Quantum Computer Architecture

Rod Van Meter
Keio University
rdv@sfc.wide.ad.jp

FIRST Project
Kyoto Summer School
2011 Aug 17

tum Architecture 

D-Wave Systems sells its first Quantum Computing System to Lockheed Martin Corporation

May 25th, 2011

VANCOUVER, BC, MAY 25, 2011 - Lockheed Martin Corporation (NYSE: LMT) has entered into an agreement to purchase a quantum computing system from D-Wave Systems Inc.

Lockheed Martin and D-Wave will collaborate to realize the benefits of a computing platform based upon a quantum annealing processor, as applied to some of Lockheed Martin's most challenging computation problems. The multi-year contract includes a system, maintenance and
vices.

g systems that leverage the physics of
r to address problems that are hard for
in a cost-effective amount of time.
include software verification and validation,
ty mapping and sentiment analysis, object
cal imaging classification, compressed


D-Wave flux qubit,
M.W. Johnson *et al.*, Nature, 2011

WIDE

KEIO UNIVERSITY CALAMVS GLADIO FORTIOR 1858

Aqua : Adva

I’m going to build a large-scale quantum computer.

Not sure yet what kind, or even when...

Let’s talk.

State of the Art in Quantum Computer Architectures

Rodney Van Meter

August 13, 2011

Abstract

Quantum computer architecture as a field remains in its infancy, but carries much promise for producing machines that vastly exceed current classical capabilities, *for certain systems designed to solve certain problems*. It must be recognized that *large systems are not simply larger versions of small systems*. These notes review the fronts on which progress must be made for such systems to be realized: experimental development of quantum computing technologies, and theoretical work in quantum error correction, quantum algorithms, and computer architecture. Key open problems are discussed from both a technical and organizational point of view, and specific recommendations for increasing the vibrancy of the architecture effort are given.

Keywords: quantum computation, quantum error correction, quantum computer architecture

1 Introduction

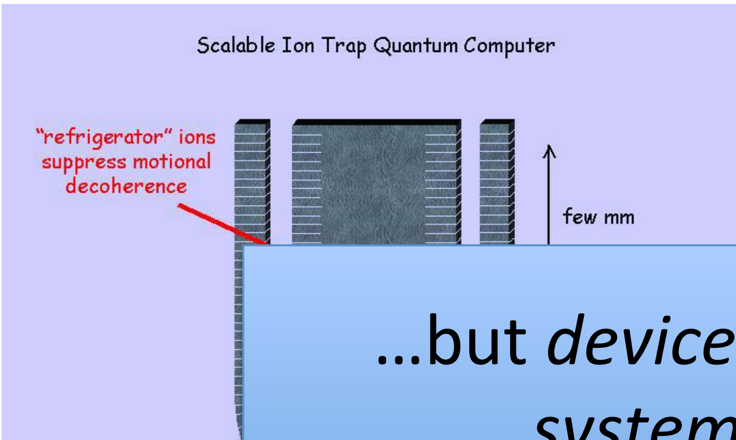
> When will the first paper appear in *Science* or *Nature* in which the point is the results of a computation, rather than the machine itself? That is, when will a quantum computer *do* science, rather than *be* science?

This question provokes answers ranging from, “Already have,” (in reference to analog quantum simulation of a specific Hamiltonian) to “Twenty years,” to “Never,” – and all these from people actually working in the field. I will try to shed a little light on how such varying answers can arise, and more importantly, how we can change that equation.

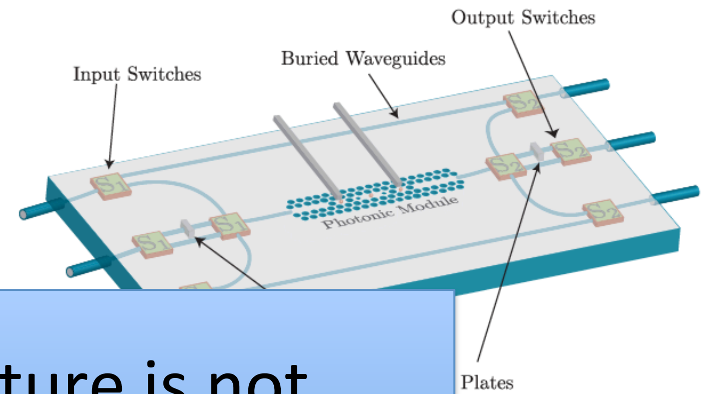
This informal set of notes accompanying the FIRST 2011 Quantum Computing Summer School is intended to convey the current state of the art in designing and building large-scale quantum computers, that is, the art of quantum computer architecture. I do not present other major areas of quantum technology such as quantum key distribution (QKD) or quantum repeater networks. I am particularly happy to discuss quantum repeaters if the occasion arises. Naturally, everything written and said is from the point of view of the author/presenter only, and occasionally is over-stated to clarify rhetorical arguments.

Aqua : Advancing Quantum Architecture

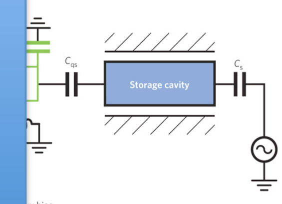
Many Types of Hardware



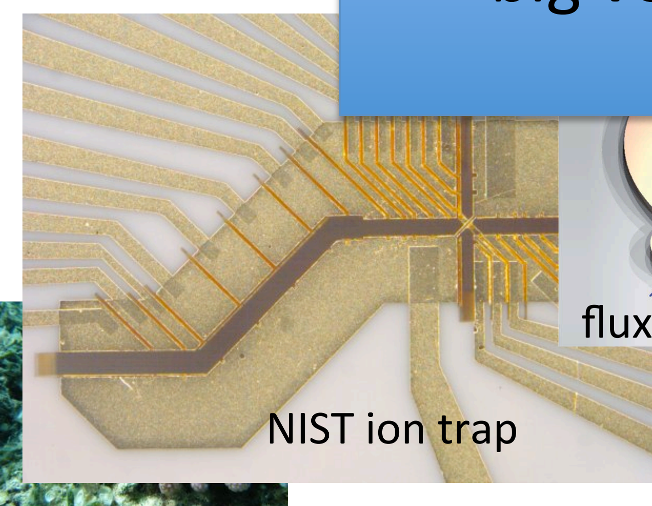
ion trap, Monroe Lab
Michigan



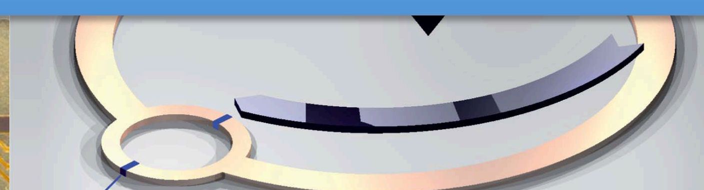
qubits



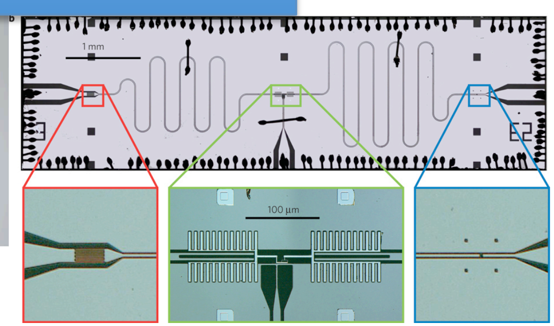
…but *device* architecture is not *system* architecture.
Large systems are much more than big versions of small systems.



