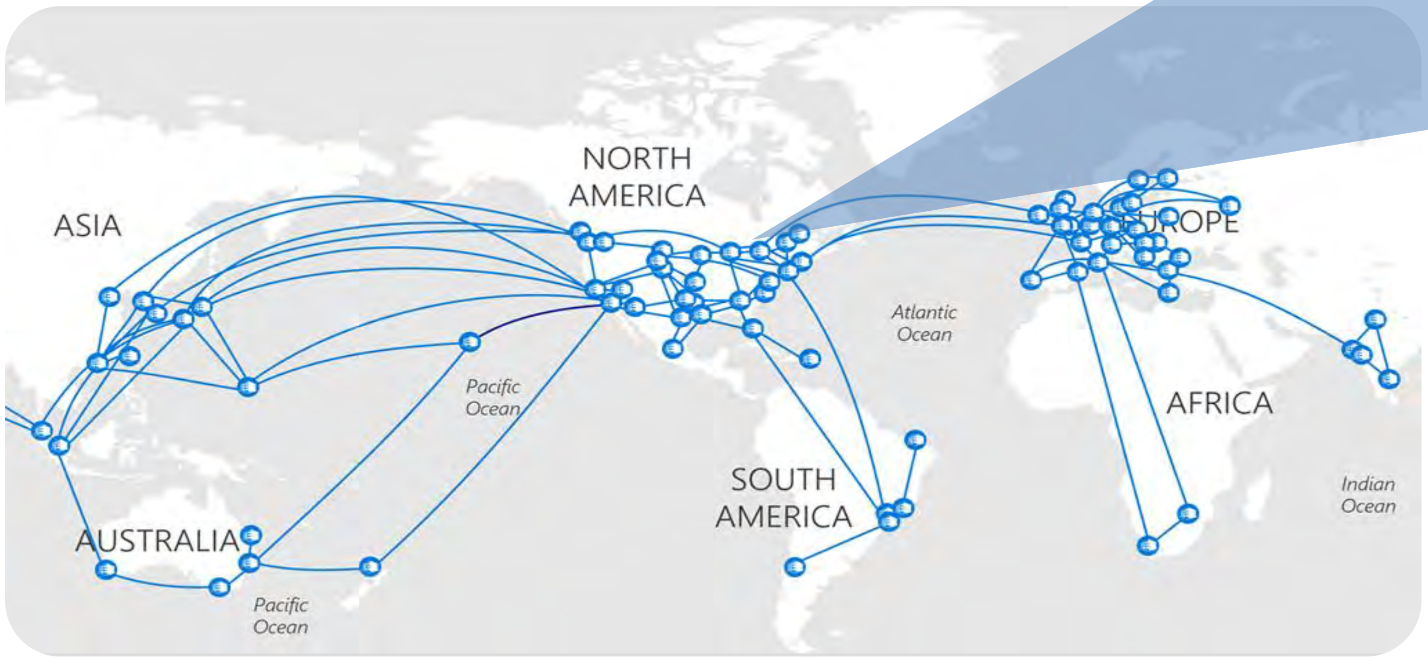
- Current HDDs can store 10 TB…
- So we need 100 million of those to store 1 ZB…
- That means 1000 data centres…
- Which would cover 20% of Manhattan

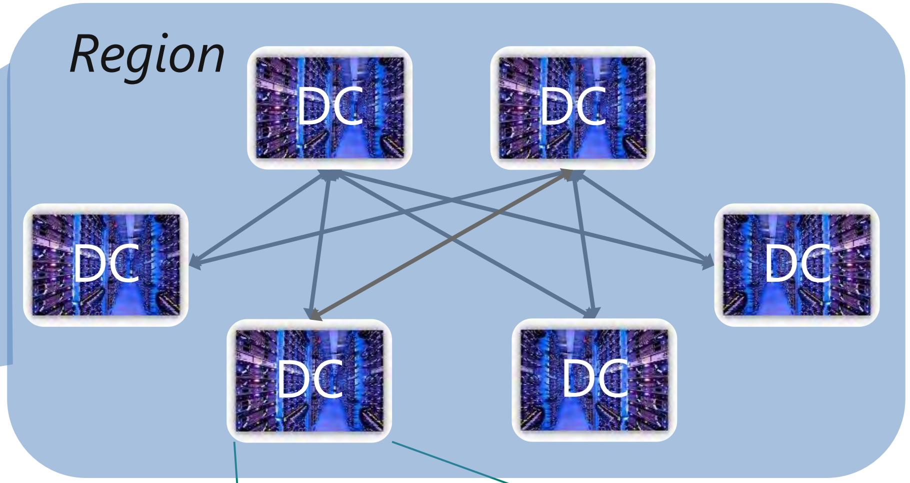
- This data also needs to be transmitted across the network…
- Data centre network links can send 100 Gbps…
- That’s 2500 years to transmit a zettabyte

Microsoft

Up to **20 data centres** per region
Up to **60 km** of cable between regions

Azure’s data centre infrastructure





Data centres in over **50 regions**
~**10,000 km** of fibre-optic cable connecting these

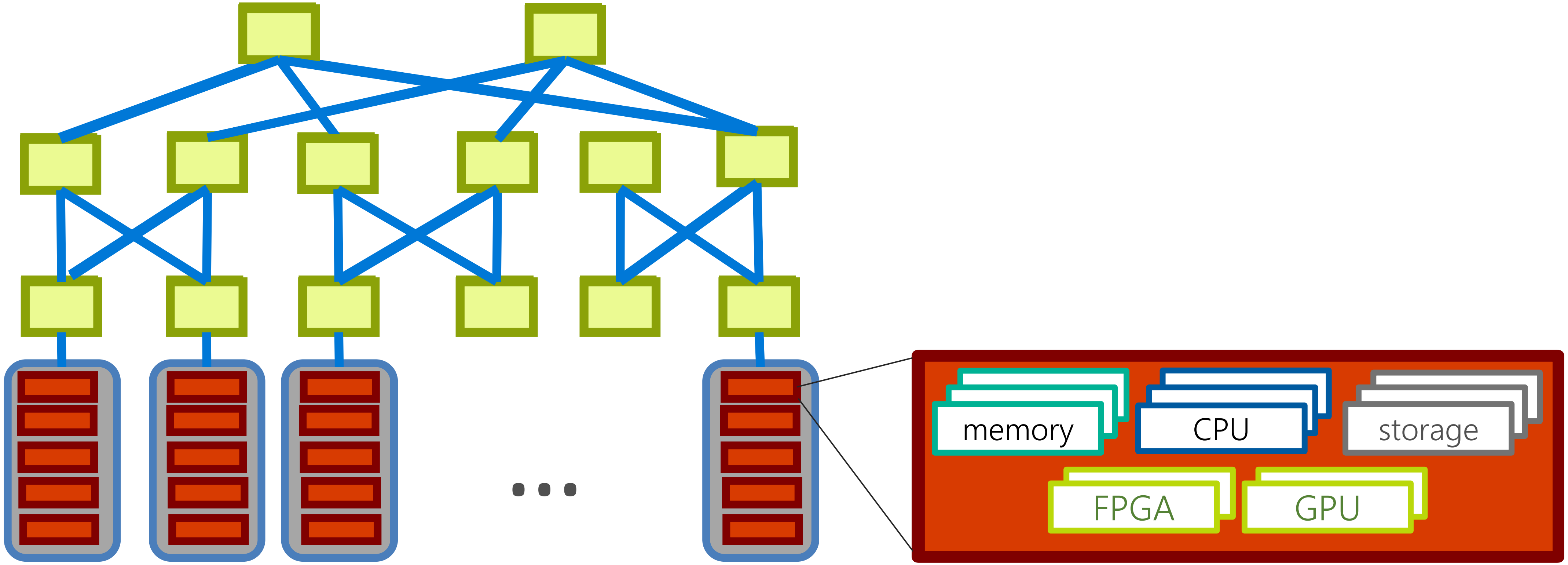
~ **100,000 servers** per data centre

Microsoft

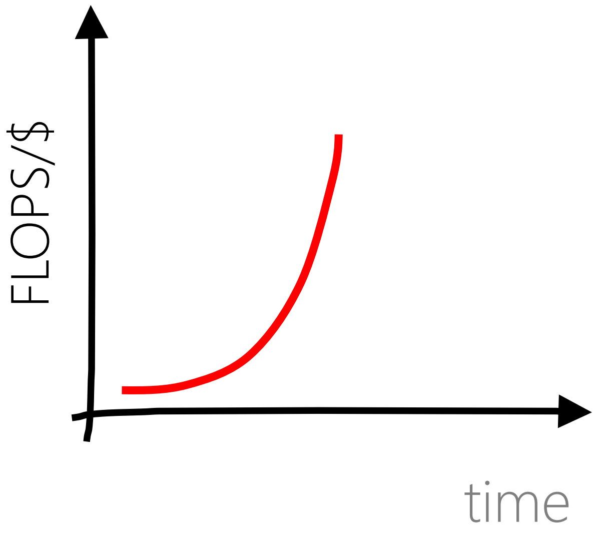
Microsoft

Microsoft

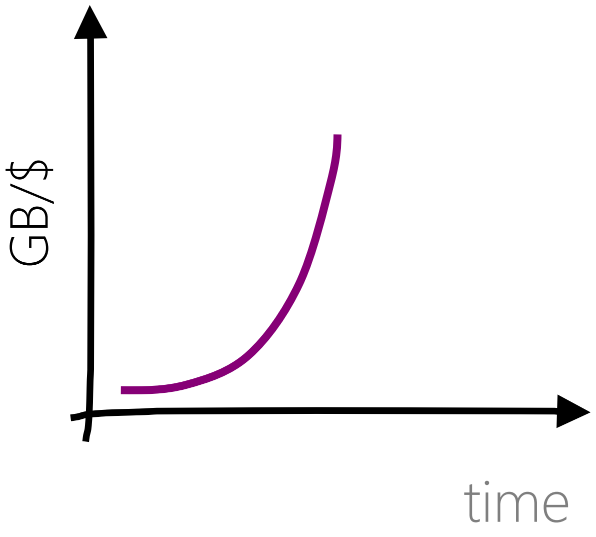
How can we scale the capacity of the cloud?



