

FIT3142 Distributed Computing

Topic 12: Limits to Application Performance

Dr Carlo Kopp
Faculty of Information Technology
Monash University
© 2009-2016 Monash University

Why? Performance Growth Matters!

- If we are planning a distributed application, we must think strategically and anticipate how the computing environment will evolve.
- Moore's Law, Bandwidth Law, Kryder's Law and other measures of "performance over time" or "performance per dollar" are important factors in planning.
- An application which is not quite feasible or economical may become feasible or economical in the near future.
- Industry is always competing to be first in the market so planning is critical to success.



Factors Limiting Application Performance

- Data dependencies in algorithms.
- Basic scalability of the application software.
- Processor performance.
- Disk storage performance.
- Memory capacity and performance.
- Network performance – link performance.
- Network performance – packet switching performance.
- *Are there any hard limits on how far distributed application performance can evolve in the future?*



Strategic Thinking When Planning

- Planning any distributed application requires careful thinking about the design of the application and its performance.
- Earlier lectures looked at the problem of what performance can be achieved with an application for a *given* algorithm, *given* CPU performance, and *given* network performance.
- The results are valid for these conditions only.
- Good planning requires that a designer think in terms of “what happens in 1, 2, 3 N years time, given what we know about performance growth?”



Scalability Vs Performance of Systems

- Amdahl's Law shows that data dependencies severely impact scalability as a result of the serial computation components, and network communications delays resulting from data dependencies.
- Consider an application which scales poorly due to the network delay.
- What happens if the performance of the network is quadrupled due to the bandwidth law?
- The scalability problem may become non-critical and the application feasible.



Impediments to Performance Growth

- While basic technology may show ongoing growth over time due to Moore's Law, Bandwidth Law, etc, it may not be feasible to exploit this quickly.
- This is because hard limits may exist in various technology standards which must be used if components are to interoperate.
- Examples – bus and interface standards like 802.3, 802.11, Firewire, USB, IED, EIDE, SATA1/2/3, SCSI-1/2/3, Thunderbolt; networking standards like ADSL 1, ADSL 2, ADSL 2+ etc.
- Good design often requires study of such standards and deployment schedules.



Digital Computer Implementation Technologies

- *Medium Scale Integration (1970s)- S-TTL, CMOS, NMOS, MPU*
- *(Very) Large Scale Integration (1980s)- LS-TTL, CMOS, NMOS*
- *1990s - Microprocessors largely displace the minicomputer.*
- *By 2000, CMOS dominates, with millions of transistors per single Silicon die.*
- *Heat dissipation becomes a major issue for processor chips post 1990.*

Moore's Law

- 1965 paper by Gordon E. Moore “*Cramming more components onto integrated circuits*” predicted that the number of transistors (switches) would double every 18 months; this has been empirically adjusted more recently to around 2 years.
- Integrated circuits encompass CPU chips, I/O chips, memory chips, digital camera imaging chips, i.e. all components used in modern digital electronics and computers.
- Moores' Law is determined by the technology used to manufacture integrated circuits – how small can an individual switch be made?