NIST ion trap

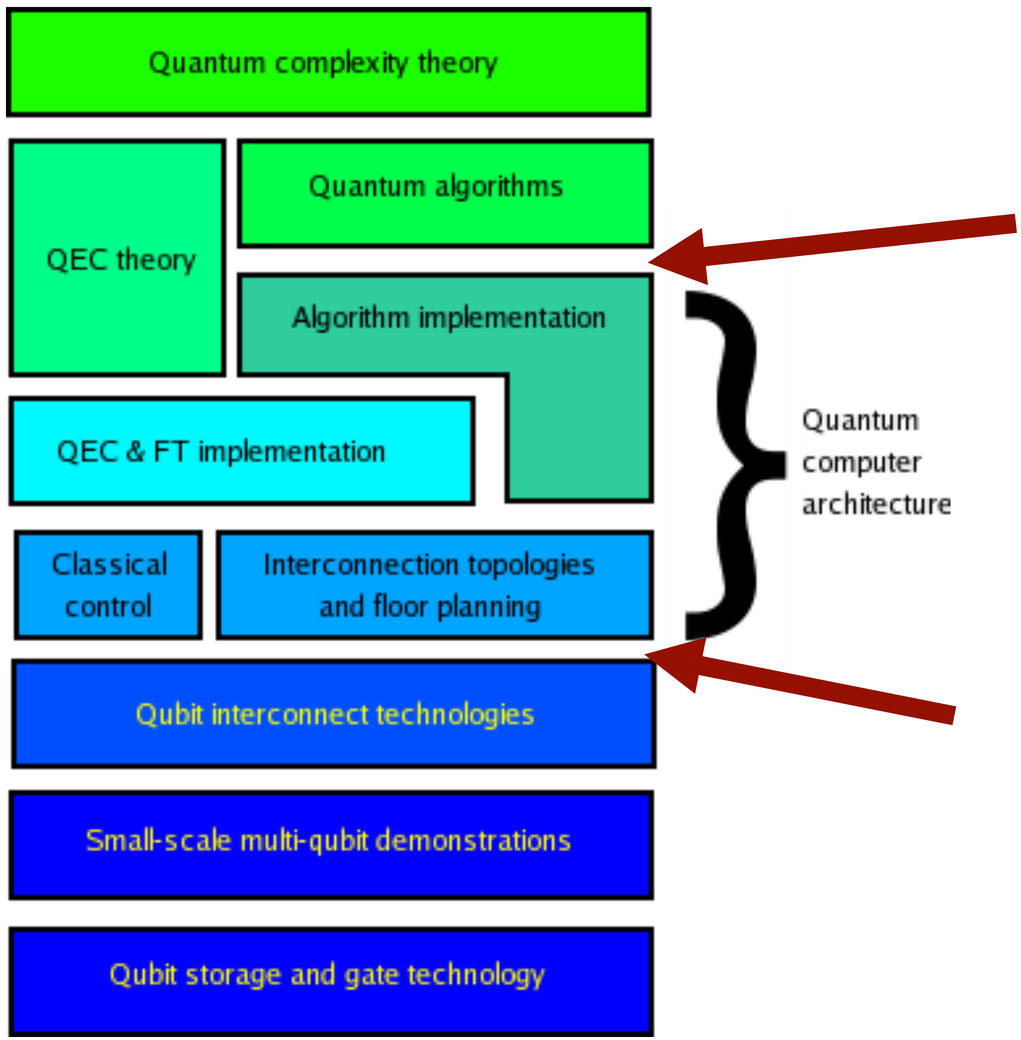


flux qubit, D-Wave, Nature, 2011



Yale transmon Nat. Phys 2010

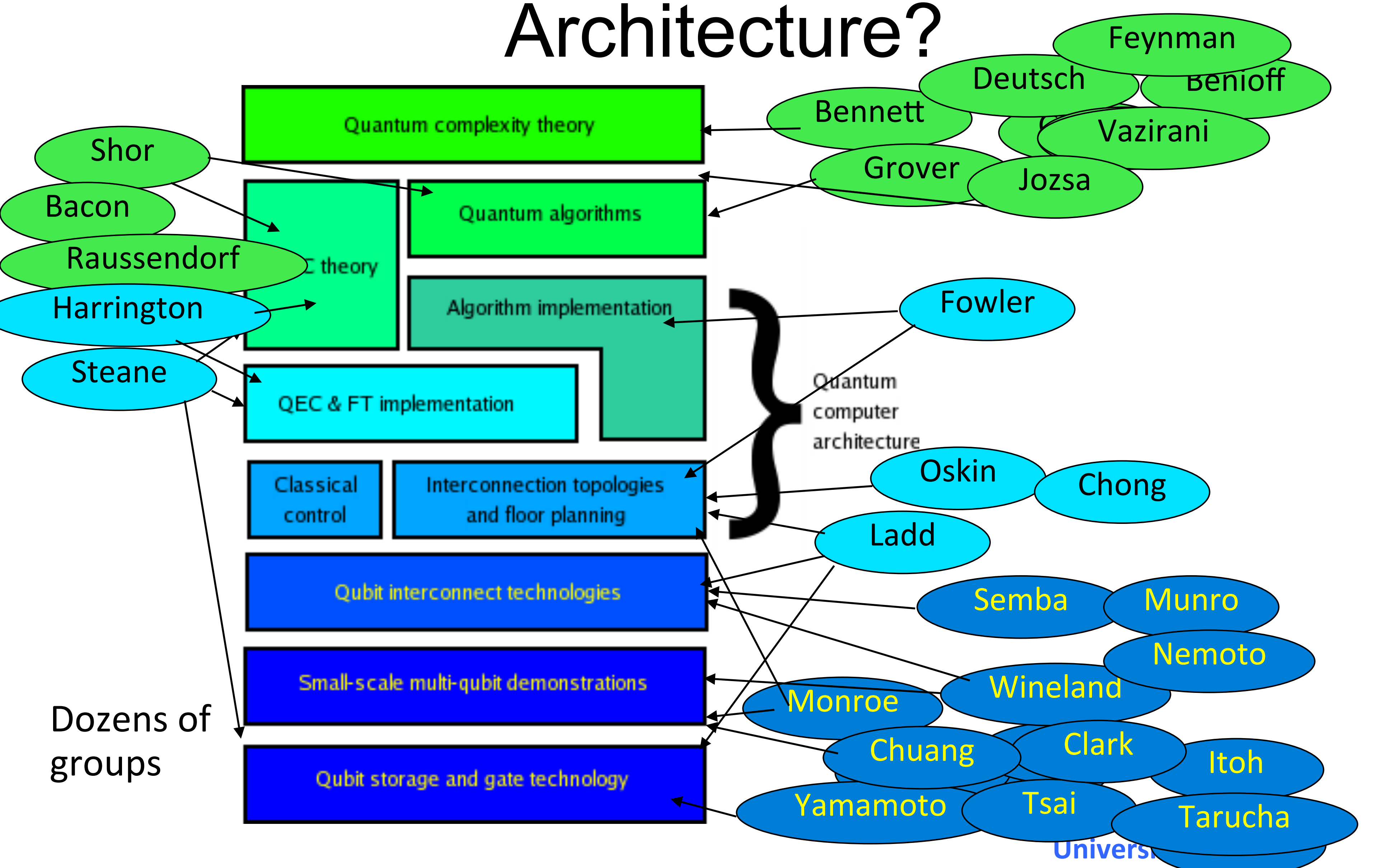
What is Quantum Computer Architecture?



Balance generality and performance

Extract commonality (e.g., qubit motion)

Who's Doing Quantum Computer Architecture?



Definition of a Quantum Computer

- A quantum computer is a machine that performs quantum error correction; quantum computation is merely a side effect.
- --paraphrased from Andrew Steane?

Outline

- Quantum computing's classical problem
- Classical computing's quantum problem
- Graduate computer architecture in six slides
- How to design a quantum computer
- Moving data
- *KQ* and what it means for you
- Putting it All Together: Understanding quantum computer performance

Quantum Computing's Classical Problem

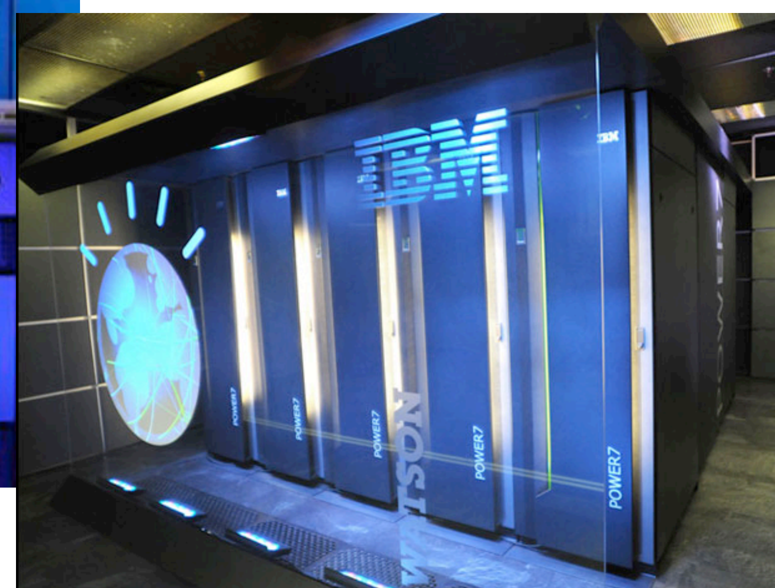
京

京

- FLOPS (float point ops/second): 8×10^{15}
- Gates per 64-bit FP multiply: $\sim1\times10^{5}$?
- Gates per second: $>1\times10^{21}$
- One month (seconds): 2.5×10^{6}
- Total computation (gates/month): 2.5×10^{27}
- Might get 1000x better in next decade
- Can QC solve bigger problem than 2.5×10^{30} classical gates in a month, for less than a billion dollars?

Supercomputing is Big Data

Supercomputing today is not about processing power *per se.* It is about turning enormous amounts of raw data into useful information.



Insurmountable Opportunity

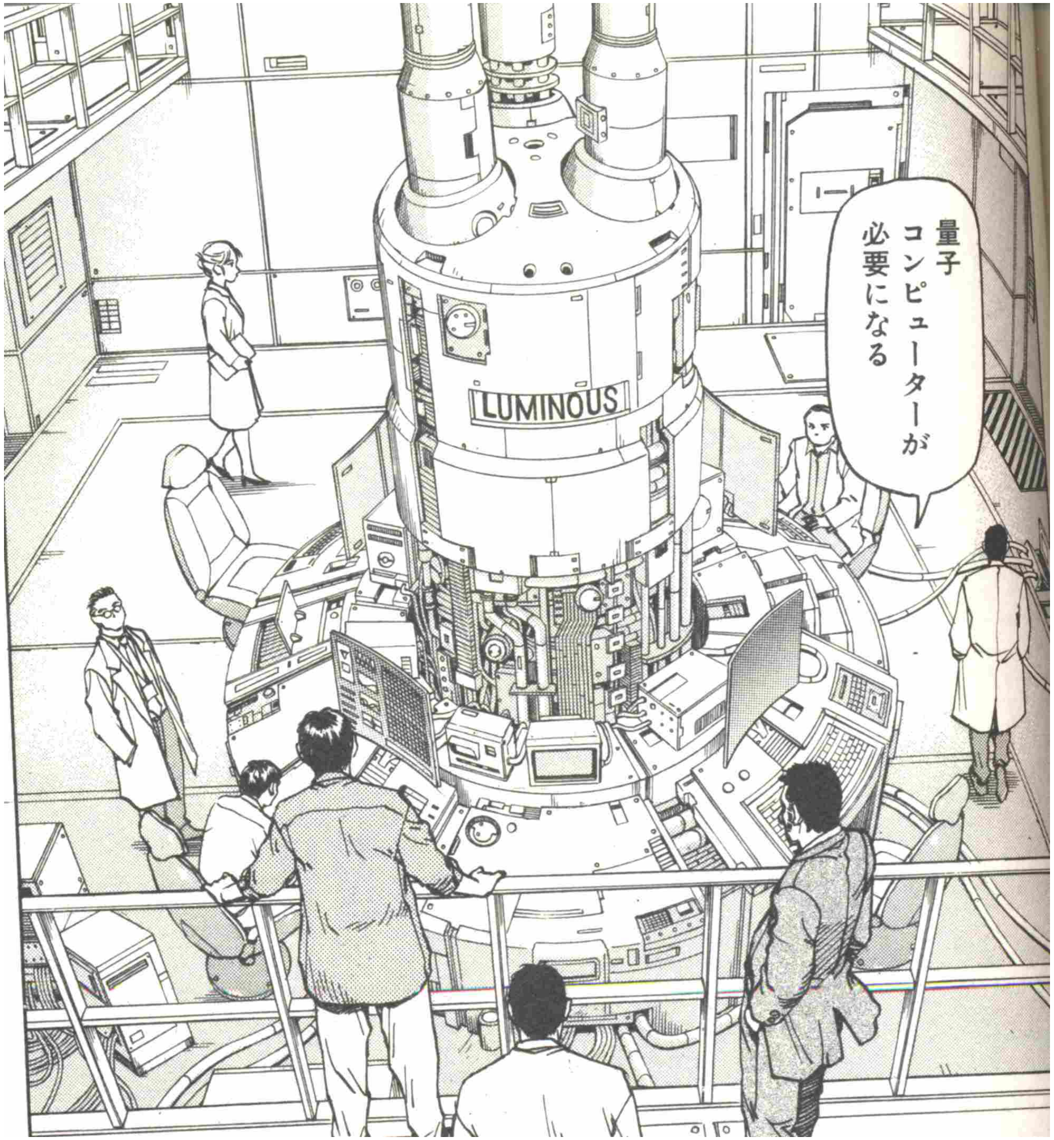
"We are confronted with insurmountable opportunities." Walt Kelly

Quantum computing will, indeed, *must*, open new fields of applications, mostly heavily mathematical. QC will probably be of minimal use on existing SC applications.

Hold that thought....

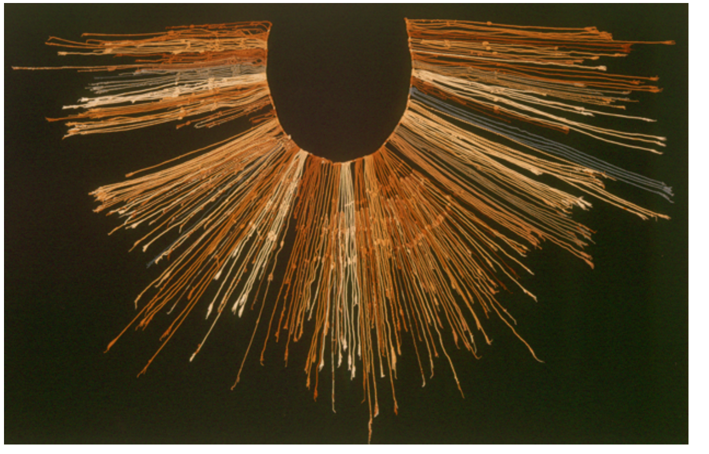
Classical Computing's Quantum Problem

One Like This?