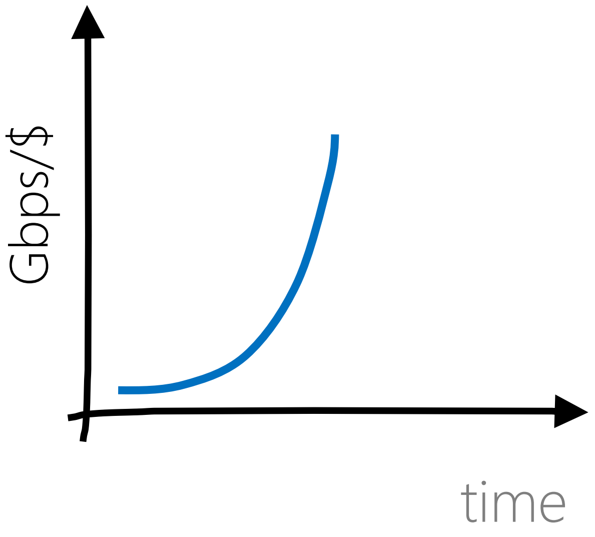
We have addressed these challenges in three areas



Compute



Storage

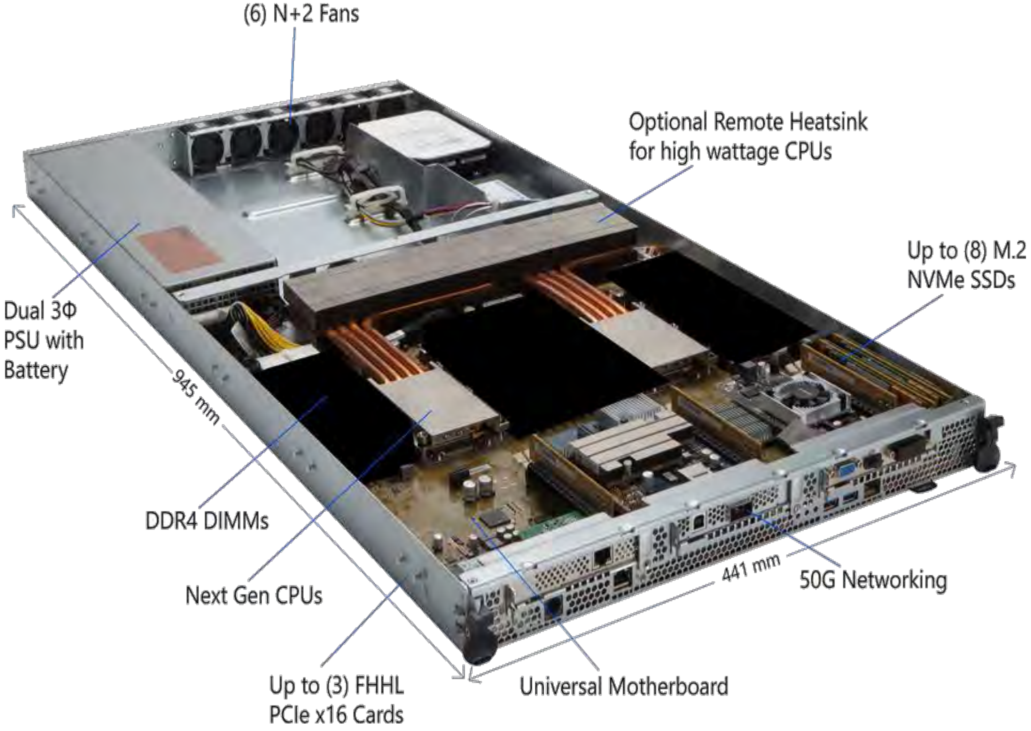


Network

Microsoft

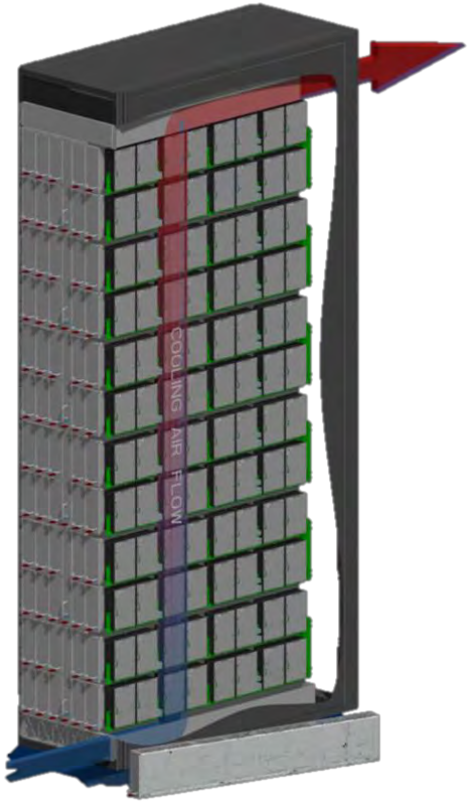
Pushing compute growth

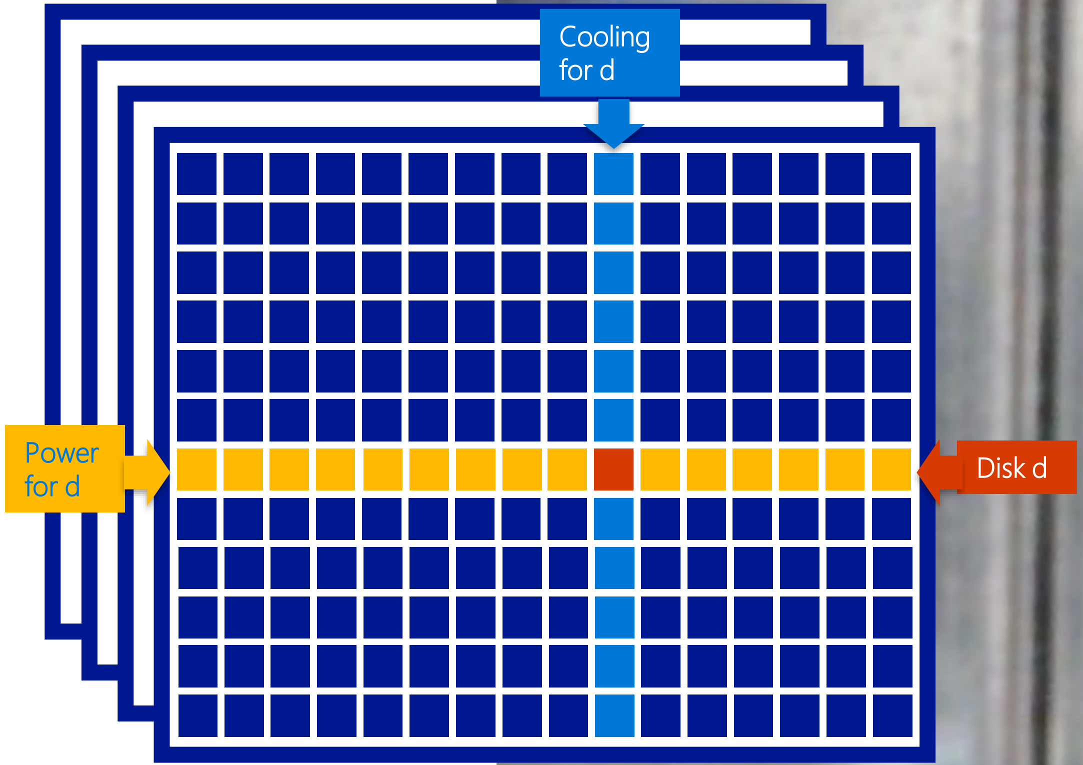
Project *Olympus*: hyperscale cloud hardware design in collaboration with OCP (Open Compute Project)