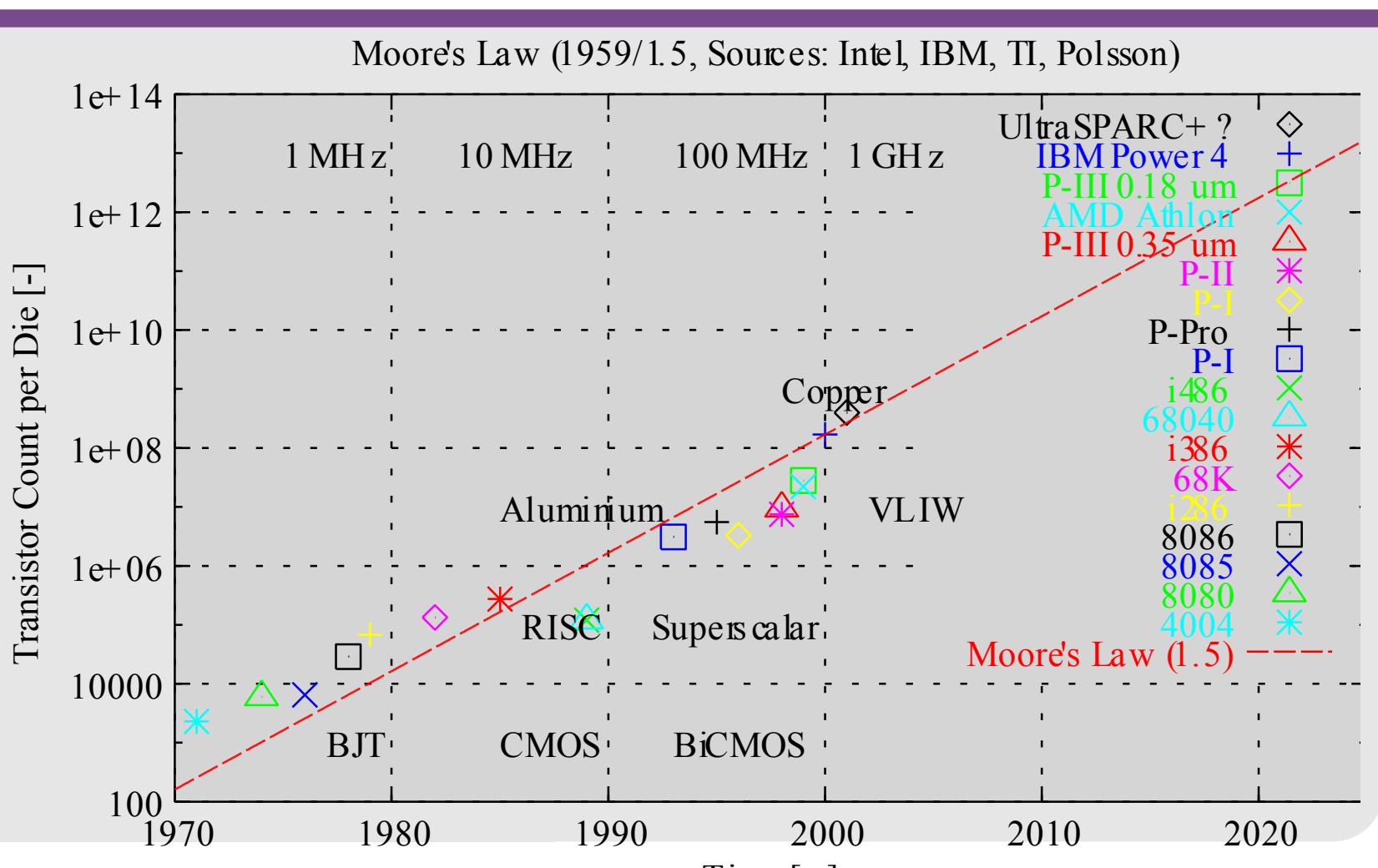


Impact of Moore's Law

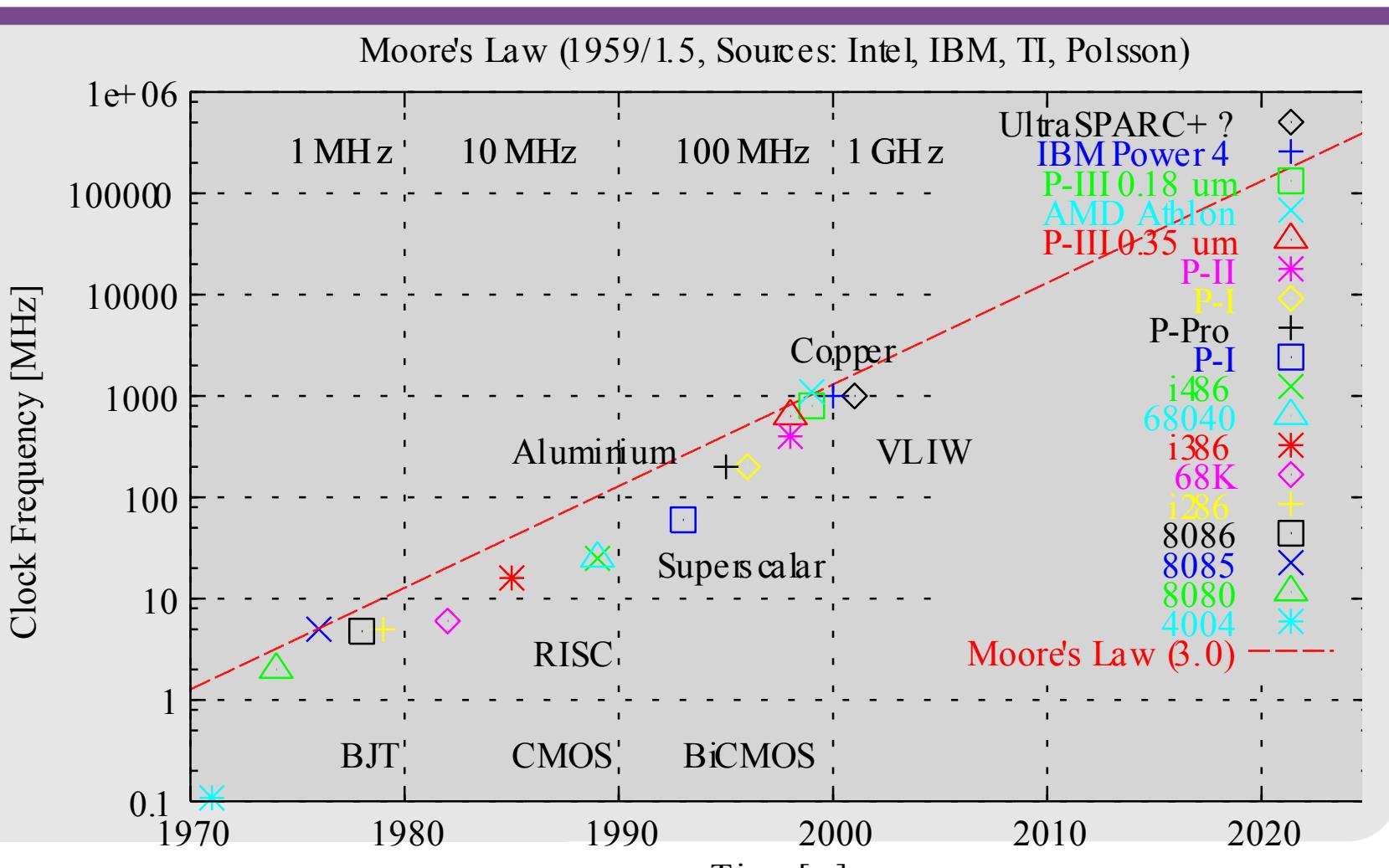
- Prof Carver Mead “clock speeds scale with the ratio of geometry sizes, as compared to transistor counts which scale with the square of the ratio of geometry sizes”.
- Square root relationship between switching speed and density growth.
- Empirical measurements validate this relationship.
- This results in two Moore's Law curves, one for density, and another for achievable clock speed.
- Density and clock speed both impact CPU performance.



Moore's Law - Chip Density

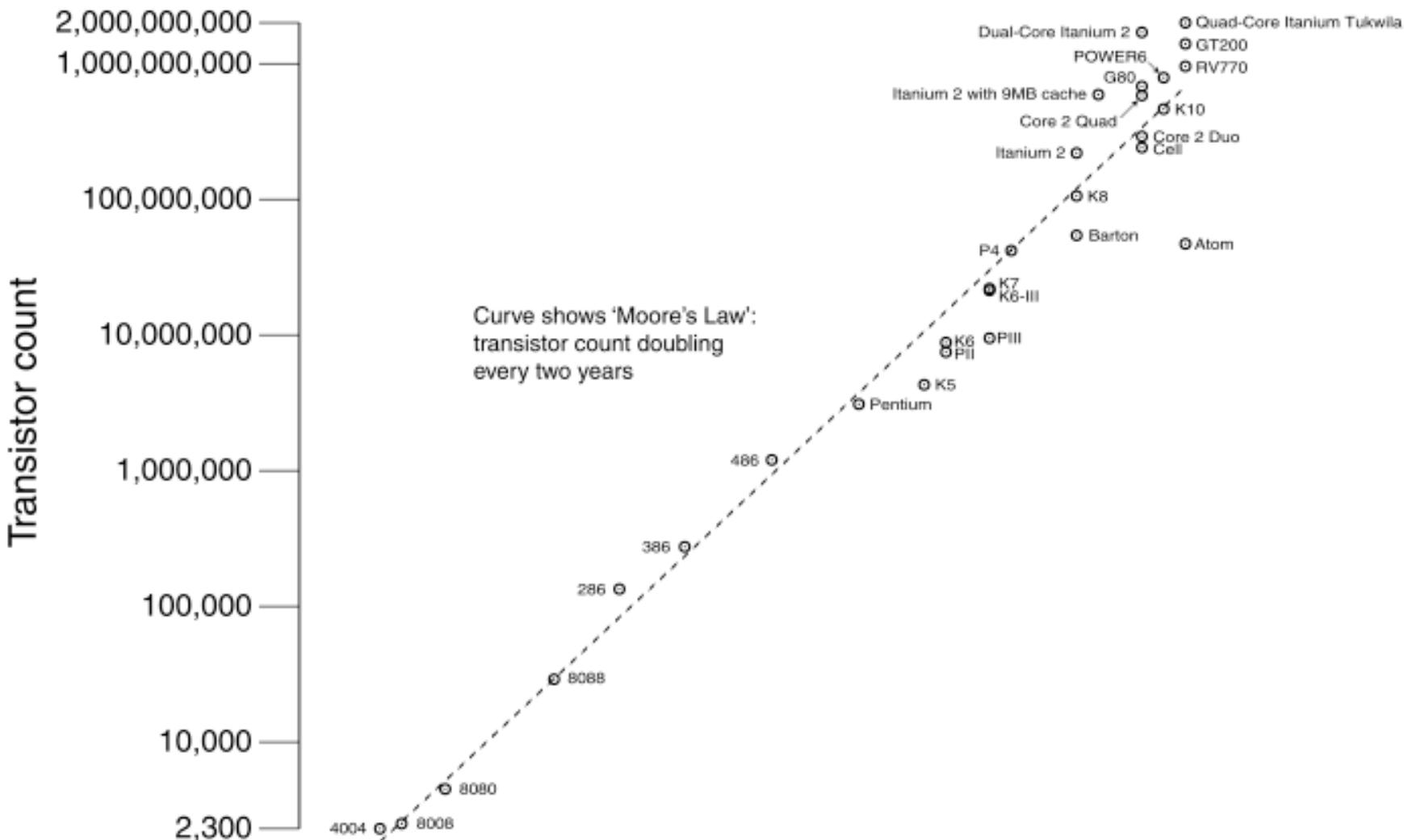


Moore's Law - Chip Clock Speed



CPU Transistor Counts 1971-2008 & Moore's Law

(Wikipedia chart – 2009)