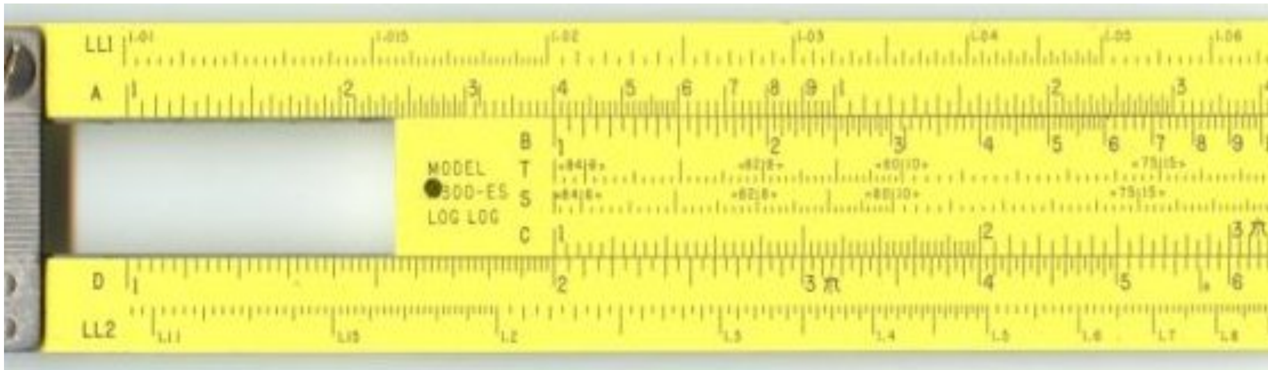
Data Recording Technology

abacus/そろばん
3-4000 years ago

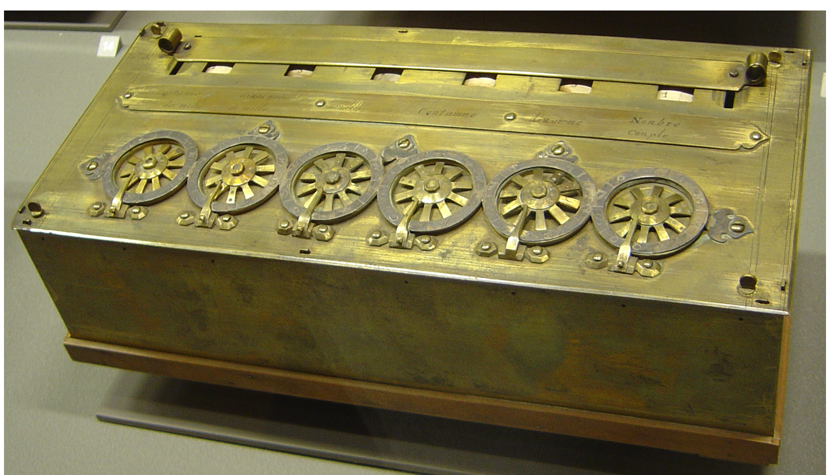


quipu
(Inca, 500 y.a.)

Semi-automatic Computation

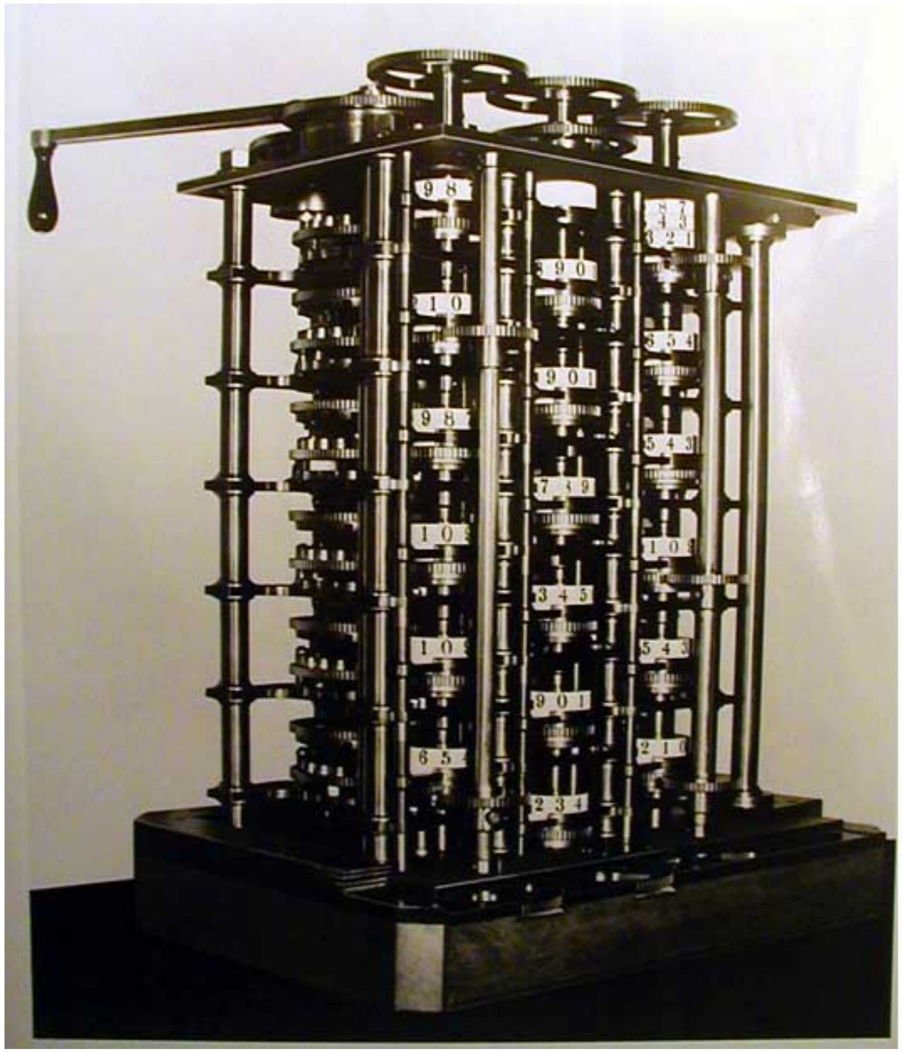


計算尺
(slide rule)
1620



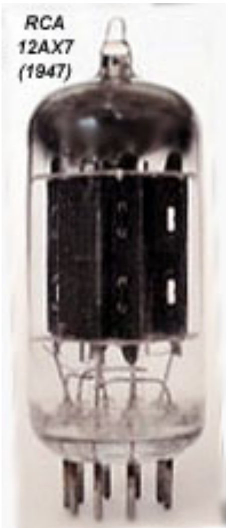
Pascaline
Blaise Pascal
addition
1643

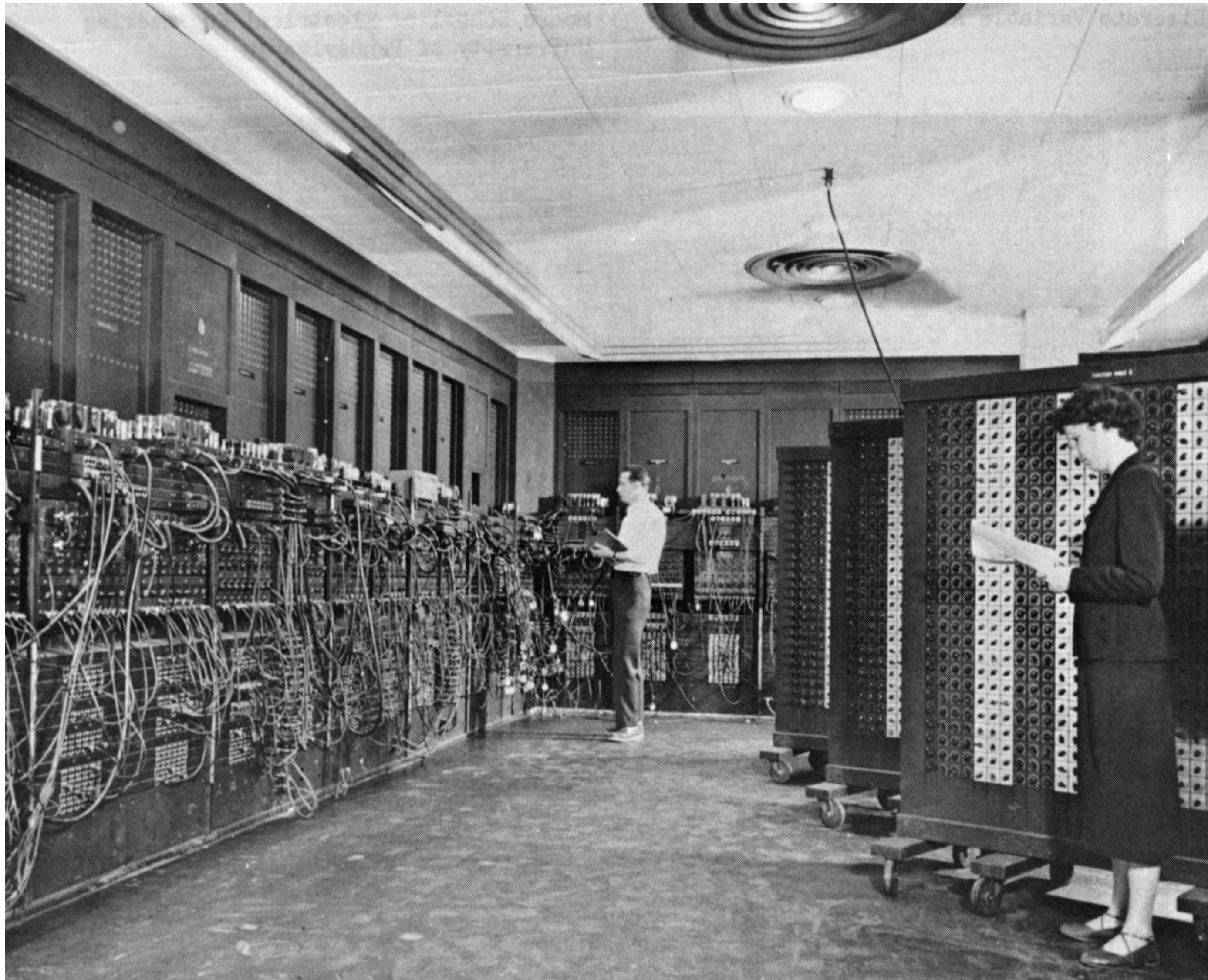
Babbage: Genius of the Steam Age: Calculating Polynomials