Pushing storage growth

Project *Pelican*: custom storage racks optimised for cold storage

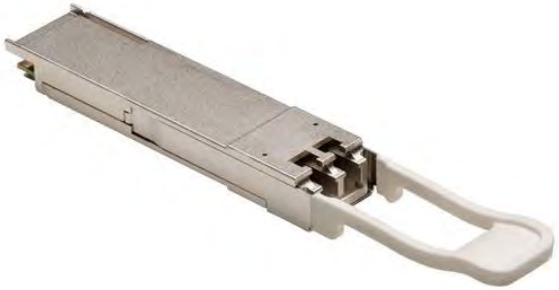


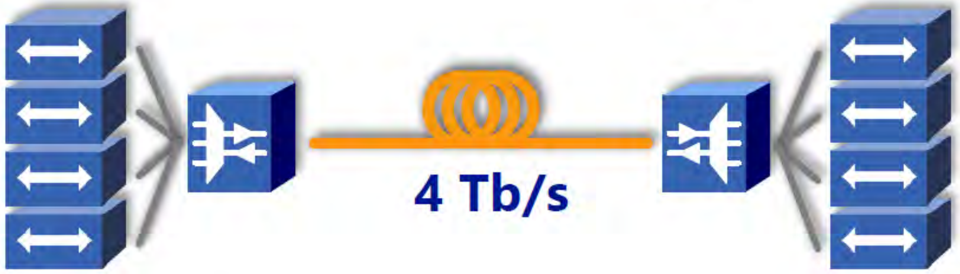


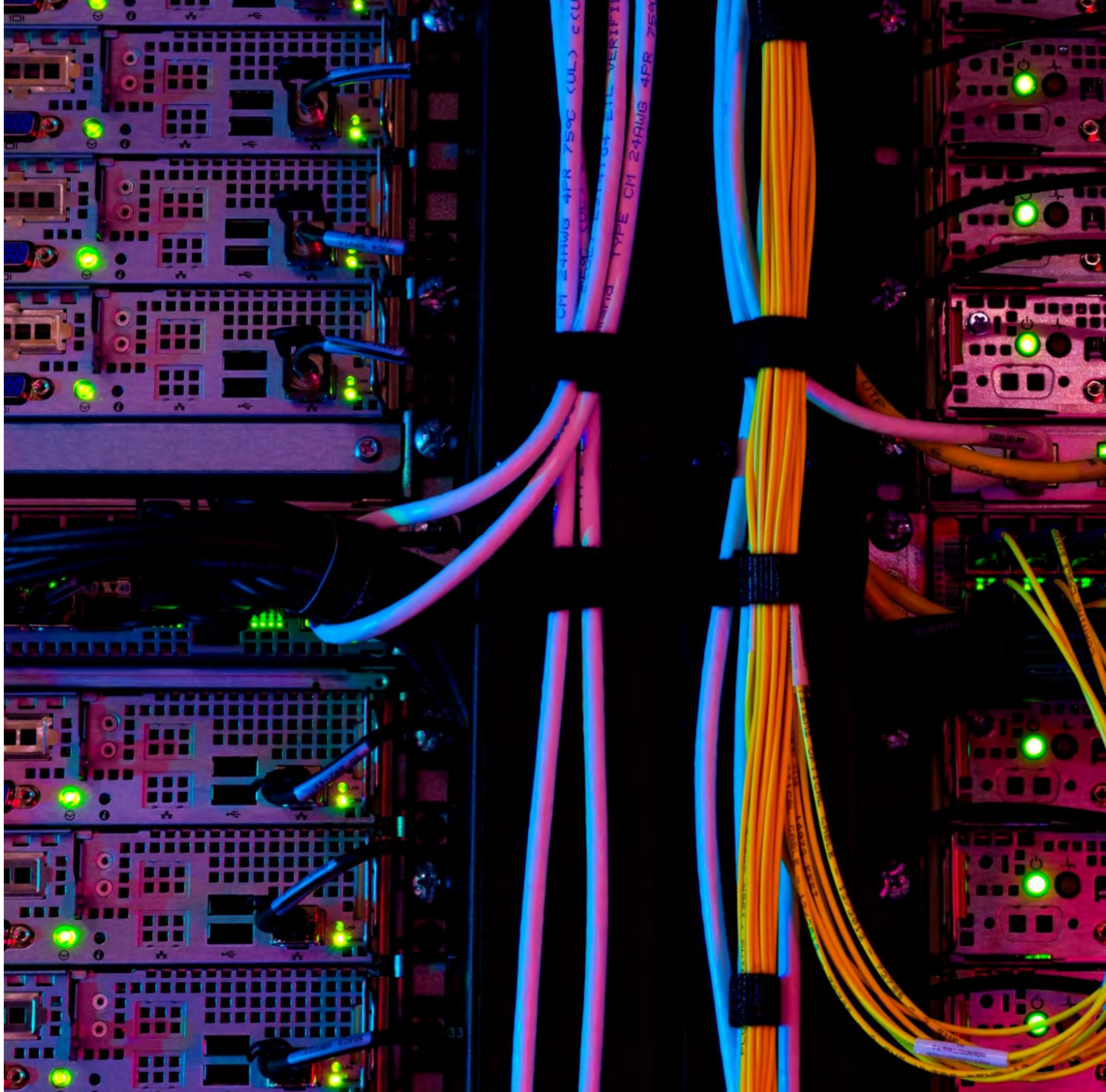
Microsoft

Pushing network growth

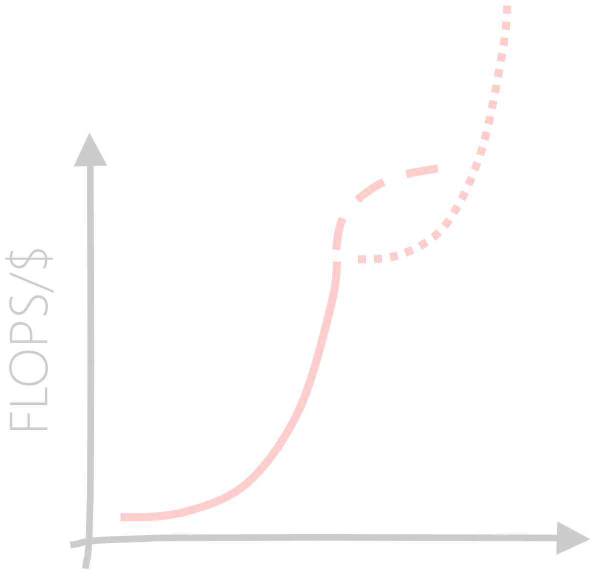
Project *Madison*: custom data centre interconnect technology at multi-terabit speeds



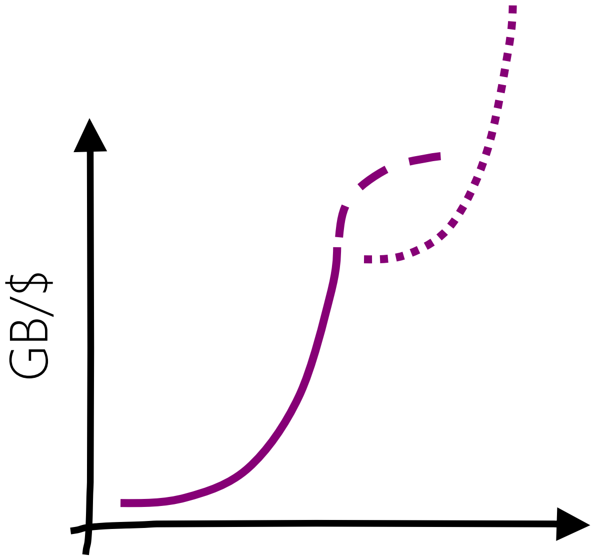




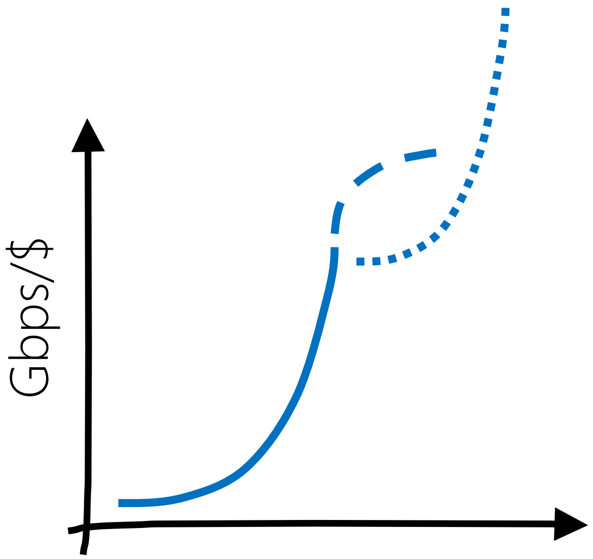
We kept up (just)... but the end is nigh... we must respond!



Compute



Storage



Network

Microsoft

Impressive advances in optical technologies in the 21st century

Breakthrough innovation in academic labs across the world:

***Optics for the Cloud* Research Alliance**

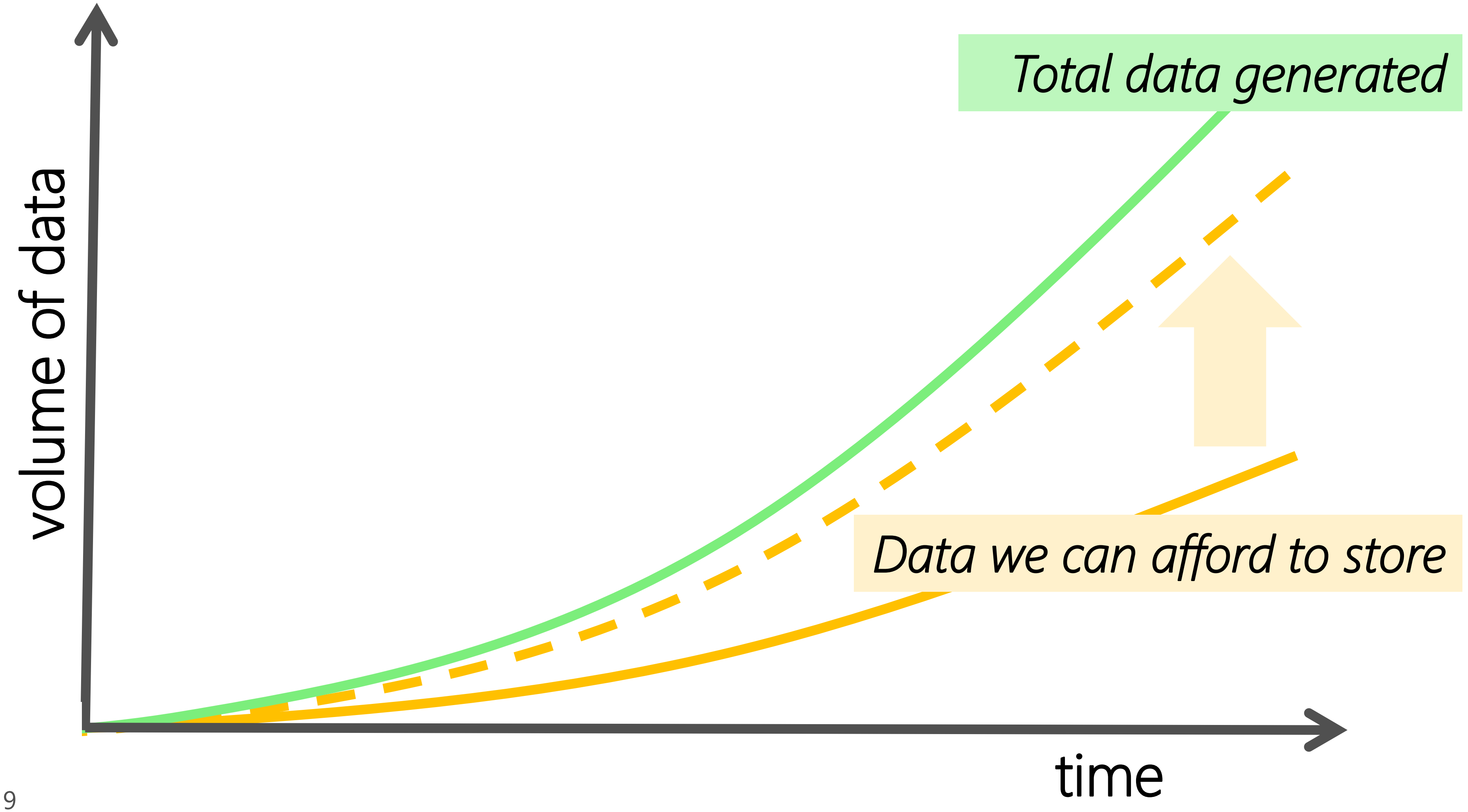
Cambridge, EPFL, Imperial, Southampton, TU/e, UCL, UCSB, etc etc