Impact of Moore's Law – Clock Speed

- For “like” machine architectures and “compute bound” jobs, performance typically scales with clock speed.
- If a processor is clocked at 3 GHz, clocking the same processor design at 3.6 GHz allows a compute bound job to complete 20% earlier.
- Exploited by amateur “overclockers”.
- Not all jobs are compute bound.
- Not all internal architectures are the same, even if the CPU brand name is the same.
- *Not all cited “GHz” are the same as actual “GHz”.*



Impact of Moore's Law – Internal Density

- **Density impacts performance by allowing for more complex internal architectures in CPU chips.**
- **Larger caches minimise slower accesses to main memory and hard disk storage.**
- **Speculative execution techniques allow CPUs to exploit instruction level parallelism.**
- **Wider internal bussing in CPUs speeds up internal data transfers.**
- **Vector processing speeds up many calculation types.**
- **Multiple core permit parallel processing.**



Impact of Moore's Law – Parallelism

- Improved density allows exploitation of instruction level parallelism, and program level parallelism.
- Caveat: Amdahl's Law and other limits on speedup apply, as a result of data dependencies at program and instruction level.
- Multicore chips with 2, 4, 8 or more CPUs are a result of Moore's Law permitting very high density.
- *Performance is always a contest between techniques to improve clock speed and the ability to use parallelism, vs data dependencies in algorithms which impair performance.*

Impact of Moore's Law – Main Memory

- **Moore's Law applies to memory chips used in computer equipment, particularly for main memory storage.**
- **Twenty years ago computers operated with Megabytes of RAM, today PCs with Gigabytes are common.**
- **Memory resident application code and data can be accessed thousands of times faster than disk.**
- **The feasibility of some applications will depend on how much memory can be installed in a single computer.**
- **Cost vs time becomes a factor for memory.**



Limits to Moore's Law

- For thirty years analysts have predicted that “the end is near” and Moore’s Law growth will end due to limits in transistor manufacturing techniques.
- So far, every fabrication limit has been beaten and Moore’s Law has continued to apply.
- Hard limits will be eventually reached once transistor sizes become so small that quantum physical effects prevent reliable switching.
- As a result different computing technologies will be required, or techniques to increase parallelism.
- *In the near future, Moore’s Law will apply.*

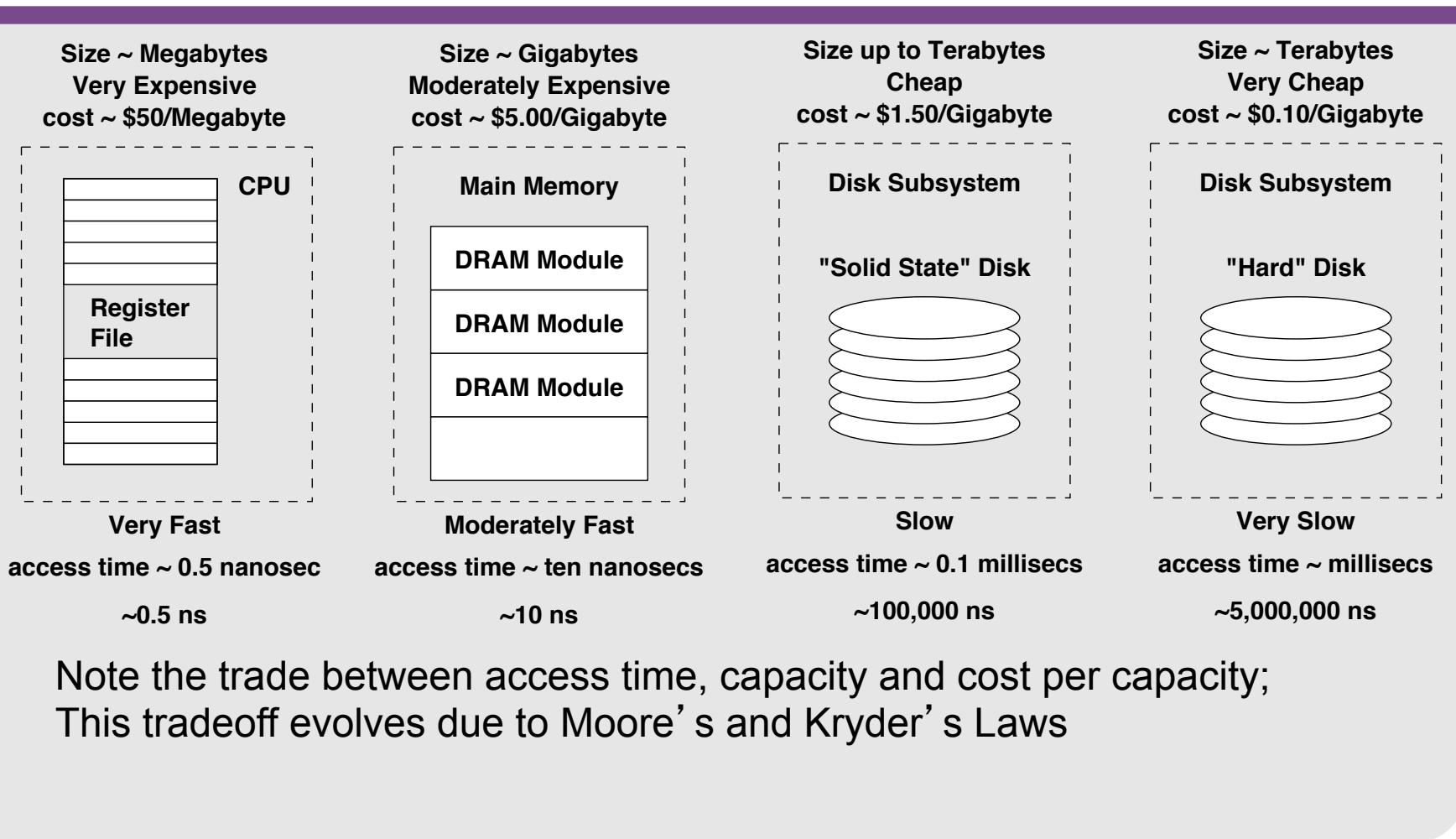


Moore's Law vs Application Design

- Moore's Law results in two important relationships for distributed applications:
 - A.The dollar outlay of providing some fixed amount of computing power declines over time.
 - B.The available computing power per CPU increases over time for a fixed dollar outlay.
- Economic feasibility of many distributed applications depends on the cost of computing power.
- Technical feasibility of many applications depends on the achievable computing power.
- ***Good planning requires these be considered.***