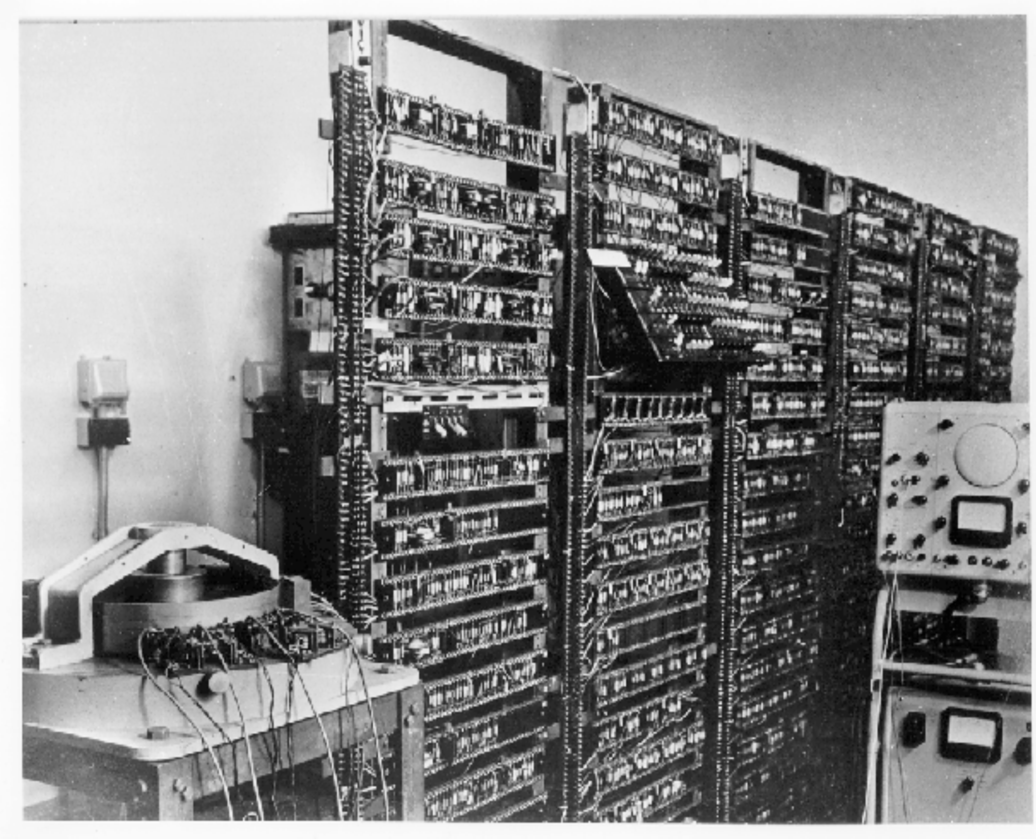
Aqua : Advancing Quantum Architecture

Vacuum Tubes: ENIAC、1948

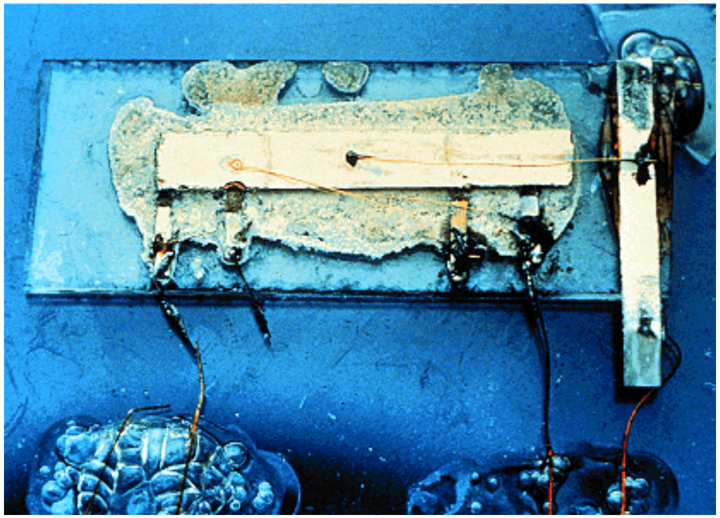


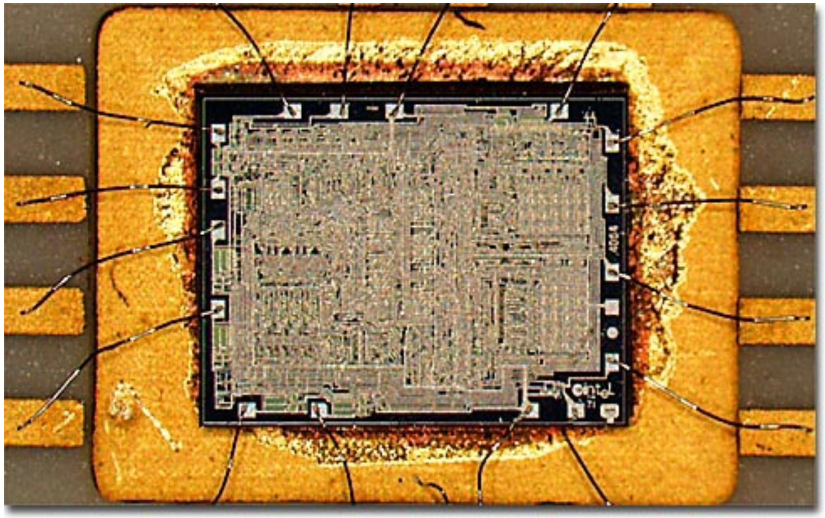


Transistor: 1948-1953



Integrated Circuit: 1958-



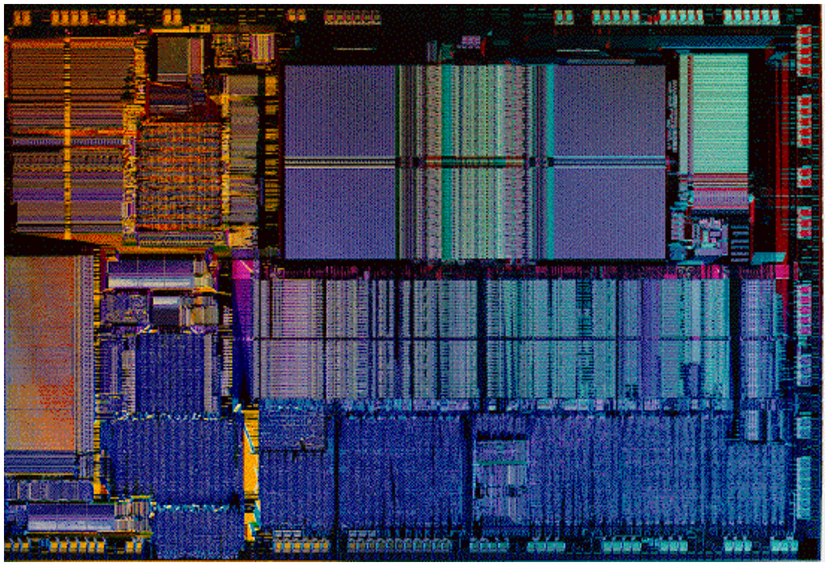


Intel 4004 uP, 2,300 transistors

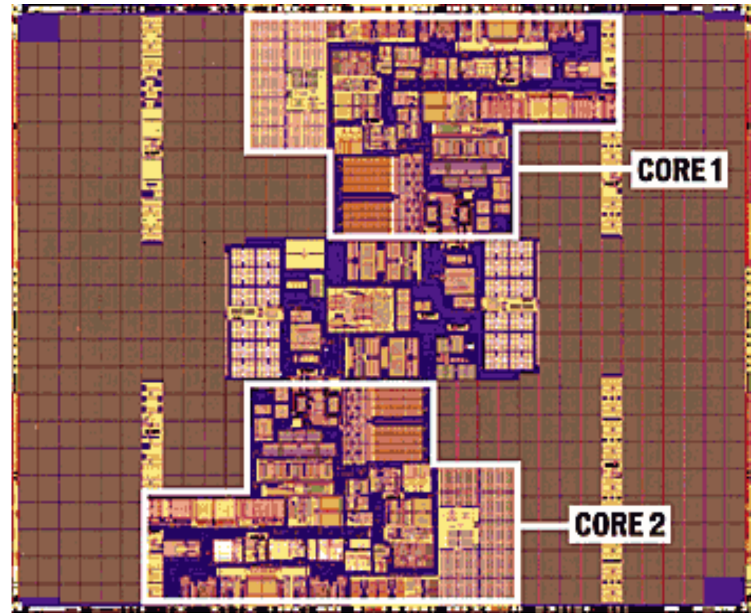
First IC
1958

First microprocessor
1971

Integrated Circuit: 1958-



Intel i486



Intel Montecito, 90nm process, ~100W

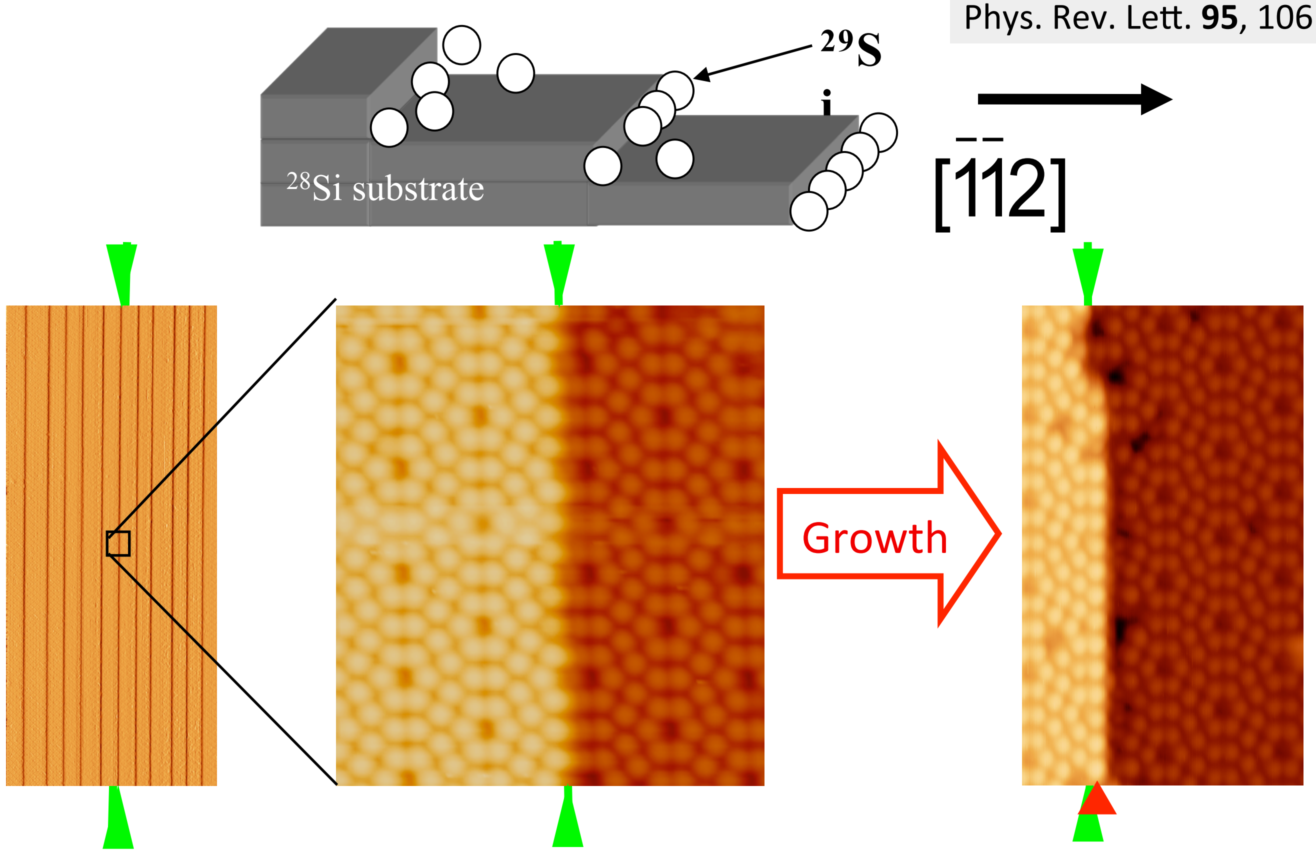
1,000,000
トランジスタ
1989

1,720,000,000
トランジスタ
2005

Si single single wire fabrication

Appl. Phys. Lett. **87**, 031903 (2005)

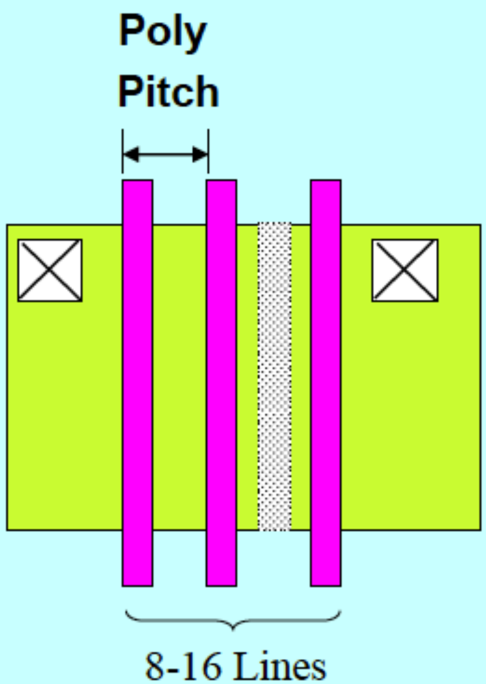
Phys. Rev. Lett. **95**, 106101 (2005)