Microsoft

Project *Silica*: optical storage for the Cloud

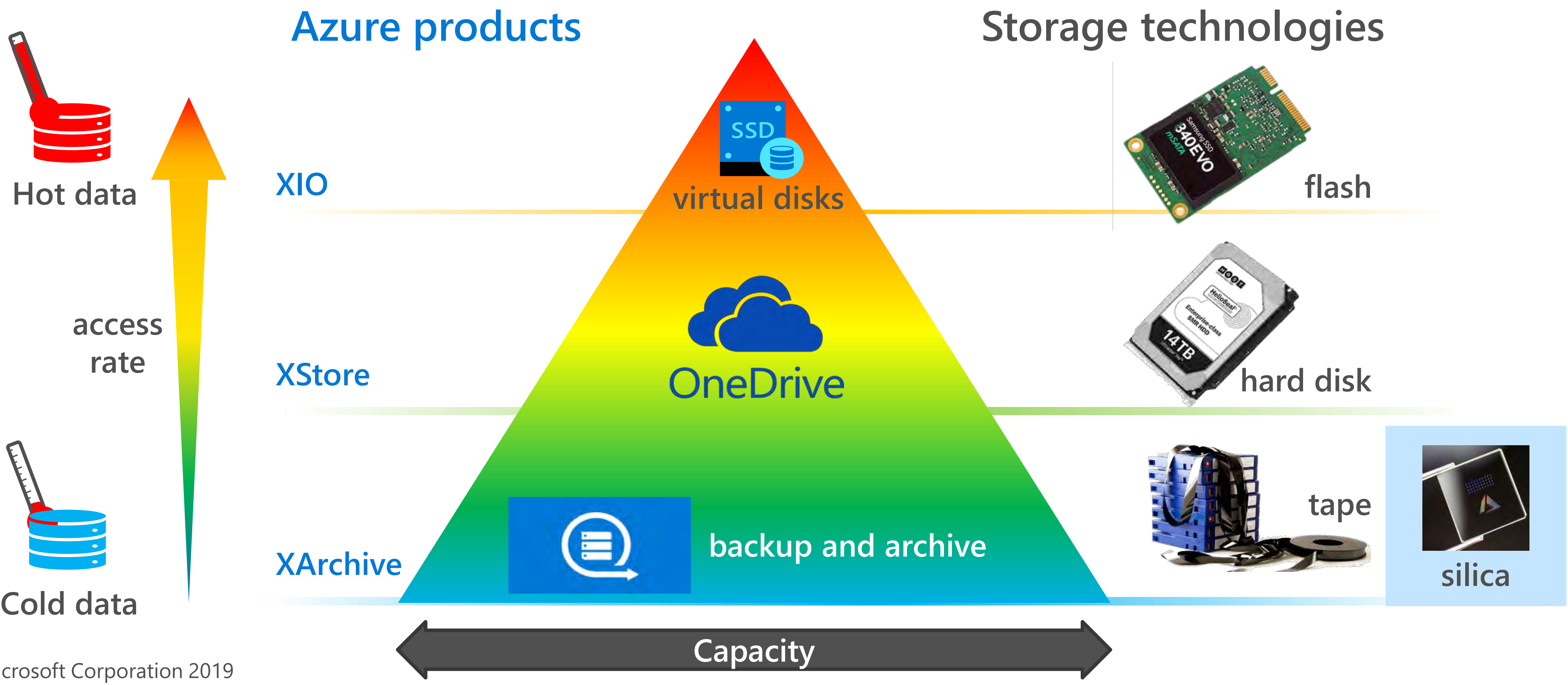
Microsoft

Data volume is growing … we need to reduce $/GB



Microsoft

Today's cloud storage landscape has three tiers



Azure products

Storage technologies

Hot data

access rate

Cold data

XIO

virtual disks

XStore

OneDrive

XArchive

backup and archive

flash

hard disk

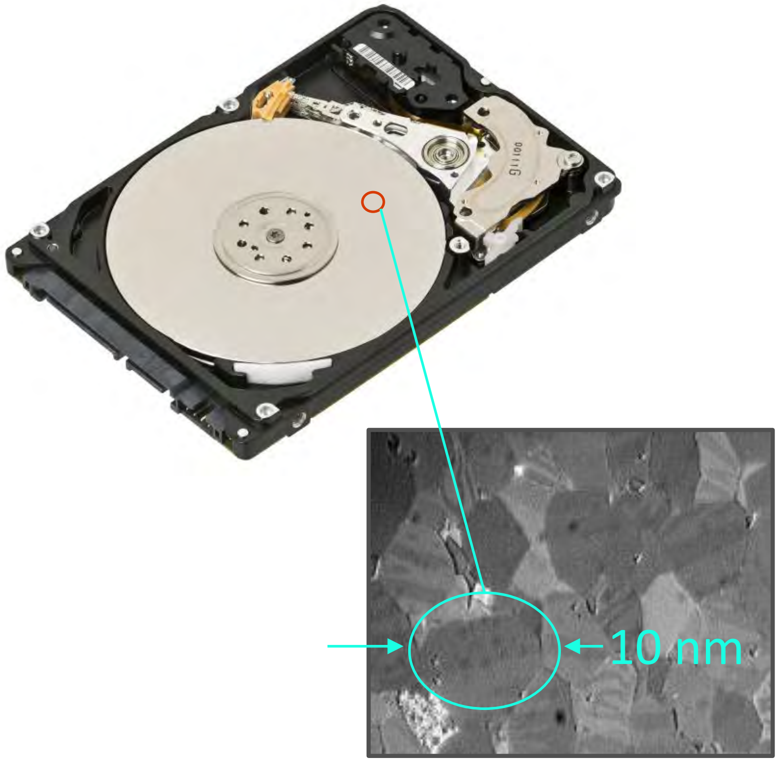
tape

silica

Capacity

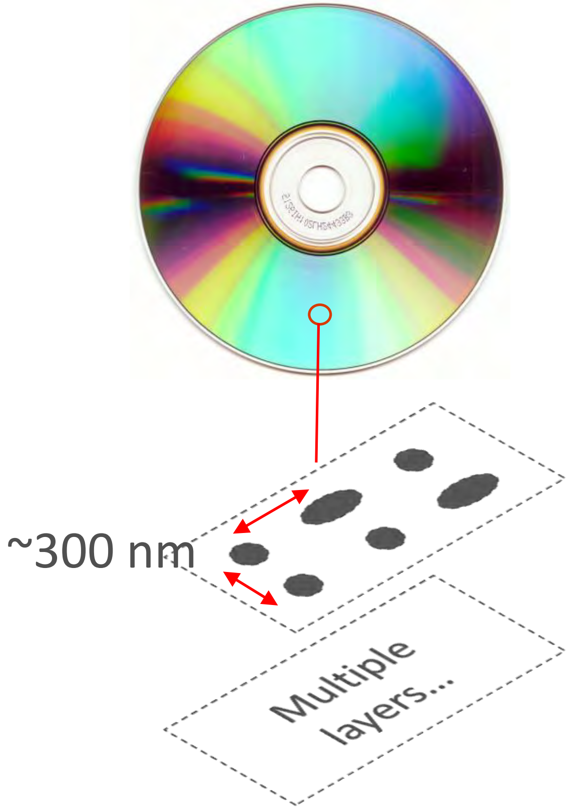
Existing archival storage media is magnetic or optical