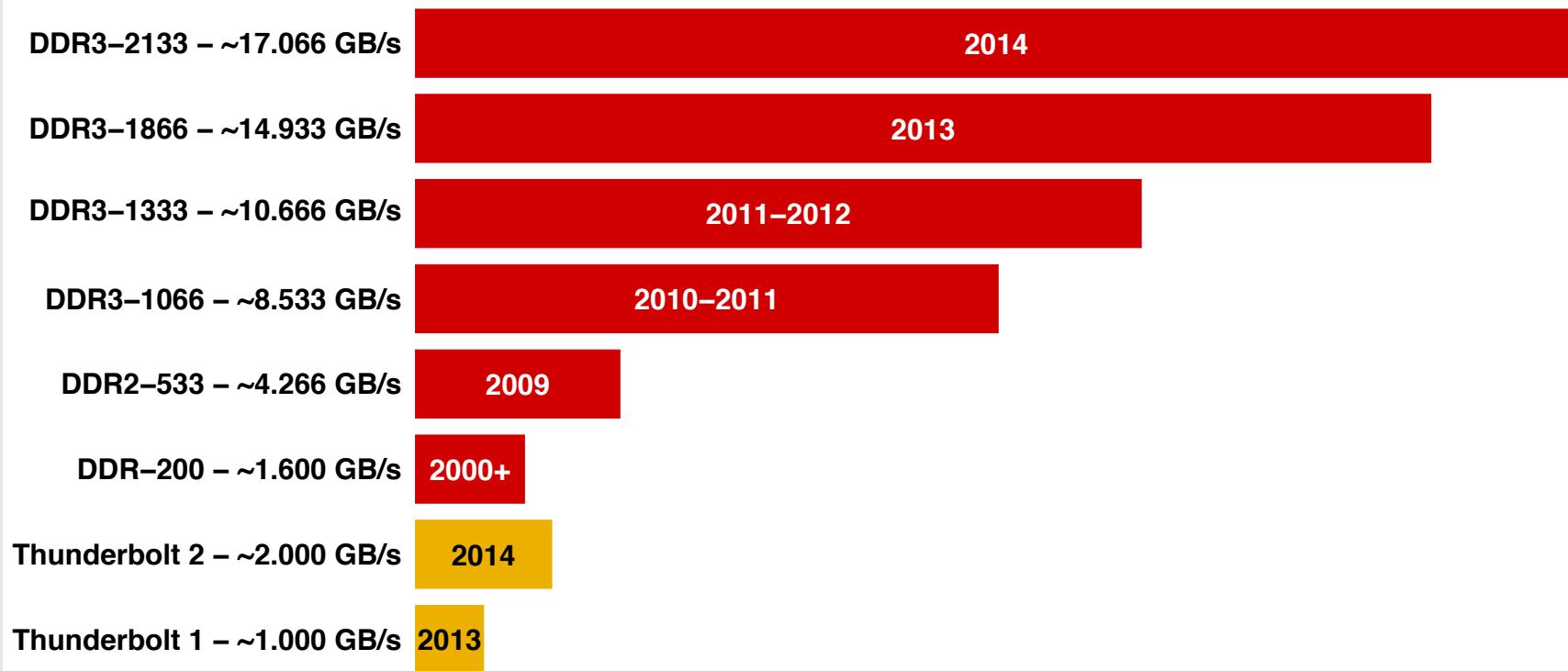


Storage Hierarchies - Today



Storage Hierarchies – SDRAM Memory Bandwidth

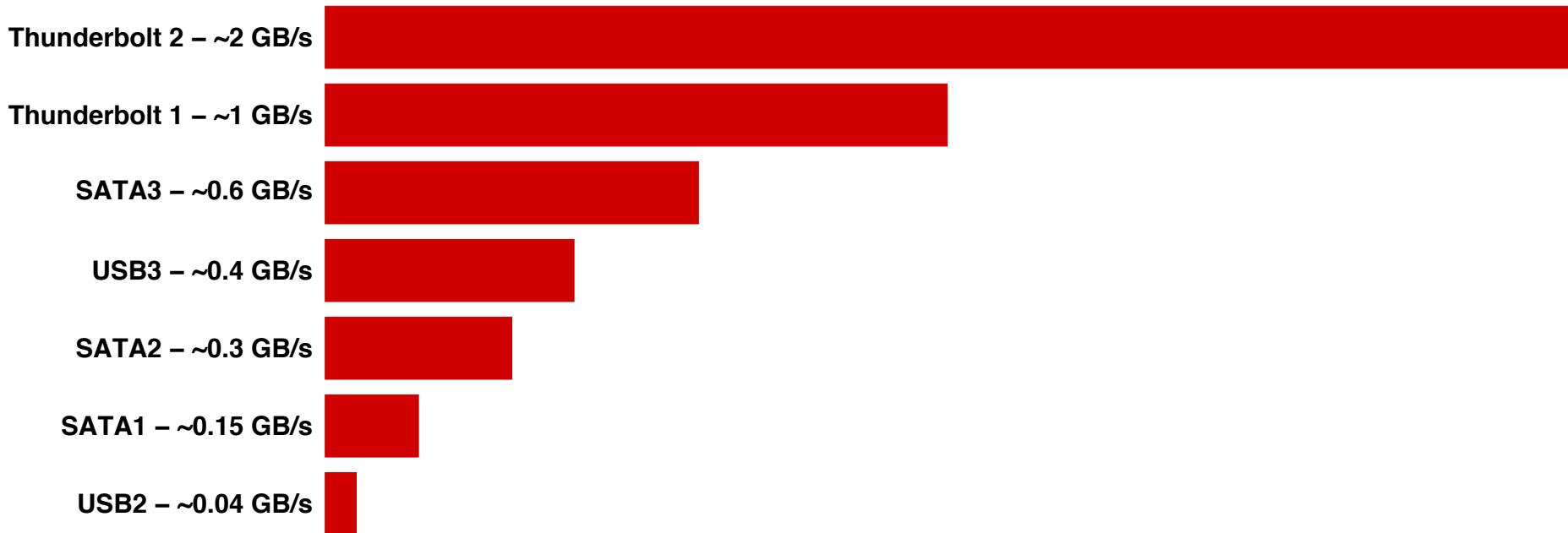
JEDEC SDRAM MODULE PERFORMANCE COMPARISON (PEAK/BURST)



Note that JEDEC standard SDRAM module availability overlaps between DDR, DDR2 and DDR3, with all three formats available in 2013. GPU SRAM – 25 GB/s+

Storage Hierarchies – Mass Storage Bandwidth

MASS STORAGE BUS PERFORMANCE COMPARISON (PEAK/BURST)



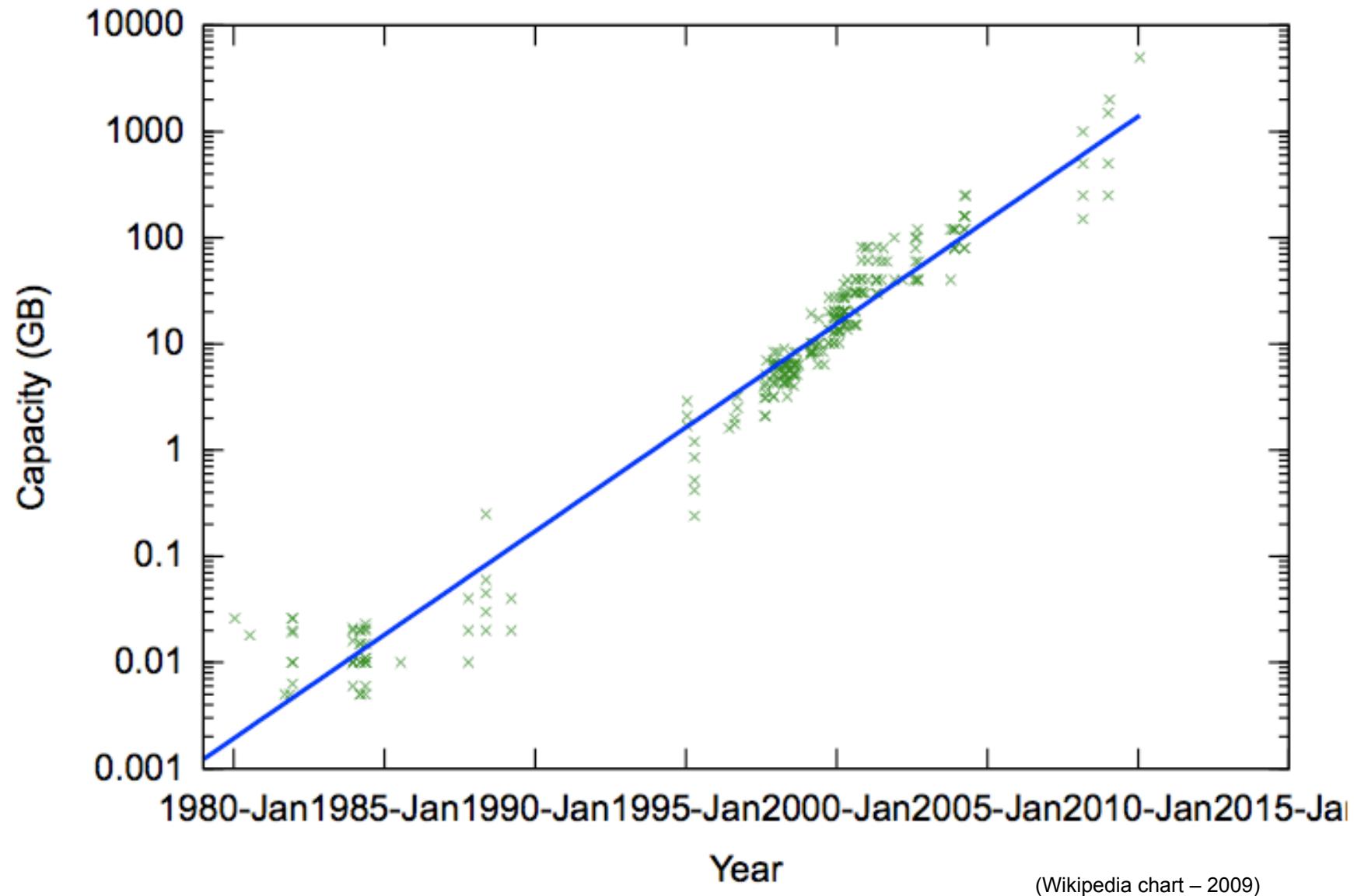
Note that achievable throughput performance varies with implementation, and in practice may fall between 50% and 90% of nominal peak/burst transfer rates