2009 Definition of the Half Pitch – unchanged

[No single-product "node" designation; DRAM half-pitch still litho driver; however, other product technology trends may be drivers on individual TWG tables]

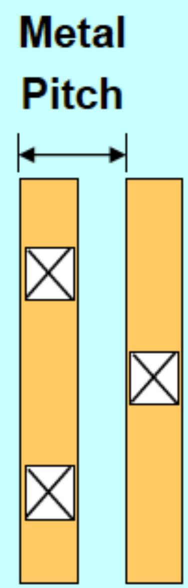
FLASH Poly Silicon ½ Pitch
= Flash Poly Pitch/2



Typical flash
Un-contacted Poly

DRAM ½ Pitch
= DRAM Metal Pitch/2

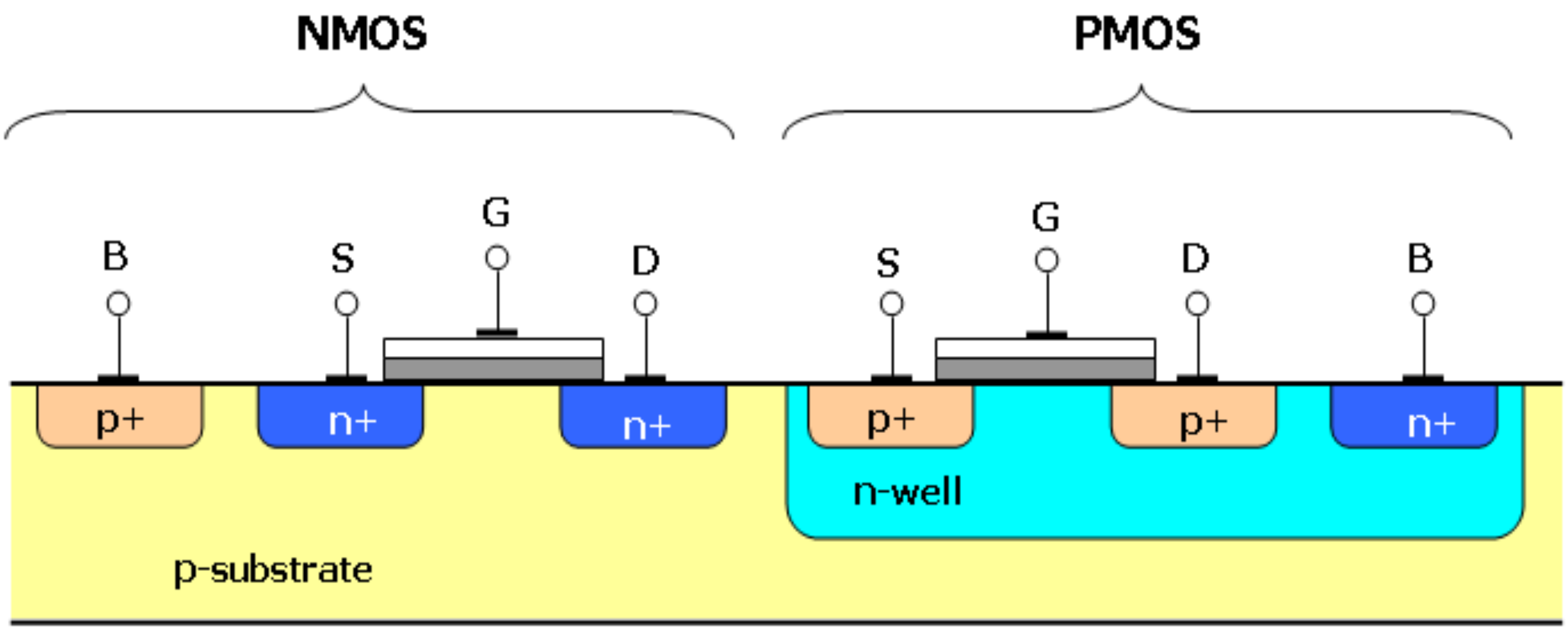
MPU/ASIC M1 ½ Pitch
= MPU/ASIC M1 Pitch/2



Typical DRAM/MPU/ASIC
Metal Bit Line

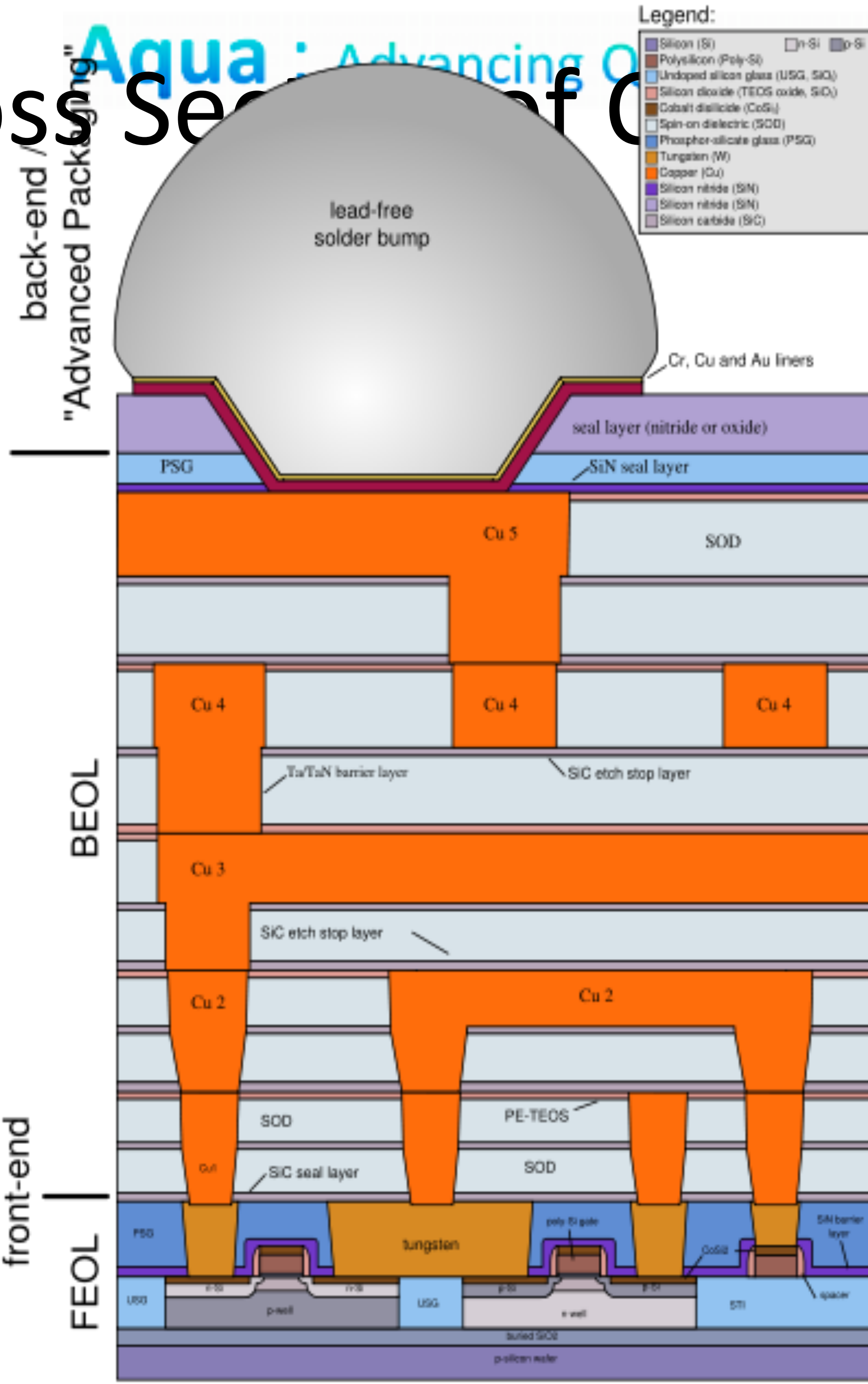
Source: 2009 ITRS - Exec. Summary Fig 1

Cross Section of CMOS



From Wikipedia

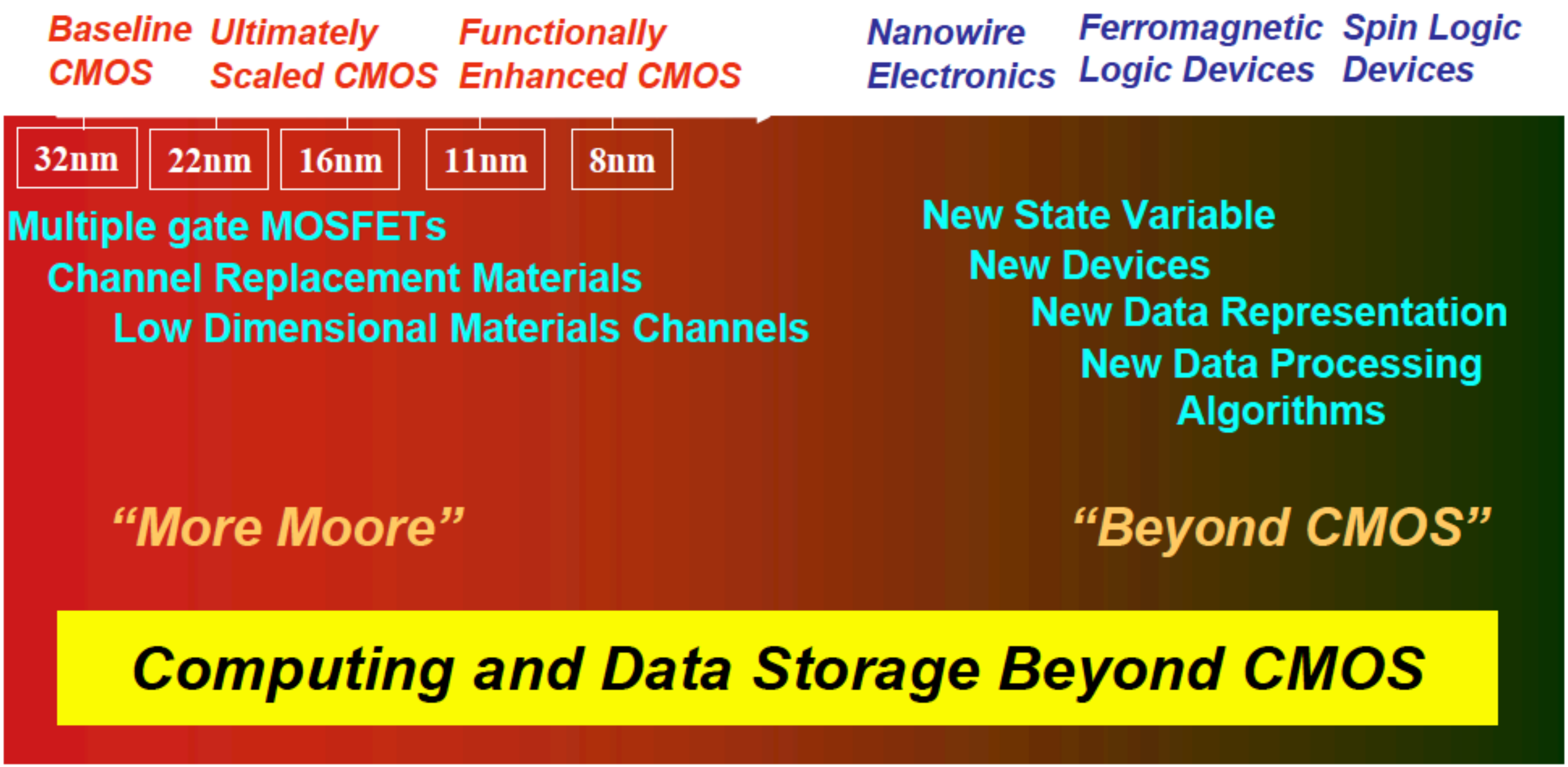
Cross Se… of C…s



From Wikipedia

[2009 – Unchanged]