hard disk



magnetic tape

optical disk



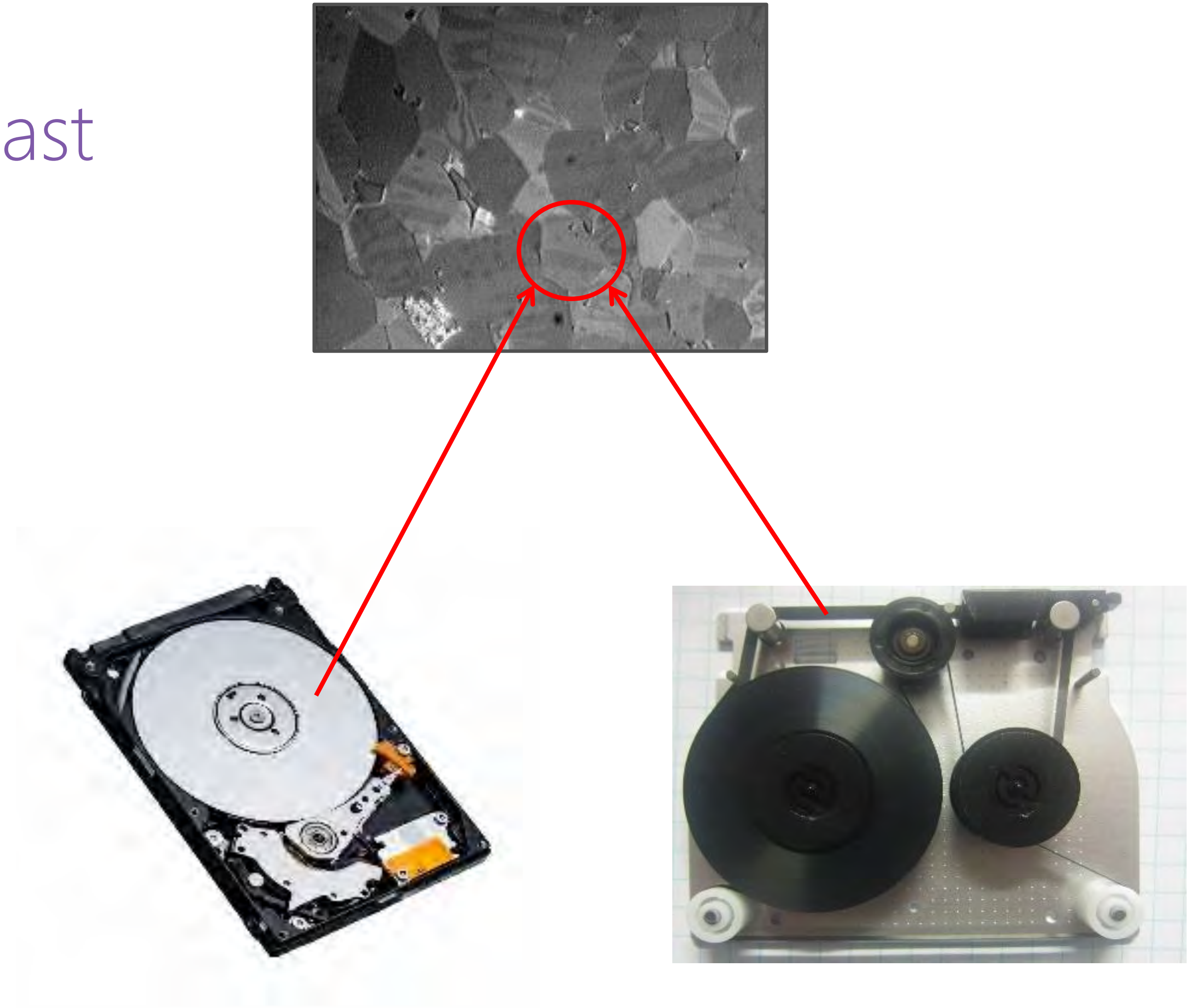
Magnetic media doesn’t last

Magnetic media degrades over time

- Latent sector errors
- Requires scrubbing

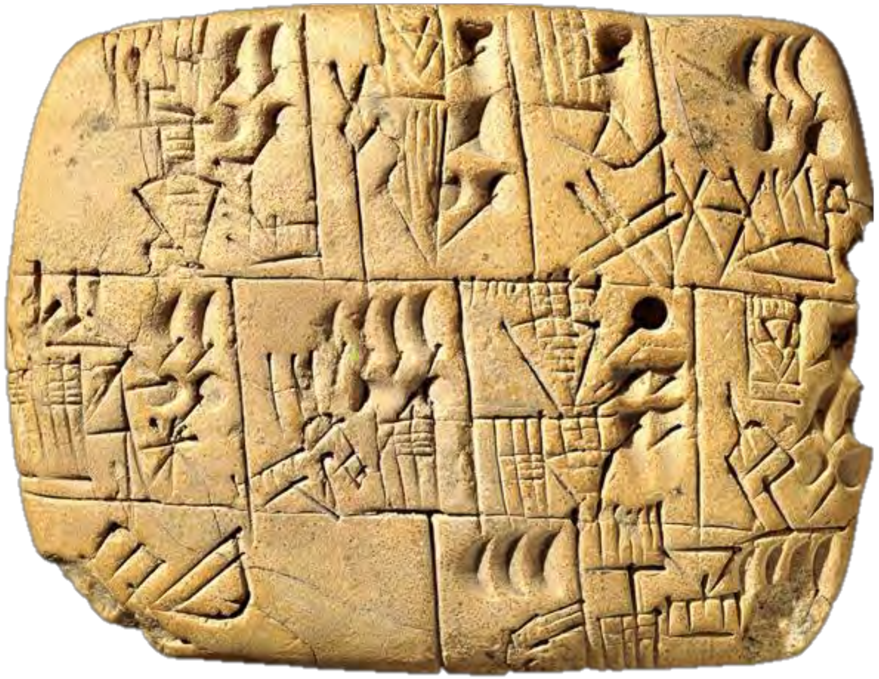
Magnetic media has limited life

- HDD: 3 - 5 years
- tape: 5 - 7 years



Data can be stored by changing the structure of silica

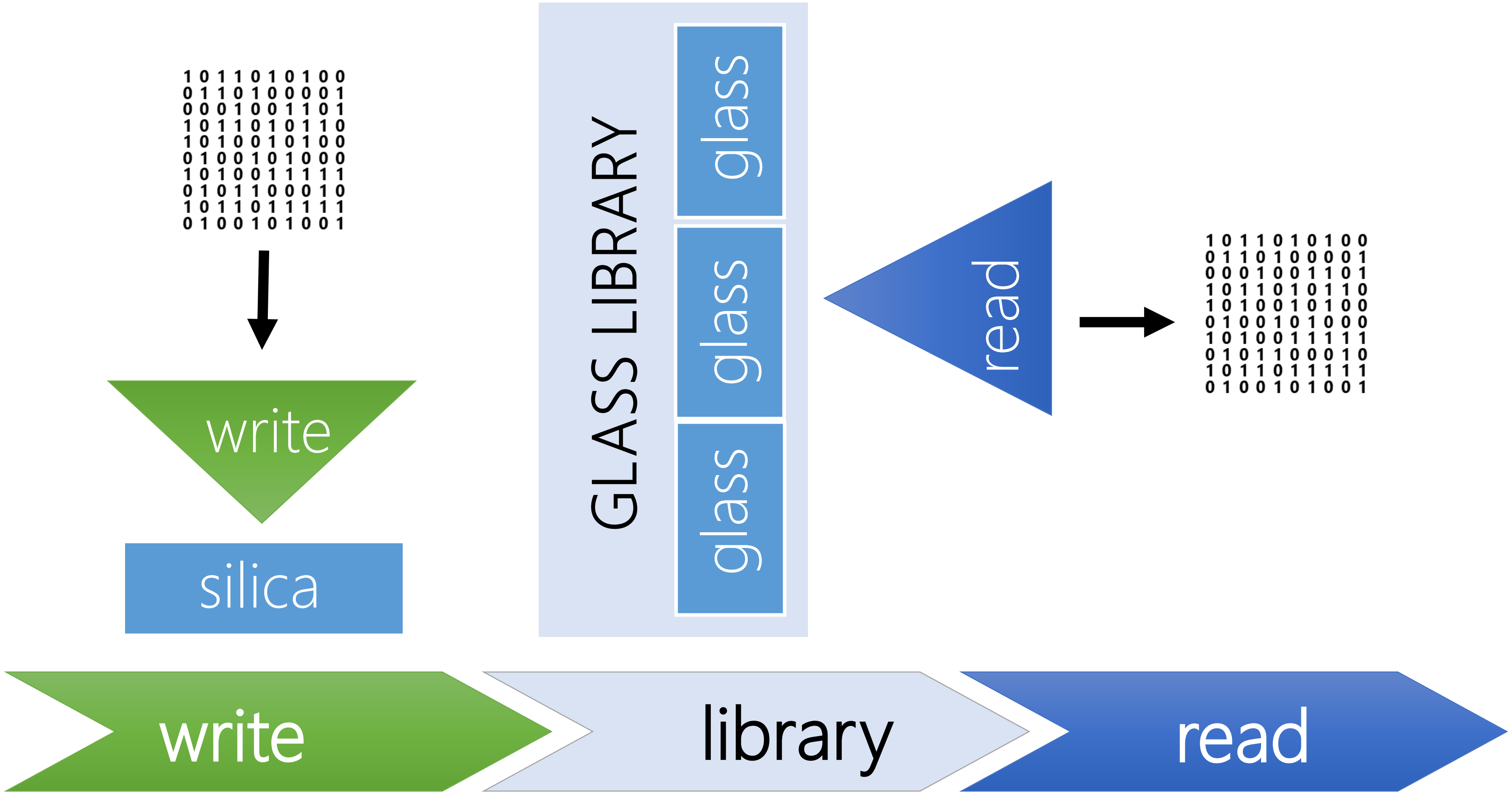
- Write once, read many
- Persistent – lifetime of millennia
- Immune to EMI/EMP
- No bit rot or disc rot or media decay…
- …hence no need to re-write periodically
- Cheap
- Better access time than tape
- Leave data in place



Cuneiform tablet recording the allocation of beer, 3100-3000 BC.
© Trustees of the British Museum.