Disk Storage Performance Growth

- **Storage capacity per dollar spent improves over time.**
- **“Bandwidth” of disks improves over time.**
- **Access time to data generally does not improve strongly over time.**
- **Disk drive density obey’s “Kryder’s Law” – achievable capacity doubles every two years.**
- **Density improvements using magntoresistive and giant magnetoresistive disk head technology.**
- **Error rates improved by ECC block encoding of data.**



Kryder's Law - Disk Capacity Vs Time



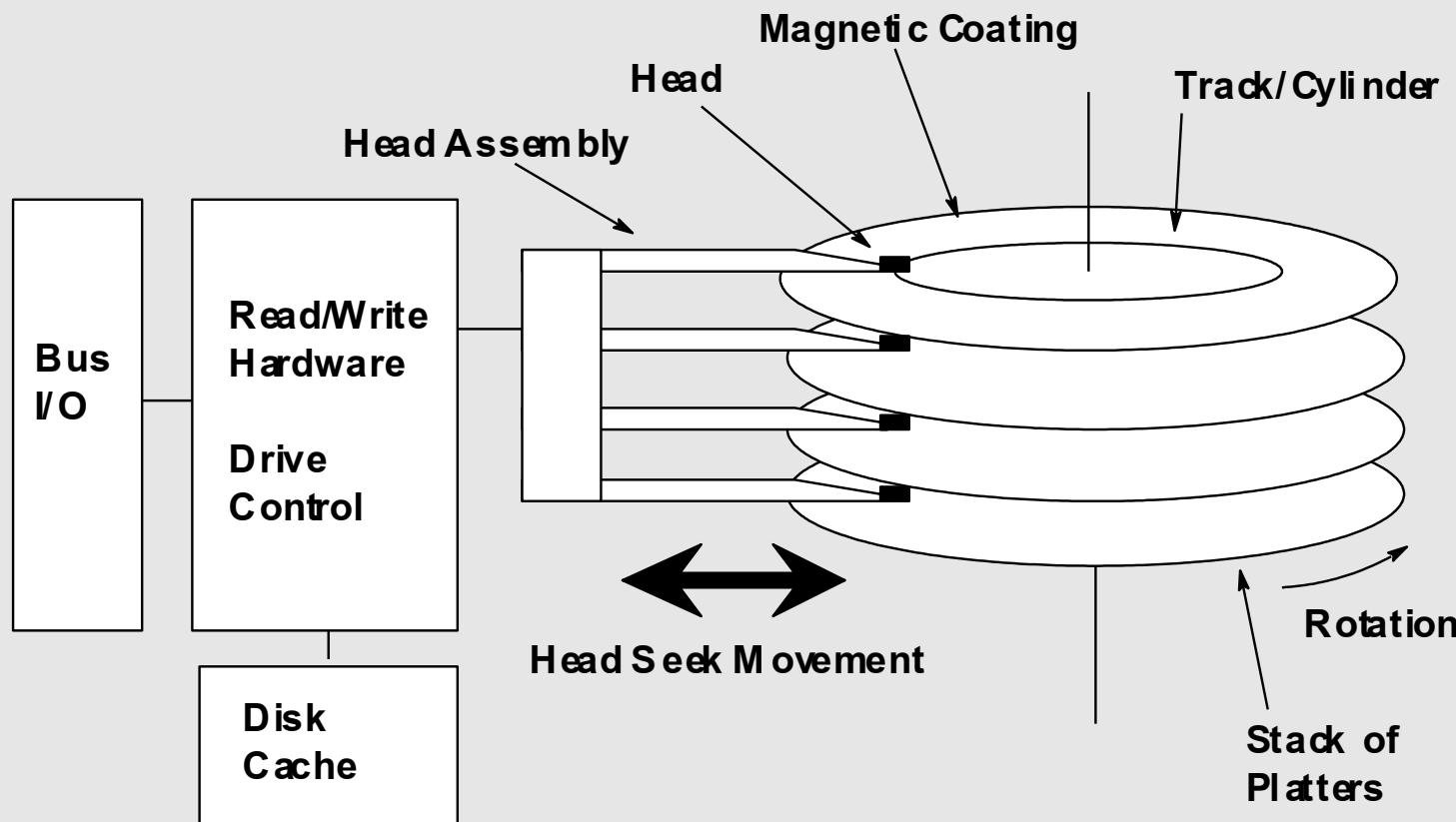
(Wikipedia chart – 2009)

Access Times in Disks

- Disk drives use a magnetic coating on a rotating disk platter to store information. Data is written in "cylinders".
- An electromagnetic transducer, termed a "head" is used to write or read from the disk surface.
- A pivoting mechanical arm or translating mechanical finger is used to position the head over the part of the disk where data is to be read or written.
- Mechanical access mechanisms put severe and hard limits on achievable access time performance.



Disk Drive Operating Principles



Disk Performance

- The stack of disk platters rotates at a constant RPM, in older disks 3,600 RPM, in newer types 4,500, 5,400 or 7,200 RPM.
- The time to access any item of data on the disk therefore depends on two parameters:
 - 1.Rotational Latency - the time it takes for platter to rotate into position under the head.
 - 2.Seek Time - the time it takes to move the head over the cylinder holding the data.

<Access Time> =

<Seek Time> + <Rotational Latency> [ms]

Disk Performance

- The average rotational latency is 50% of the time it takes for a disk to rotate through 360 degrees:
 - 3600 RPM - 8.33 milliseconds
 - 7200 RPM - 4.17 milliseconds
- The average seek time between two adjacent cylinders (tracks) depends on the mechanical design of the head system:
 - Typically 1.5-5 milliseconds
- <Access Time> is usually 6 -> 13 [ms]
- Extremely slow compared to main memory !



Disk Performance Vs Applications

- Many applications are “storage intensive” involving database lookup, multimedia retrieval, or data archive access or storage.
- Latency in access to disk resident data can have a major impact on application performance.
- Strategies may be required to address disk latency.
 1. Exploitation of memory caches – memory obeys Moore’s Law and is relatively “cheap”.
 2. Prefetching of disk resident data which is expected to be used soon.
 3. Speculative prefetching of data.