2008 ITRS "Beyond CMOS" Definition Graphic

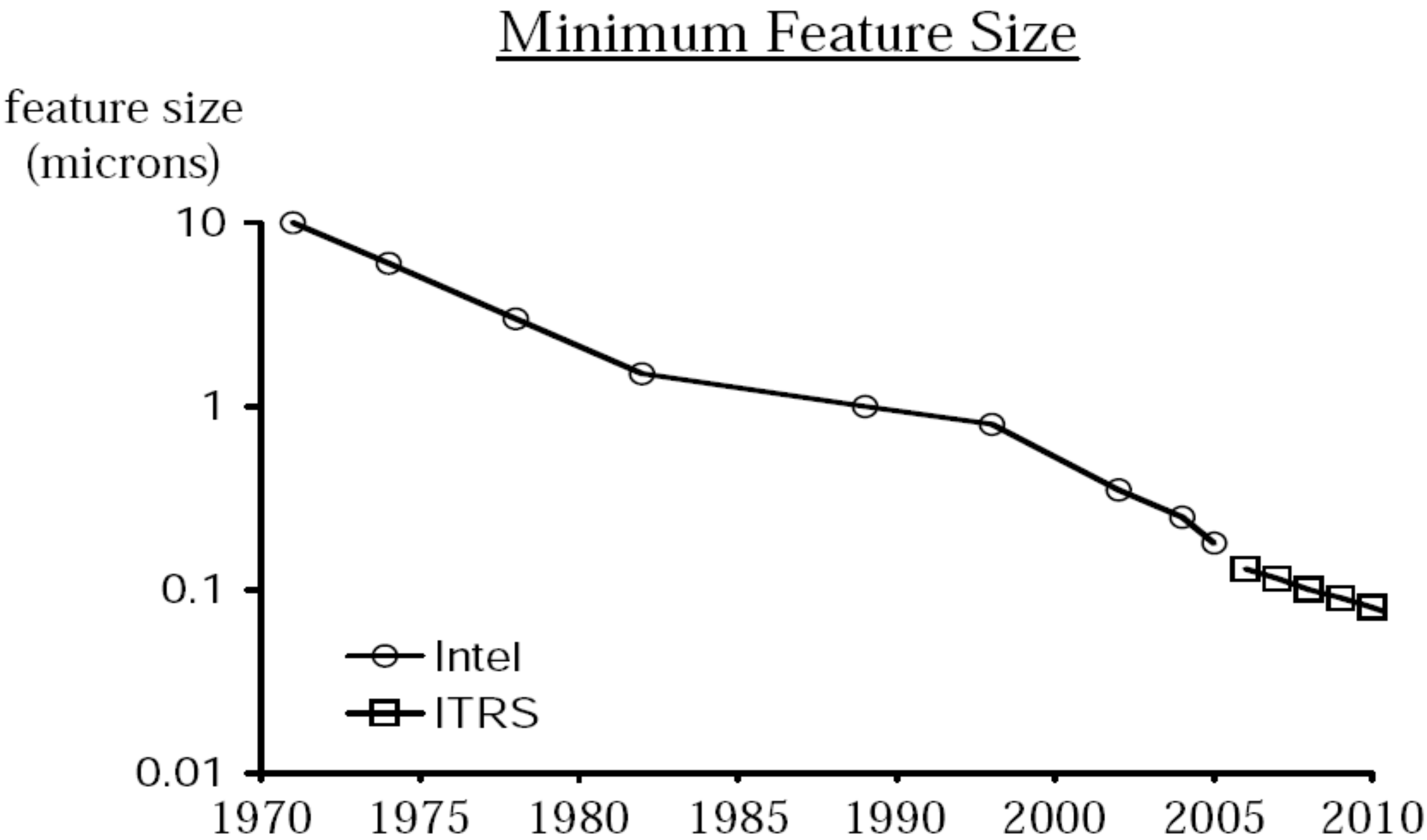


Source: Emerging Research Device Working Group

Work in Progress – Do Not Publish!

Aqua : Advancing Quantum Architecture

Limits to Moore's Law

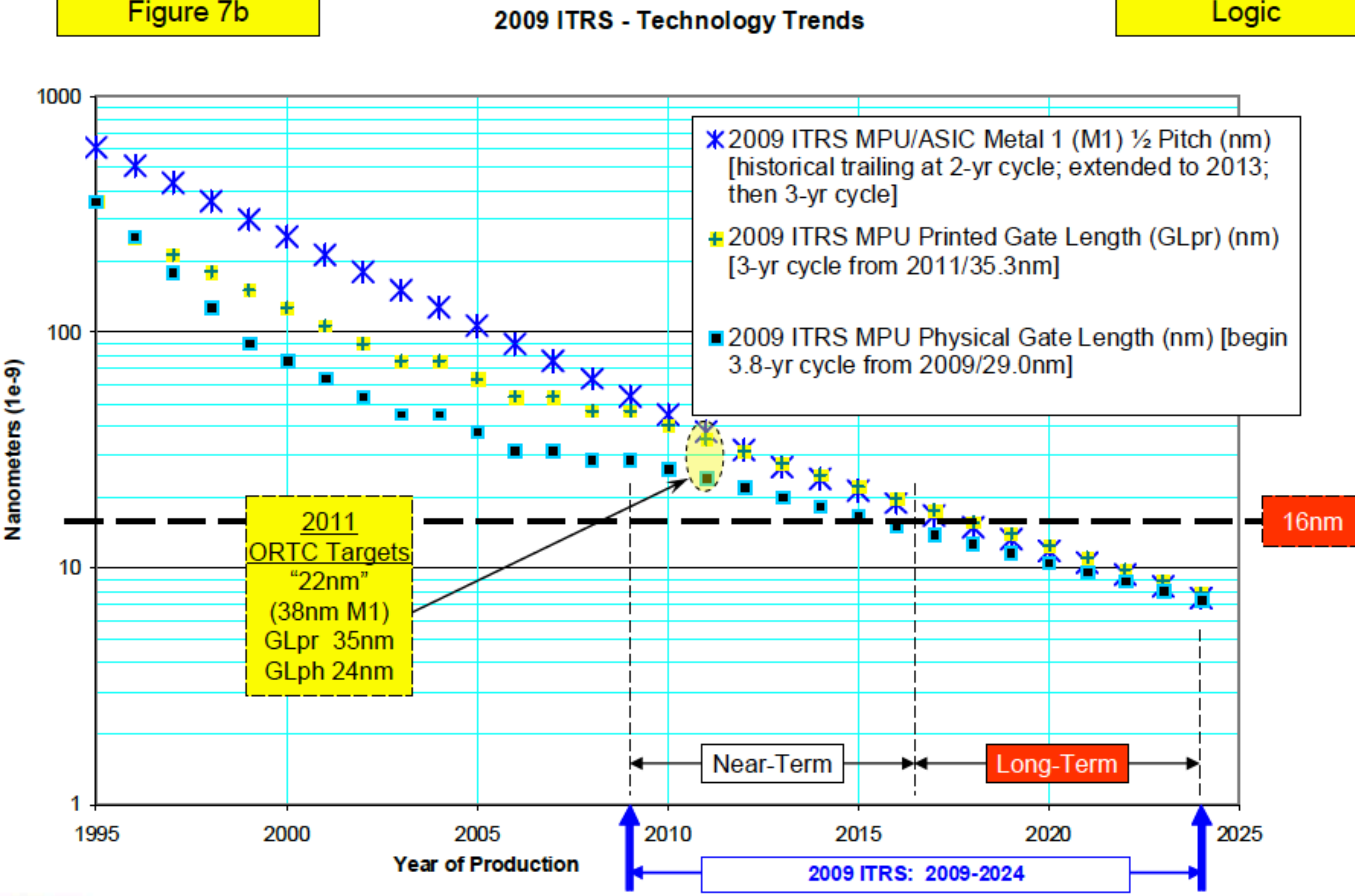


The decreasing minimum feature size of transistor components is shown for both Intel products and data reported by the International Technology Roadmap for Semiconductors (ITRS).

Including '11 snapshot Analysis

Figure 7b

Logic



Source: 2009 ITRS - Executive Summary Fig 7b

Work in Progress – Do Not Publish!

Table B ITRS Table Structure—Key Lithography-related Characteristics by Product

Near-term Years

YEAR OF PRODUCTION	2009	2010	2011	2012	2013	2014	2015	2016
Flash Uncontacted Poly Si ½ Pitch (nm)	38	32	28	25	23	20	18	15.9
DRAM stagger-contacted Metal 1 (M1) ½ Pitch (nm)	52	45	40	36	32	28	25	22.5
MPU/ASIC stagger-contacted Metal 1 (M1) ½ Pitch (nm)	54	45	38	32	27	24	21	18.9
MPU Printed Gate Length (nm)	47	41	35	31	28	25	22	19.8
MPU Physical Gate Length (nm)	[illegible]	27	24	22	20	18	17	15.3

2019: 11.3 nm
20x Si lattice
cell size

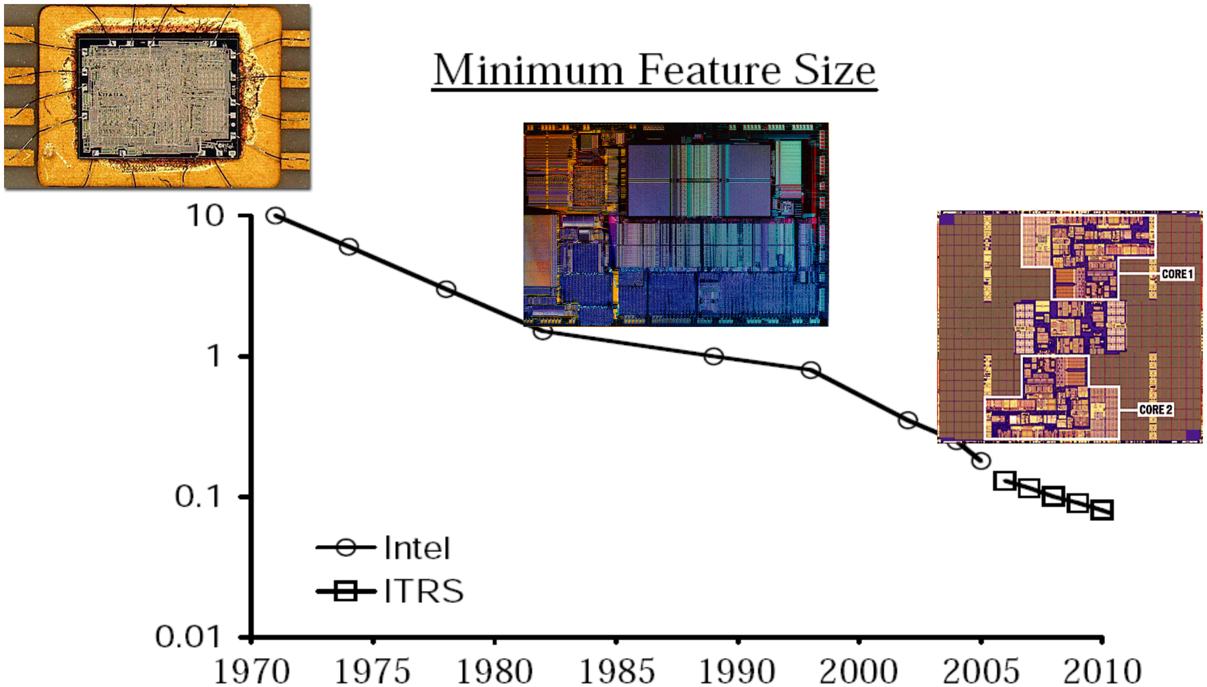
Long-term Years

YEAR OF PRODUCTION	2017	2018	2019	2020	2021	2022	2023	2024
Flash Uncontacted Poly Si ½ Pitch (nm)	14.2	12.6	11.3	10.0	8.9	8.0	7.1	6.3
DRAM stagger-contacted Metal 1 (M1) ½ Pitch (nm)	20.0	17.9	15.9	14.2	12.6	11.3	10.0	8.9
MPU/ASIC stagger-contacted Metal 1 (M1) ½ Pitch (nm)	16.9	15.0	13.4	11.9	10.6	9.5	8.4	7.5
MPU Printed Gate Length (nm)	17.7	15.7	14.0	12.5	11.1	9.9	8.8	7.9
MPU Physical Gate Length (nm)	14.0	12.8	11.7	10.7	9.7	8.9	8.1	7.4