Microsoft

In *Silica*, write, store & read are disaggregated

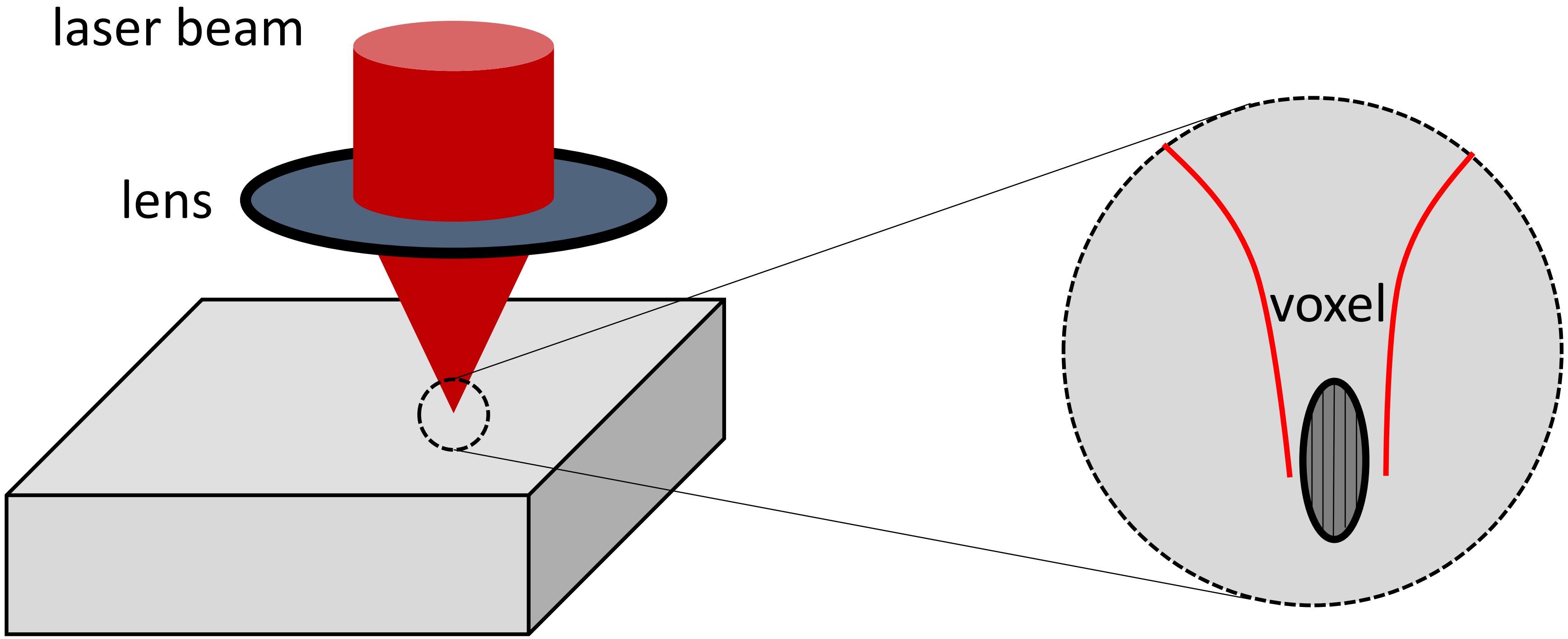


```
1 0 1 1 0 1 0 1 0 0
0 1 1 0 1 0 0 0 0 1
0 0 0 1 0 0 1 1 0 1
1 0 1 1 0 1 0 1 1 0
1 0 1 0 0 1 0 1 0 0
0 1 0 0 1 0 1 0 0 0
1 0 1 0 0 1 1 1 1 1
0 1 0 1 1 0 0 0 1 0
1 0 1 1 0 1 1 1 1 1
0 1 0 0 1 0 1 0 0 1
```

```
1 0 1 1 0 1 0 1 0 0
0 1 1 0 1 0 0 0 0 1
0 0 0 1 0 0 1 1 0 1
1 0 1 1 0 1 0 1 1 0
1 0 1 0 0 1 0 1 0 0
0 1 0 0 1 0 1 0 0 0
1 0 1 0 0 1 1 1 1 1
0 1 0 1 1 0 0 0 1 0
1 0 1 1 0 1 1 1 1 1
0 1 0 0 1 0 1 0 0 1
```

Microsoft

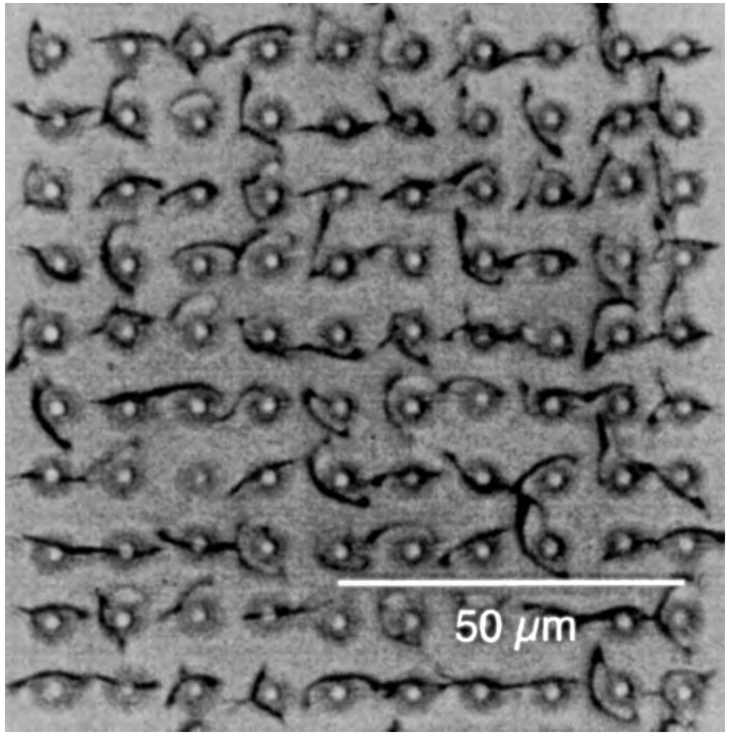
Data is written into silica by femtosecond laser pulses