Advancing Magnetic Storage Technology

- *Giant Magneto-Resistive Effect (GMR)* heads – IBM a decade ago, now all manufacturers using licenced GMR technology;
- *Helium filled drive enclosures* – Western Digital; Helium permits cooler operation, allowing for more drive platters ~50-100% increased capacity;
- *Heat Assisted Magnetic Recording (HAMR)* – Seagate; HAMR uses a laser to preheat the point at which the magnetic head records, to change material properties and defeat “super-paramagnetic limit” – 100 x density gain;
- *Self-Ordered Magnetic Arrays (SOMA)* – Seagate; self-assembled, ordered and uniform nano-magnet arrays using FePt and CoPt compounds as recording medium;
- *It is expected these technologies will extend Kryder’s Law for at least another decade;*



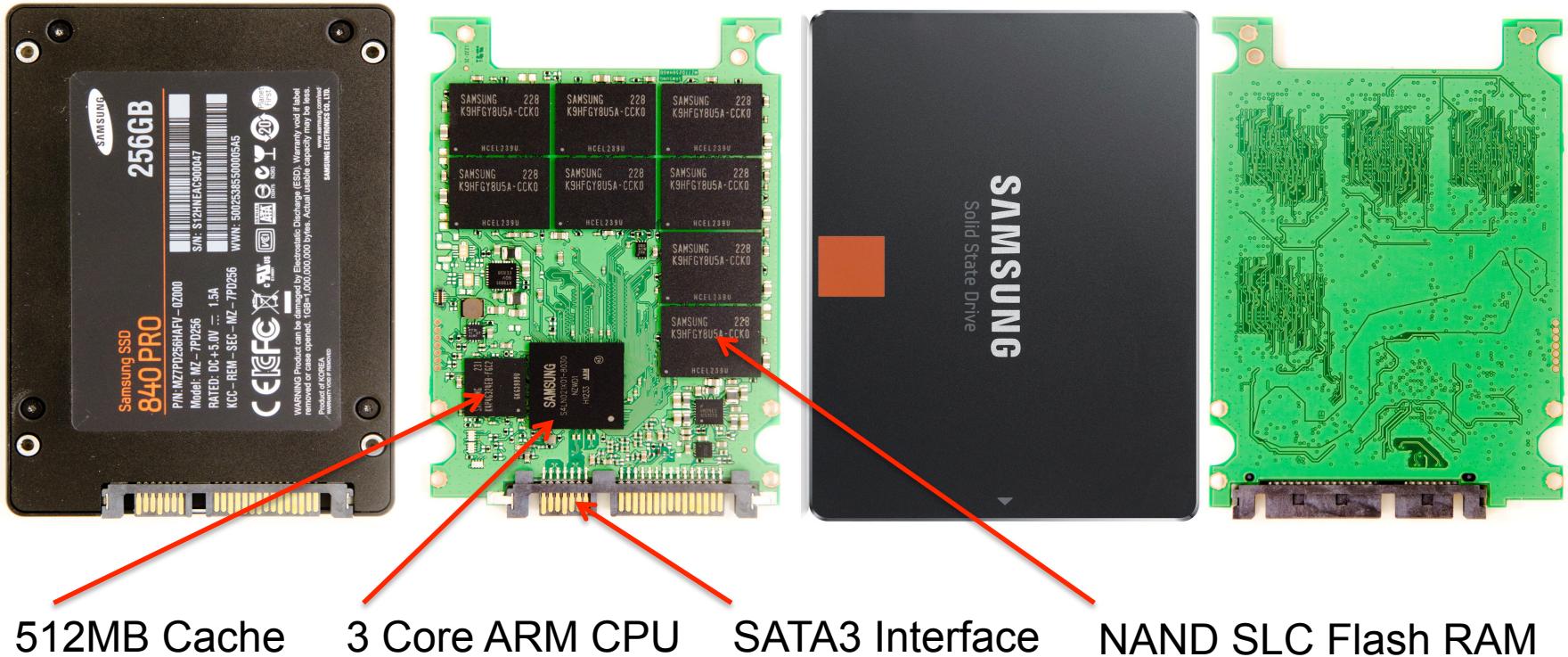
Solid State Disks

- **SSDs employ arrays of Flash RAM chips for nonvolatile mass storage of data;**
- **Flash RAM chips are used in commodity USB and SDHC media;**
- **An SSD device is accessed via a specialised embedded SSD controller chip, or chipset, with an industry standard SATA3 storage bus interface, and are packaged into the same “form factor” as industry standard “2.5 inch hard disks”;**
- **In most respects, an SSD qualifies as a lower power and quicker “drop in” replacement for a traditional rotating hard disk in the same physical and electrical format;**
- **Flash RAM employs a *write-read-erase-write-read* operating cycle, unlike magnetic media which use a *write-read-write-read* operating cycle – this has important ramifications.**