2024: 6.3 nm
12x Si lattice
cell size

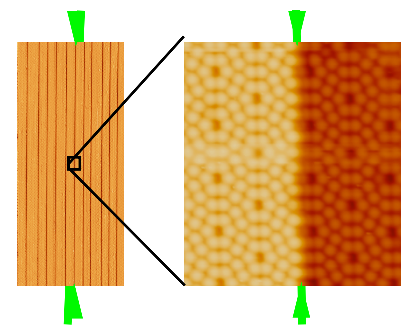
The ORTC and technology requirements tables are intended to indicate current best estimates of introduction timing for specific technology requirements. Please refer to the Glossary for detailed definitions for Year of Introduction and Year of Production.

Limits to Moore's Law



The decreasing minimum feature size of transistor components is shown for both Intel products and data reported by the International Technology Roadmap for Semiconductors (ITRS).

Atomic level
2020s?



Classical Computing's Quantum (and Atomic) Problem

- 2024 goal for flash half-pitch: 6.3 nm
- Si lattice constant: 0.54 nm
- Biggest current problem thermodynamic (hmm, reversible?)
- Doping becomes discrete phenomenon
- Transistor charging becoming quantum
- Tunnelling
- When does current become quantum?

Insurmountable Opportunity

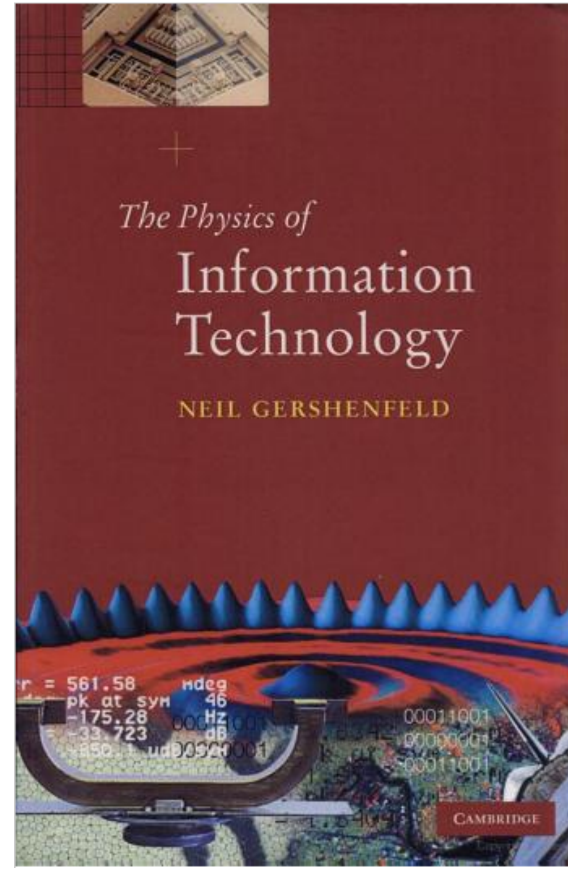
"We are confronted with insurmountable opportunities." Walt Kelly

Classical needs quantum.
Quantum needs classical.

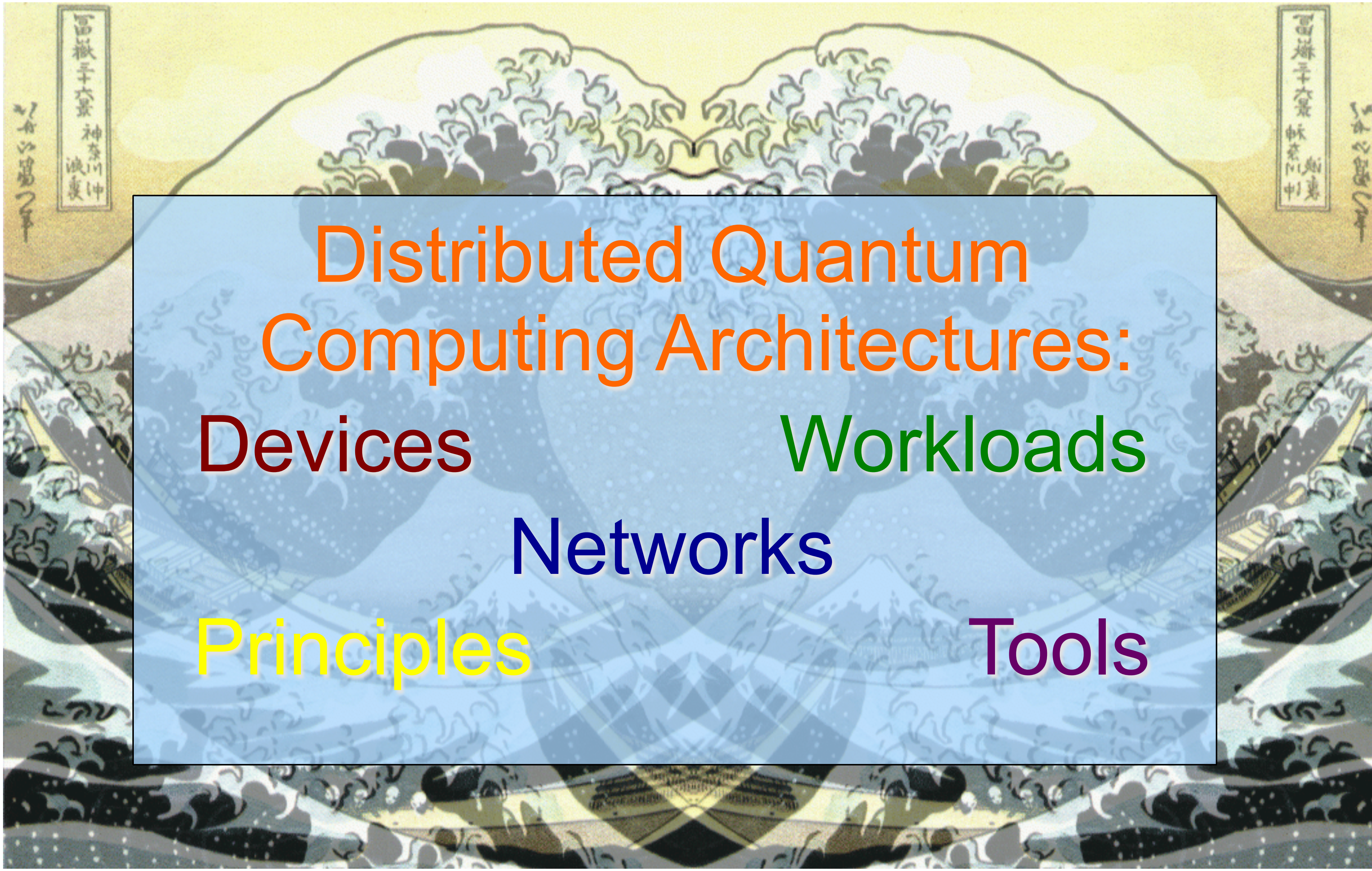
We are going to have tremendous tools
to go with our tremendous challenges.

Intermission

Recommended Books



AQUA: Advancing Quantum Architecture



Distributed Quantum Computing Architectures:

Devices

Workloads

Networks

Principles

Tools

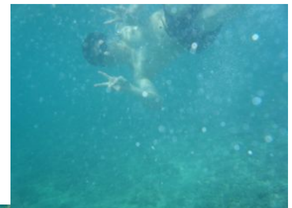
WIDE

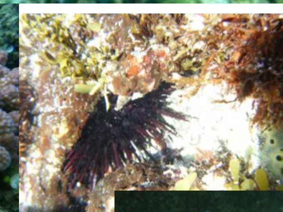
KEIO UNIVERSITY

Aqua : Advancing Quantum Architecture

Aqua & Friends

Aqua & Friends





WIDE

KEIO UNIVERSITY CALAMVS GLADIO FORTIOR 1858

Aqua : Advancing Quantum Architecture

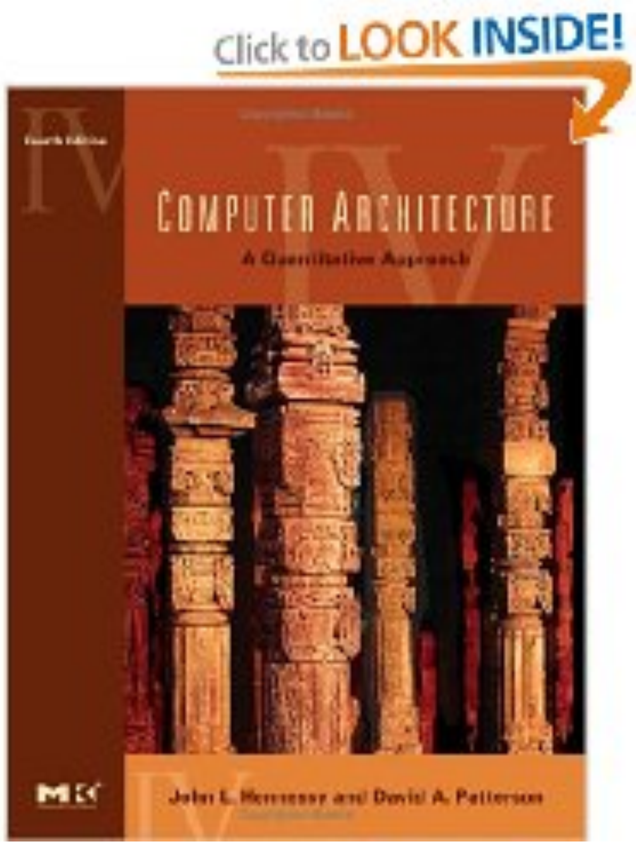
Aqua : Advancing Quantum Architecture

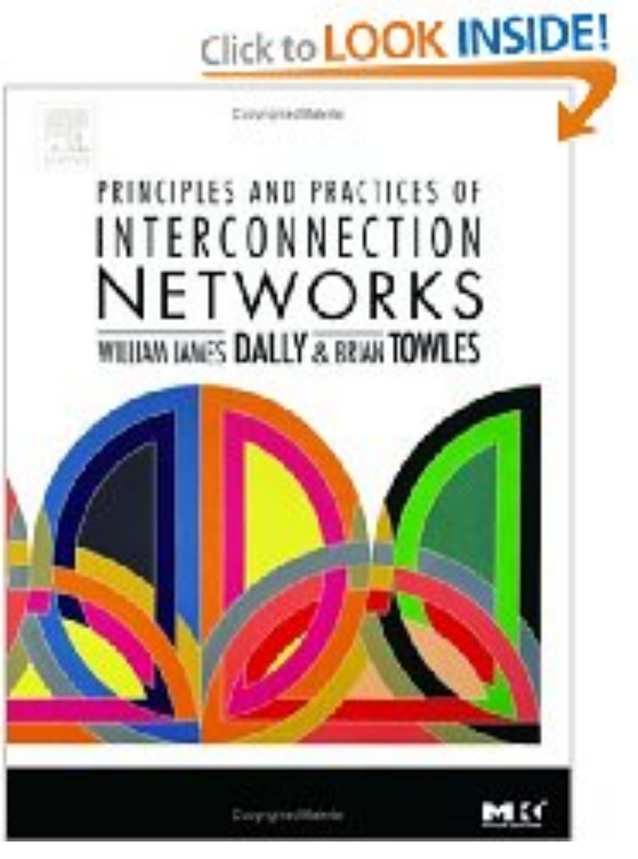
AQUA: Large-Scale Distributed Quantum Computing