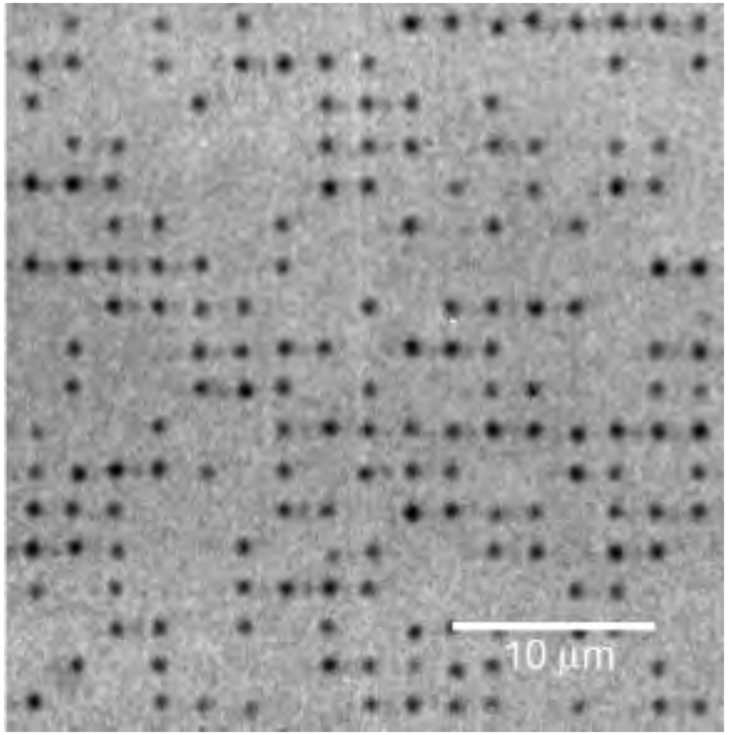
Microsoft

Pulse duration is critical

picosecond laser induces
voids with external stress

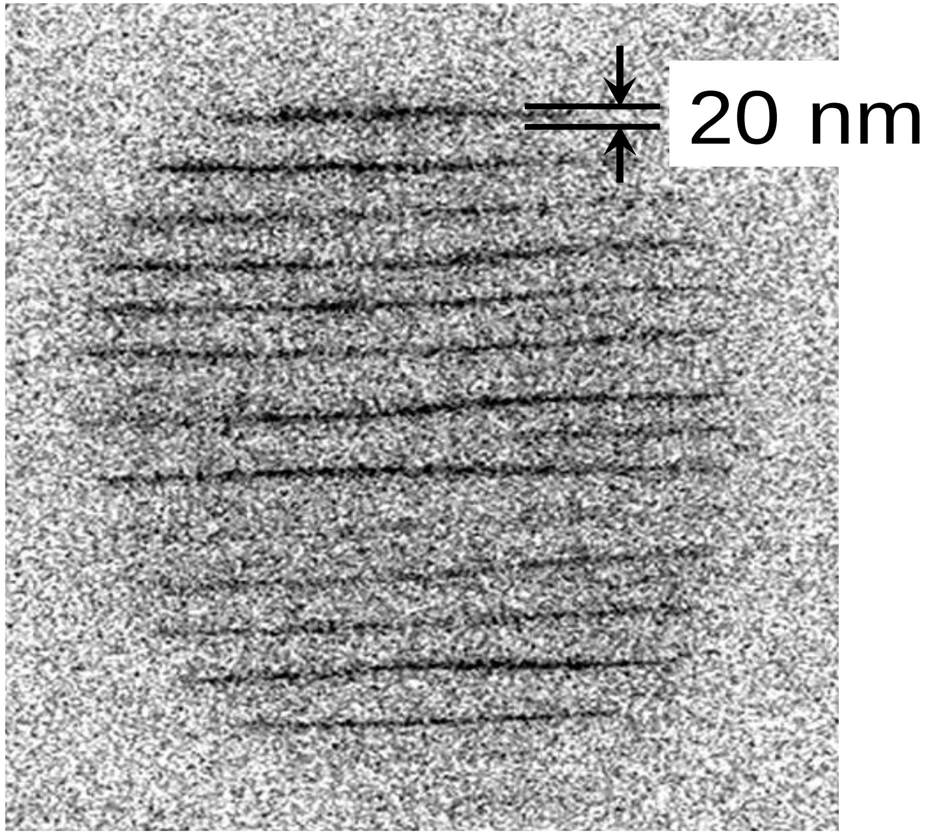


femtosecond laser induces
sub-wavelength nanogratings

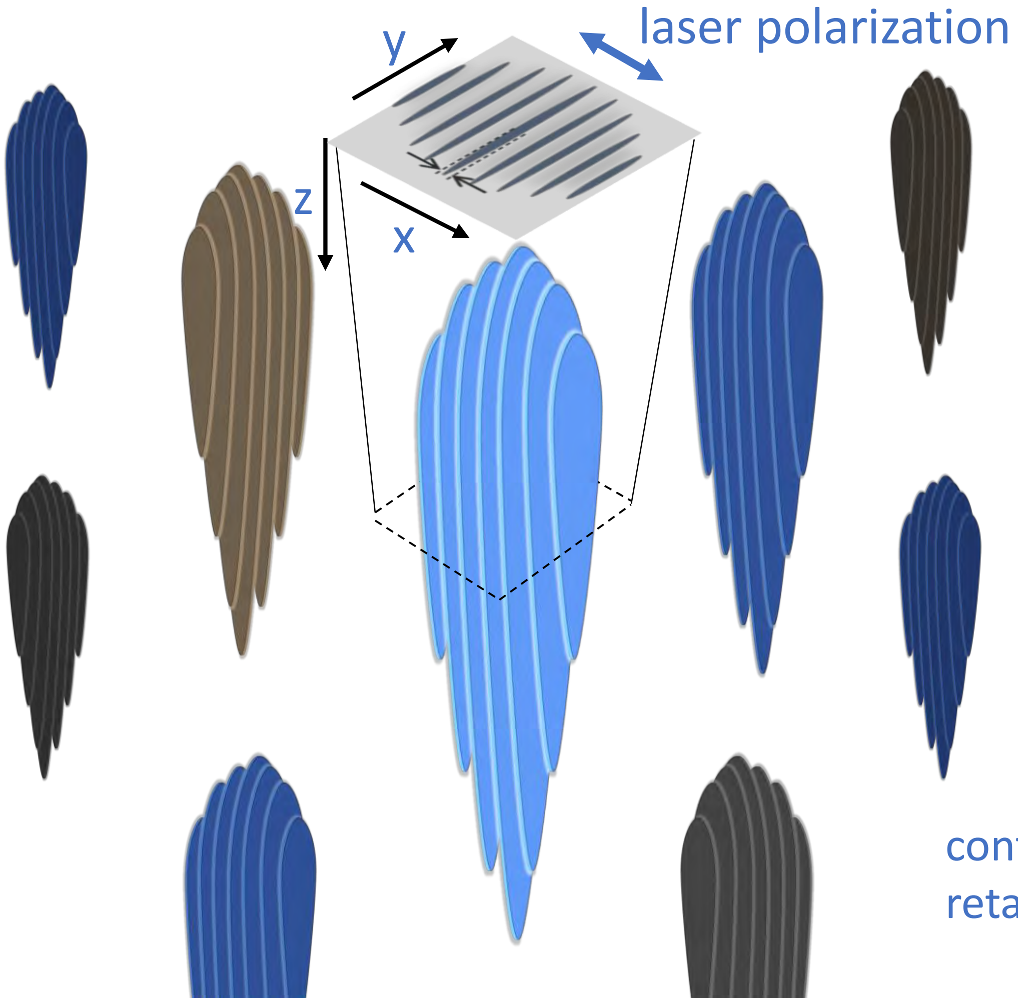


Three-dimensional optical storage inside transparent materials. E. N. Glezer, M. Milosavljevic, L. Huang, R. J. Finlay, T.-H. Her, J. P. Callan, and E. Mazur .
Optics Letters Vol. 21 Issue 24 (1996)

The voxels are sub-wavelength nanogratings



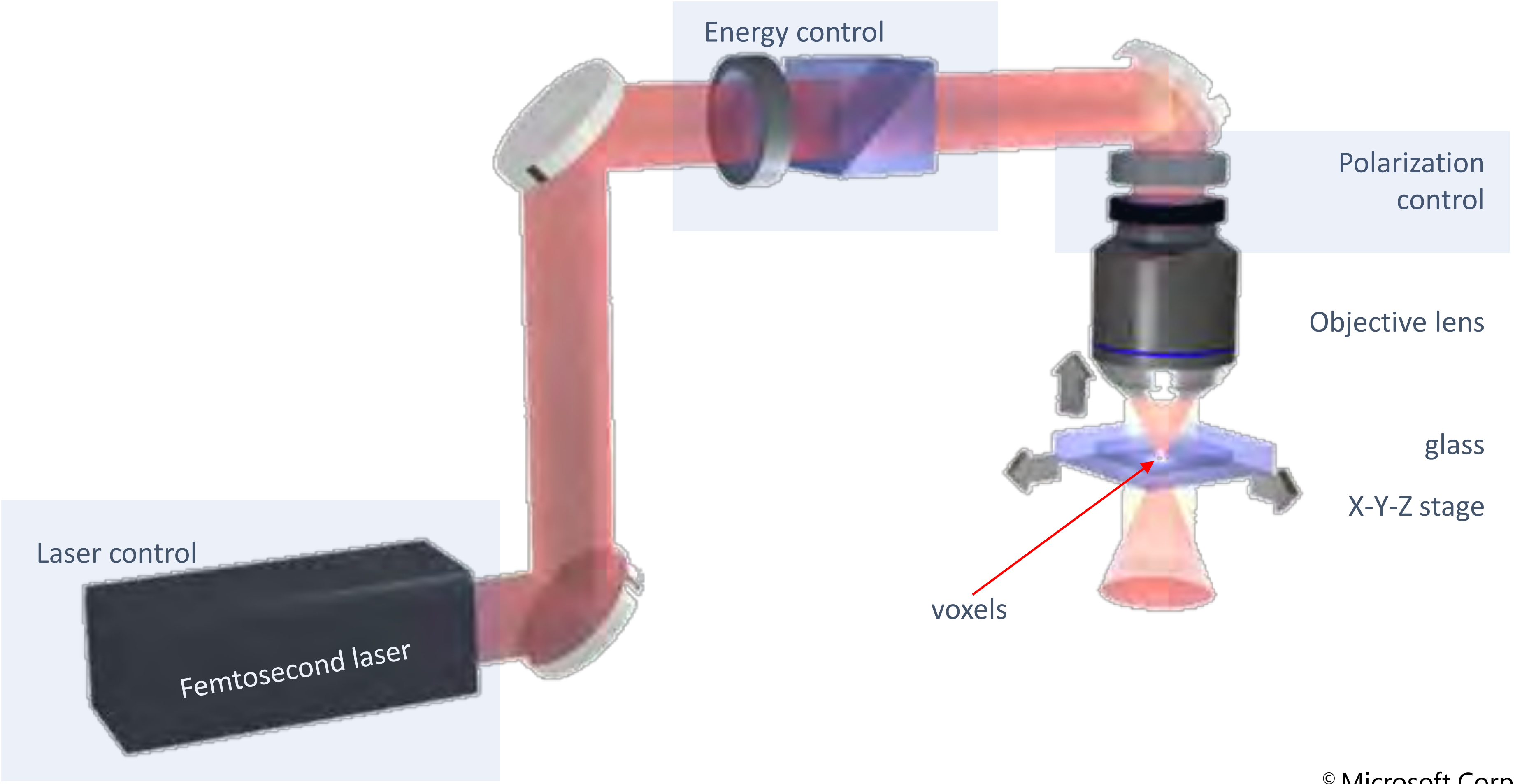
Self-Organized Nanogratings in Glass Irradiated by Ultrashort Light Pulses, Yasuhiko Shimotsuma, Peter G. Kazansky, Jiarong Qiu, and Kazuoki Hirao Phys. Rev. Lett. **91** (2003)