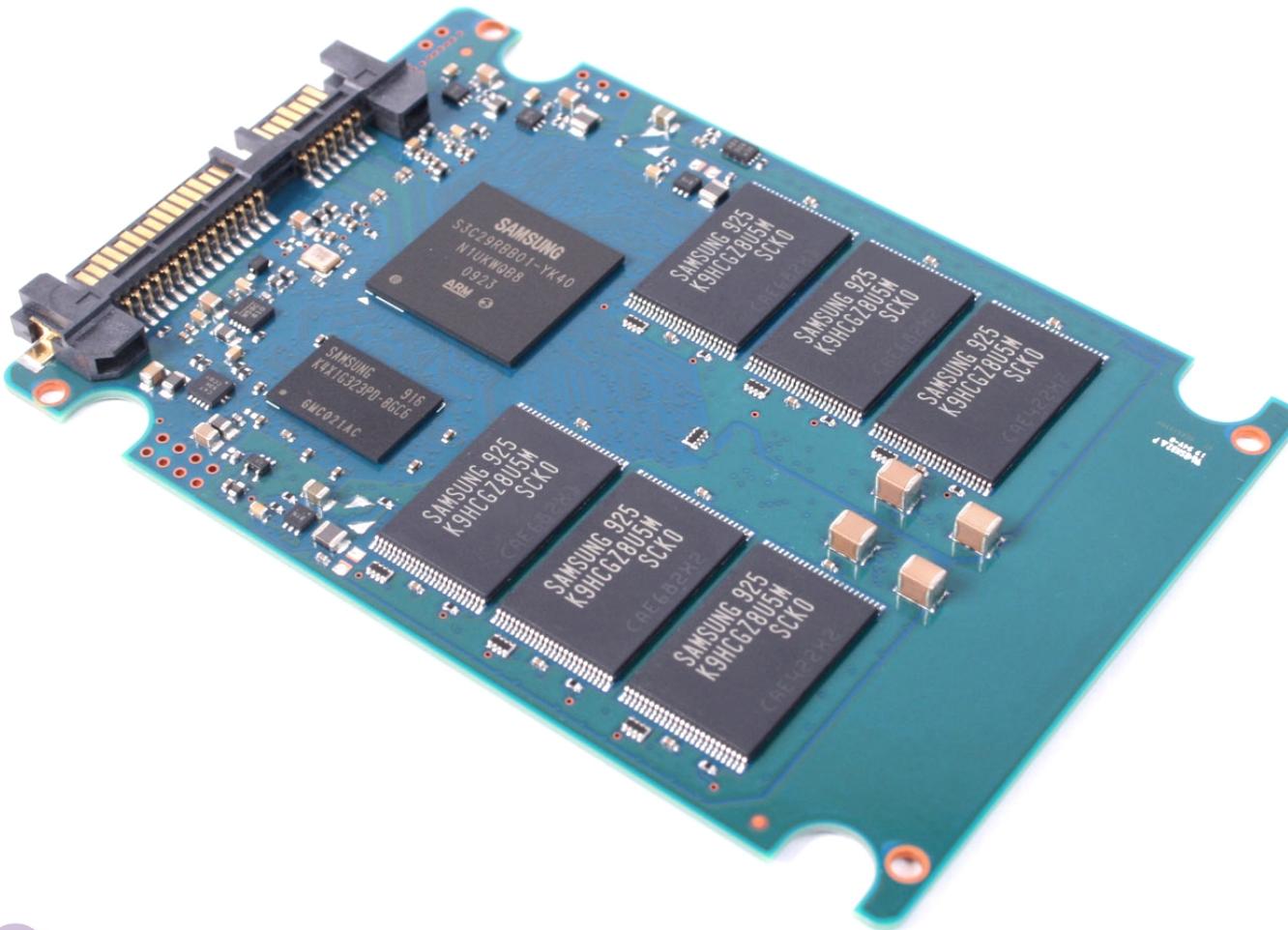


Samsung 840 Pro SSD 256 Gigabytes

- High performance “Professional grade” SSD developed for desktop systems;
- Lacks write endurance of much more expensive “Enterprise grade” SSDs;



Samsung 830 Series SSD – 2.5 inch SATA3



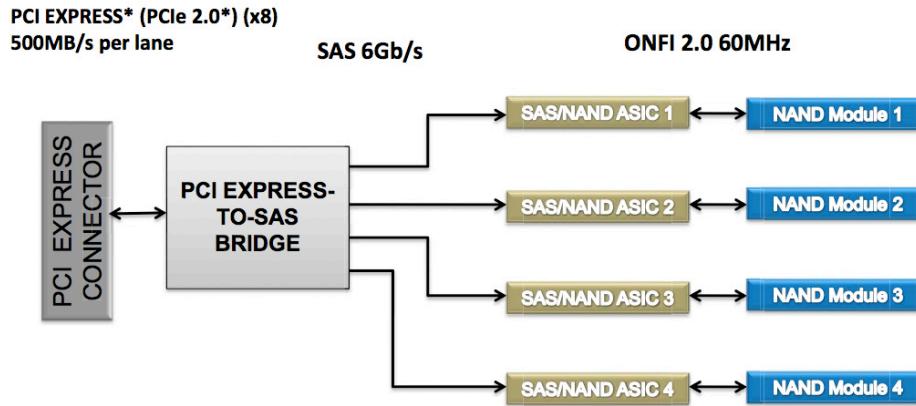
DDR3 I/O to SSD Storage - SMART ULLtraDIMM



- DDR3 memory bus providing 1 and 0.76 Gigabytes/s Read/Write aggregate transfer rates for sequential data – limited by Flash RAM/Controller hardware;
- Capacity of 200 GB – 400 GB, using RAID-like “Flexible Redundant Array of Memory Elements” overprovisioning;
- Very low access latency due to parallel I/O via DDR3 memory bus;
- Intended for Enterprise server systems with large numbers of DDR3 slots;
- Limitations: drivers available only for Windows Server 2008/12, Centos, RedHat, SUSE Linux, and VMware;
- First DDR3 packaged SSD to appear in the market.



PCIe I/O to SSD Storage – Intel 910 Series



- 8 Lane PCIe 2.0 providing 1 – 2 Gigabytes/s Read/Write aggregate transfer rates;
- Capacity of 400 GB – 800 GB with overprovisioning (896GB / 1792GB) to extend SSD life;
- Pricing is ~5.6 X commodity SSDs due to enterprise grade storage devices, redundancy;



What Next in SSDs?

- Commonly used SATA3 and mini-SATA3 I/O is now at its performance limits for commodity SSDs, and borderline for higher performing SSDs;
- PCIe 2 and 4 Lane SSDs for internal use inside chassis are now available in the market, supported by proprietary drivers mostly for MS Windows;
- DDR3 DIMM main memory bus SSDs – *Sandisk ULLtraDIMM* uses standard DDR3 SDRAM packaging and bus interfaces, supported by proprietary drivers for Linux, Windows, VMs;
- As with previous advances in hardware, uptake is slow as new hardware is poorly supported by software due to the absence of a common interface standard;
- *There is a significant gap between the operating systems and applications base, and advancing SSD / bussing technology!*



Optimising for SSD Performance

- Several steps necessary:
 1. Understand data access patterns to disk and to main memory, identify extant optimisations for legacy hardware;
 2. Identify most critical time complexity behaviours in mass storage access;
 3. Modify and adapt to exploit SSD;
 - During legacy application design, implicit assignment of per operation costs in time complexity modelling - assumed high latency for all mass storage;
 - Some memory resident structures may be migrated to SSD from main memory without major performance penalties;
 - Many mass storage resident structures will display dramatically different time complexity on SSD storage;

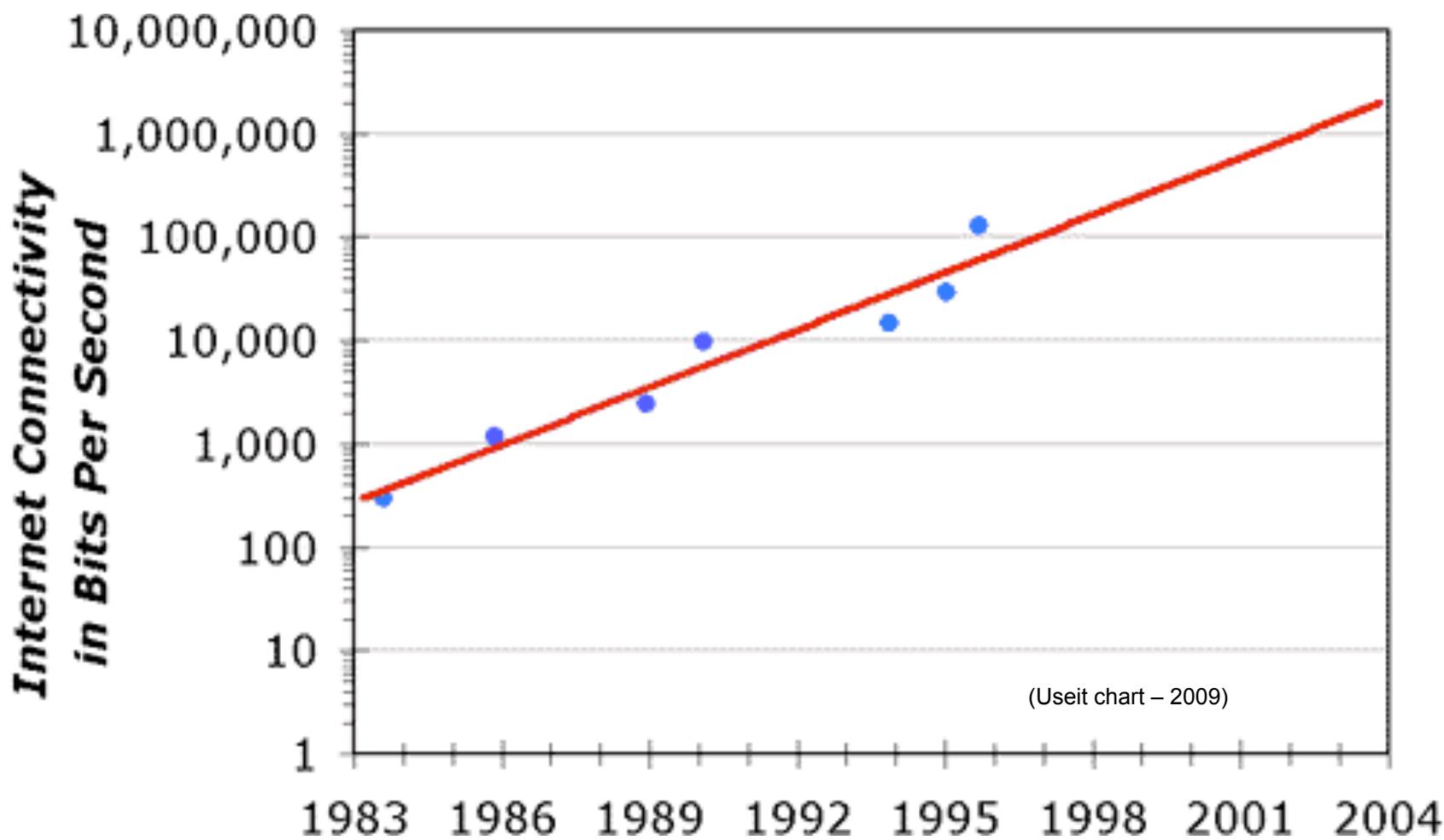


Network Performance – Bandwidth Law

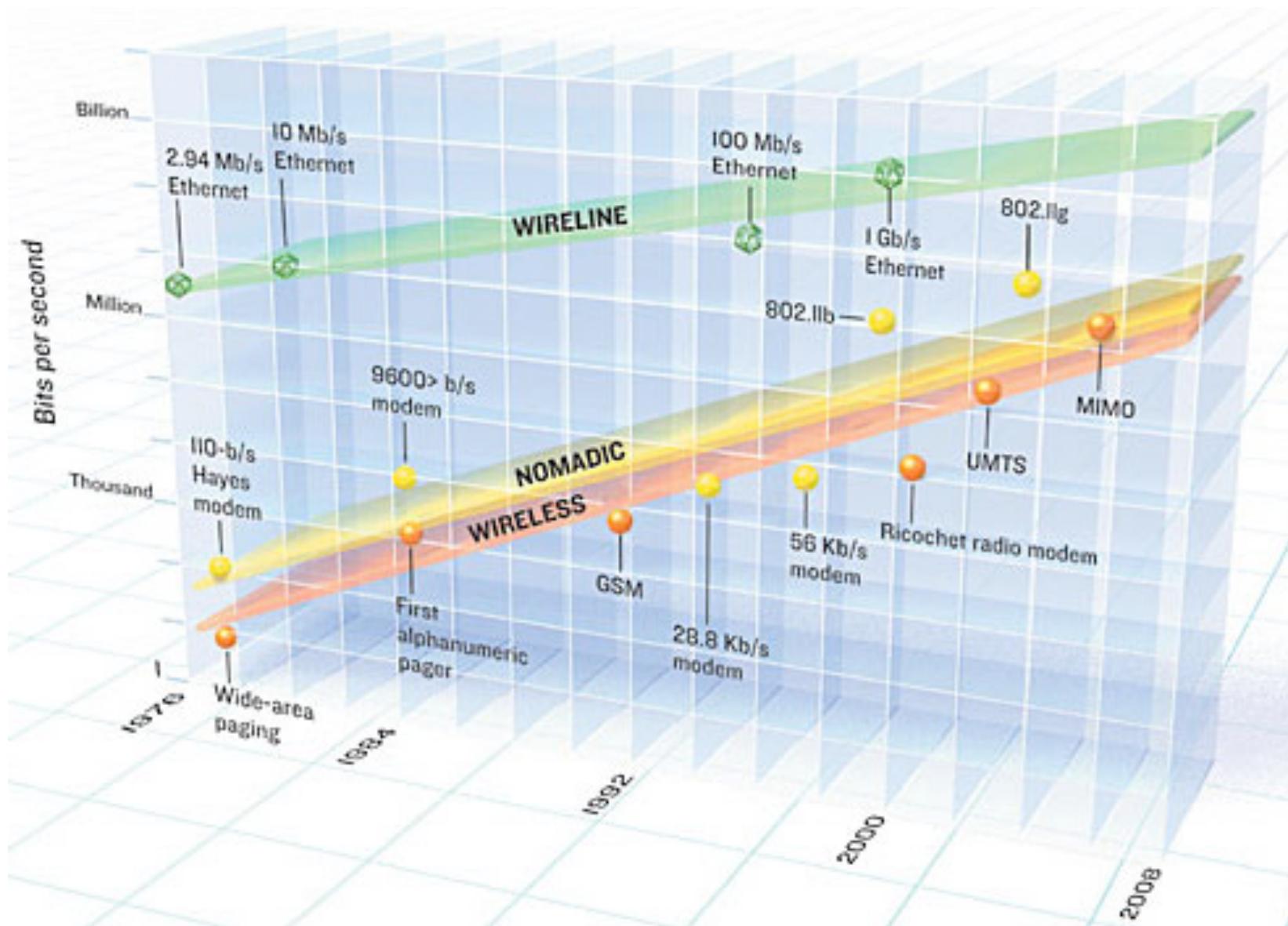
- Network data rate performance generally follows the “Bandwidth Law” which states that “Available bandwidth doubles every two years”.
- This is a result of performance improvements in optical fibre cable design, optical transmitters and receivers (and modulators), switching hardware (Moore’s Law driven) and the application of coding and compression techniques.
- Edholm's Law states that “three telecomm categories march in lock step: their data follow exponential curves, the slower rates trailing the faster ones by a predictable time lag” [IEEE].



Bandwidth [Nielsen's] Law



Edholm's Law of Bandwidth [IEEE Spectrum]



Bandwidth Law

- Bandwidth Law applies primarily to “cabled” network technologies using optical fibres or copper cables – examples being telecommunications links and LAN links.
- In general, the Bandwidth Law does not apply to wireless radio frequency links, due to the impact of antenna / transmitter performance, radio propagation environment and spectral congestion.
- *Edholm's Law is likely to break down as spectral congestion increases with the progressive saturation of useful wireless radio bands.*
- No limits in sight at this time for optical fibre links.



Bandwidth Law Vs Application Performance

- Like Moore's Law, the Bandwidth Law produces two important effects:
 - A. The dollar outlay of providing some fixed amount of bandwidth declines over time.
 - B. The available bandwidth per link increases over time for a fixed dollar outlay.
- Economic feasibility of many applications depends on the cost of networking.
- Technical feasibility of many applications depends on the available link bandwidth.
- *Good planning requires these be considered.*

Other Considerations – Metcalf' s Law

- Metcalf' s Law was defined for client server network applications such as the W3, it states that “the utility of a network increases with the square of the number of nodes in the network”.
- The law has been empirically validated for commercial sales and advertising impact.
- *Metcalf' s Law is not generally true for all applications, in many it simply does not apply.*
- Metcalf' s Law depends on the “utility function” which is “how do we measure the usefulness or productivity of a networked application?”.
- ***Metcalf' s Law often used to overstate benefits!***

Conclusions

- The performance and economic viability of distributed applications can depend strongly on compute performance, storage capacity and network performance per dollar outlay.
- Moore's Law, Kryder's Law and Nielsen's Law strongly impact distributed application performance and economics.
- Planning distributed applications requires that we consider the impact of all three laws, and where they might or might not apply in the application.
- *Performance evolution may not overcome problems with algorithm data dependencies.*

Reading/References

- [ftp://download.intel.com/museum/Moores Law/
Articles-Press Releases/
Gordon Moore 1965 Article.pdf](ftp://download.intel.com/museum/Moores_Law/Articles-Press_Releases/Gordon_Moore_1965_Article.pdf)
- <http://www.spectrum.ieee.org/jul04/3922>