controllable retardance

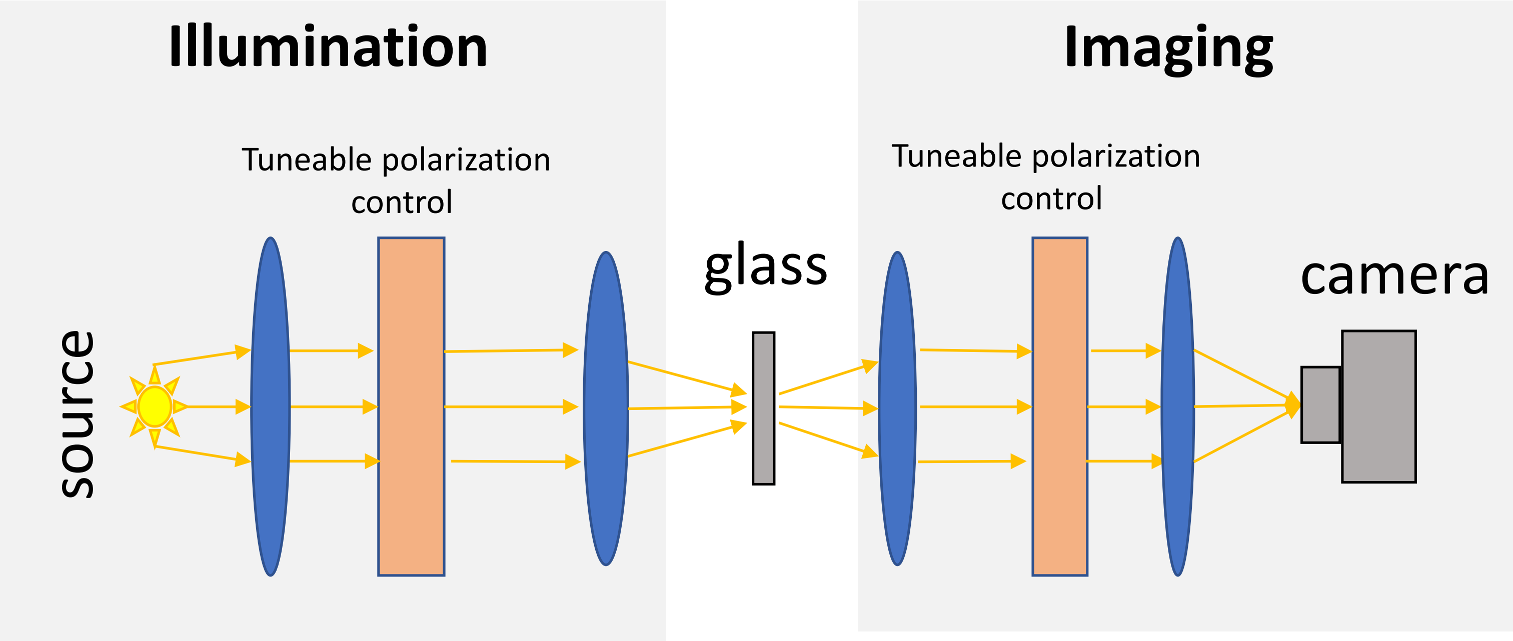
Microsoft

Schematic of WRITE system

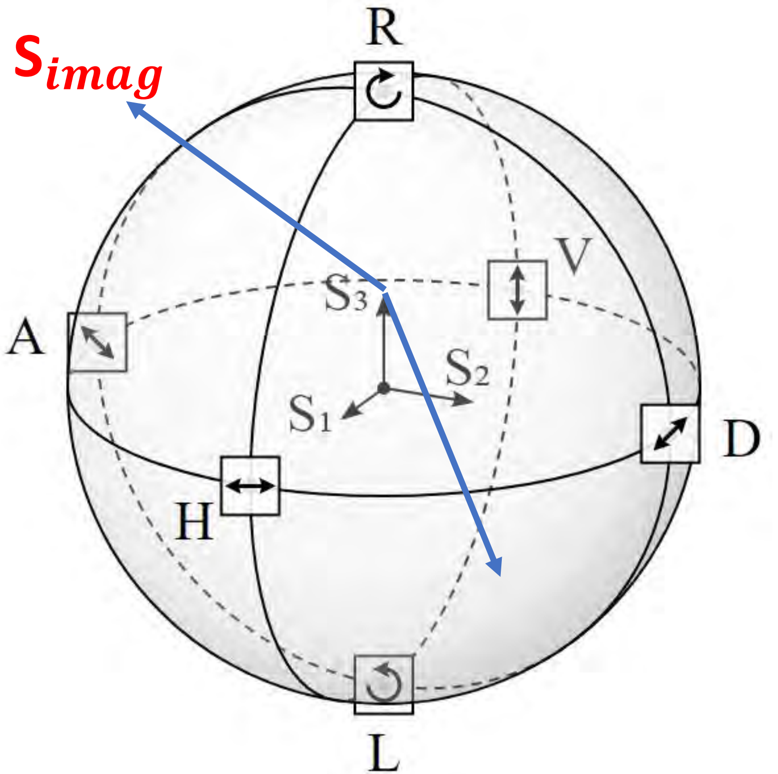


The READ head is a polarisation microscope

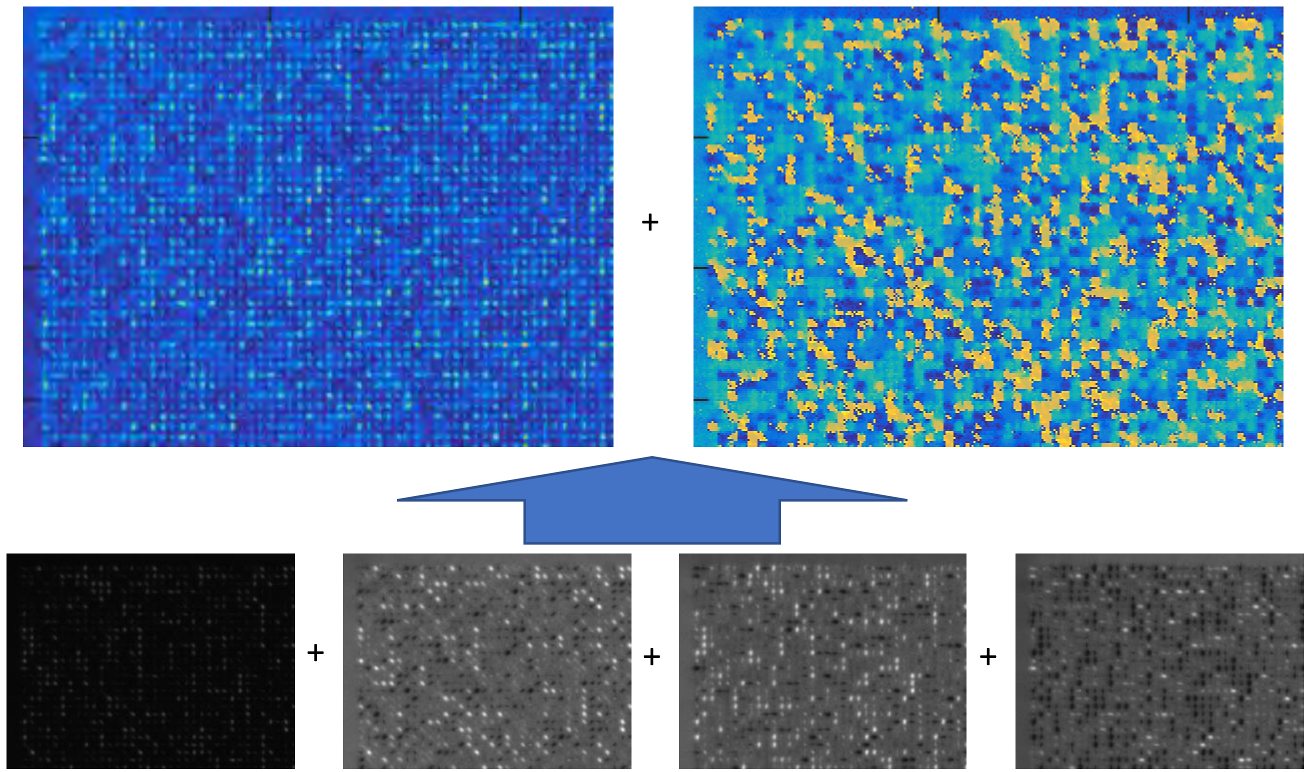
- Measure nanogratings under several polarisations of light
- Infer retardance and angle



Poincare sphere

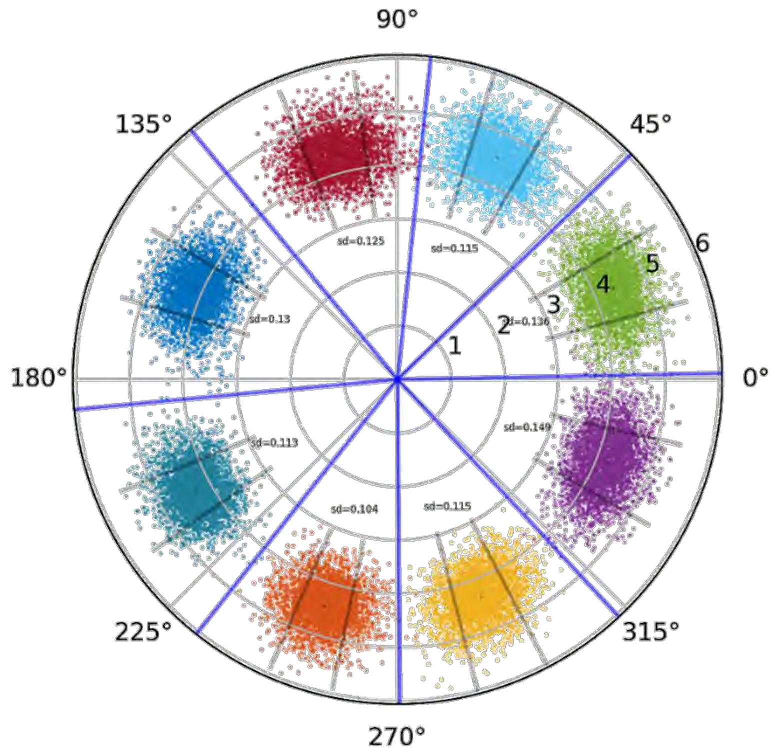


We extract the retardance and orientation values



Retardance

Orientation



Constellation diagram

Thank you!

- Tom Empson
- v-toemps@microsoft.com
- aka.ms/silica