- Surface code architectures & workloads:
  - Arithmetic: Byung-Soo Choi (U. Seoul), Agung Trisetyarso
  - Compilers & efficient data movement: Byung-Soo, Clare Horsman, Kaori Ishizaki (B4), Pham Tien Trung (B4)
  - Defects in Surface Codes: Shota Nagayama (M2)
- Networks:
  - Quantum Dijkstra (path selection for repeater networks): Takahiko Satoh (M2, Todai Imai-ken)
  - Multiplexing in repeater networks (& new network simulator): Luciano Aparicio (M2, Todai Esaki-ken)
  - IPsec with QKD (keying the Internet): Shota
- Architectures for Q simulators: Clare

Okay, back to architecture (classical first, then quantum)

Graduate Computer Architecture in Six Slides





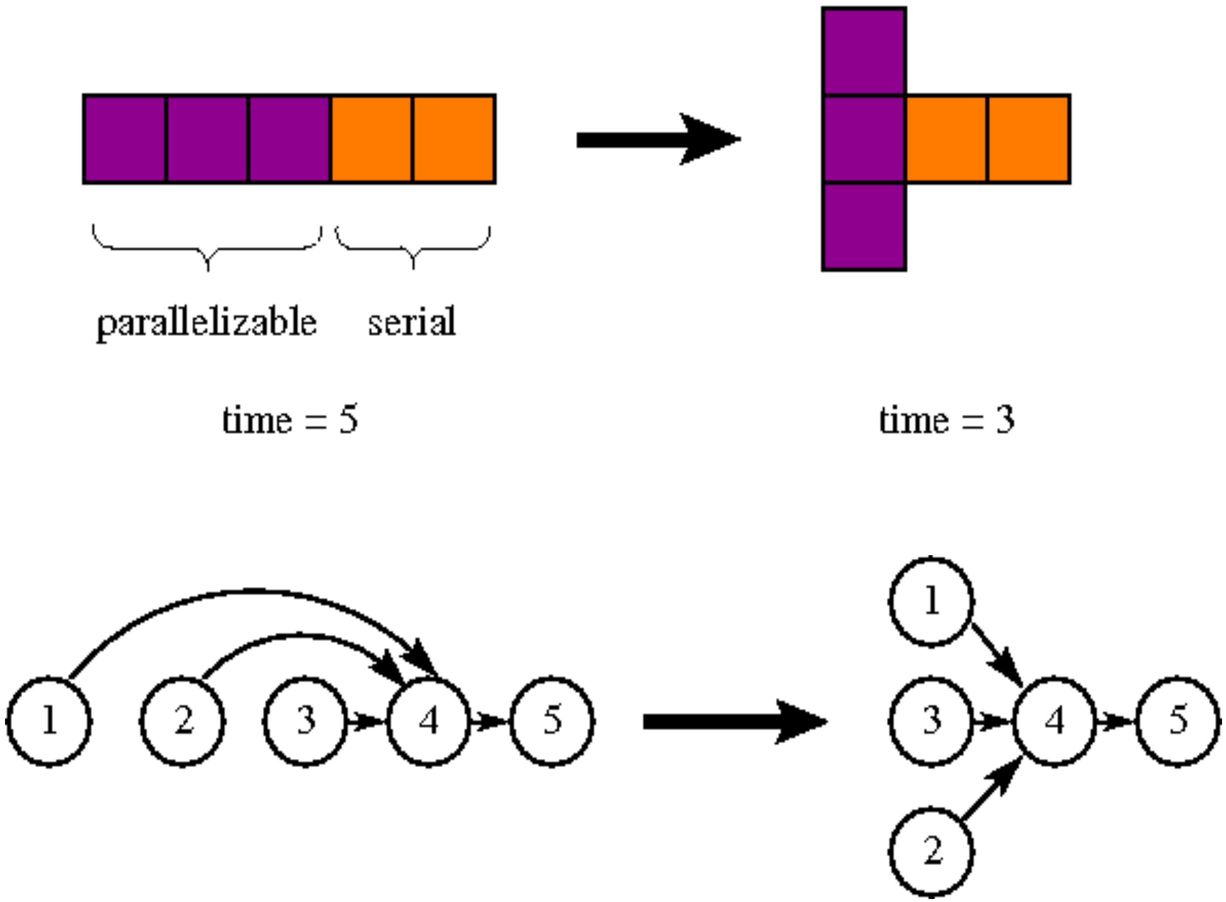
So What Does a Computer *Do*?

- Process data
- Move data
- Store data
- Manage data

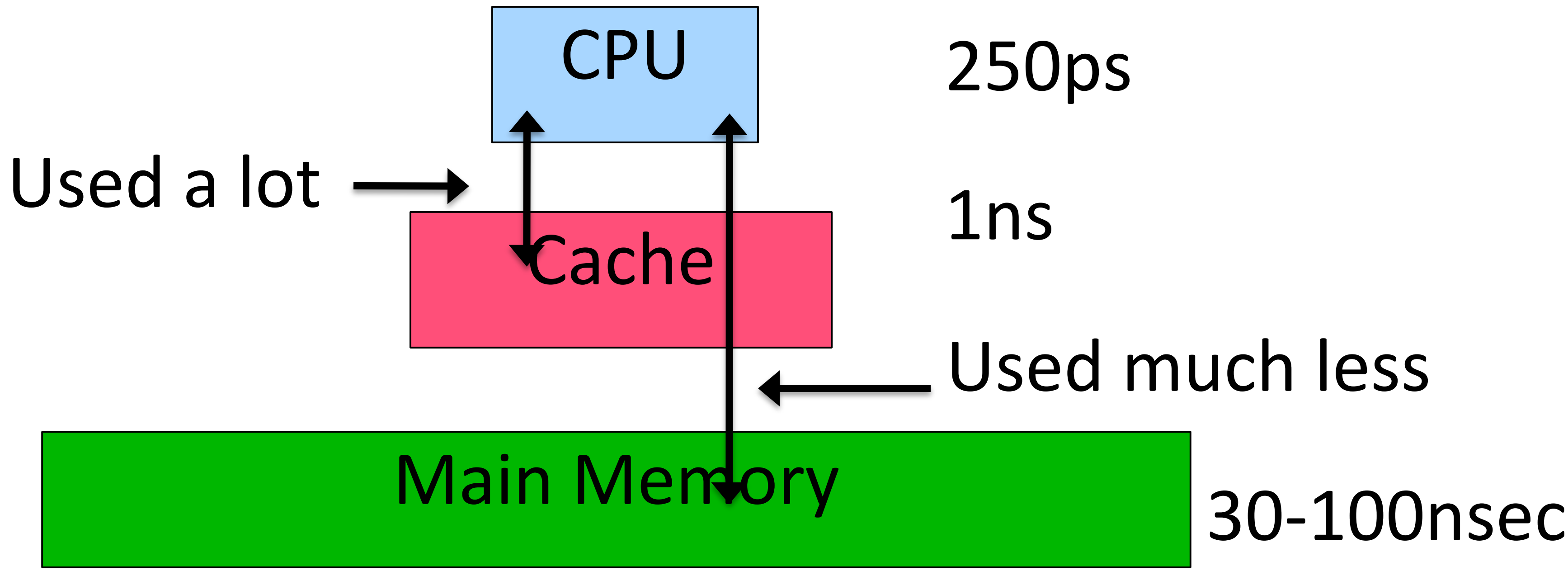
Hennessy & Patterson's Five Principles of Computer Design

- Take advantage of parallelism
- Amdahl's Law
- Principle of locality
- Focus on the common case
- The Processor Performance Equation

Amdahl's Law



Principle of Locality



A *cache memory* is *smaller* and *faster* (both in *latency* and *bandwidth*) than some "main" memory; it takes advantage of *spatial* and *temporal locality* in program behavior.

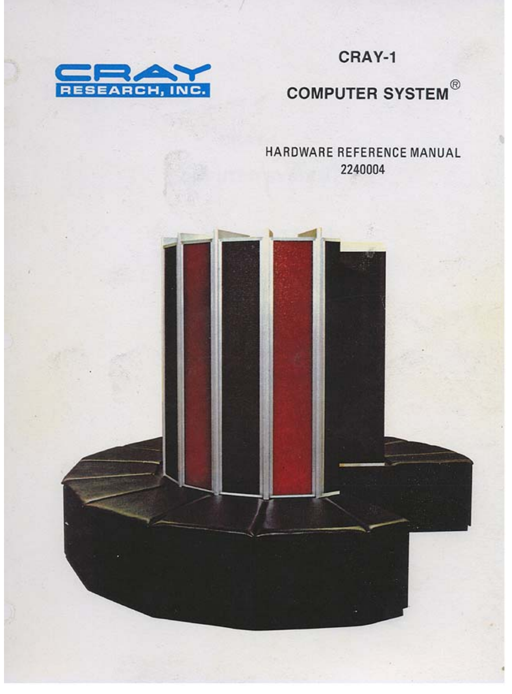
Processor Performance Equation

$$\text{CPU time} = \frac{\textit{seconds}}{\textit{program}} = \frac{\textit{Instructions}}{\textit{program}} \times \frac{\textit{Clock cycles}}{\textit{Instruction}} \times \frac{\textit{Seconds}}{\textit{Clock cycle}}$$

How much work is the problem for a machine line this?

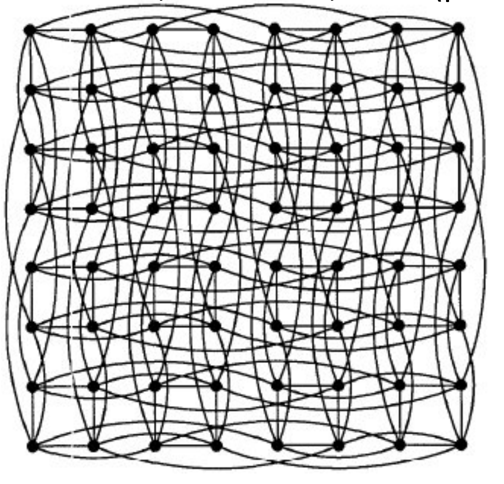
How much work per second?

Two Paths to Scalability



Cray 1, 80MFLOPS, 8MB RAM, $9M, 1976

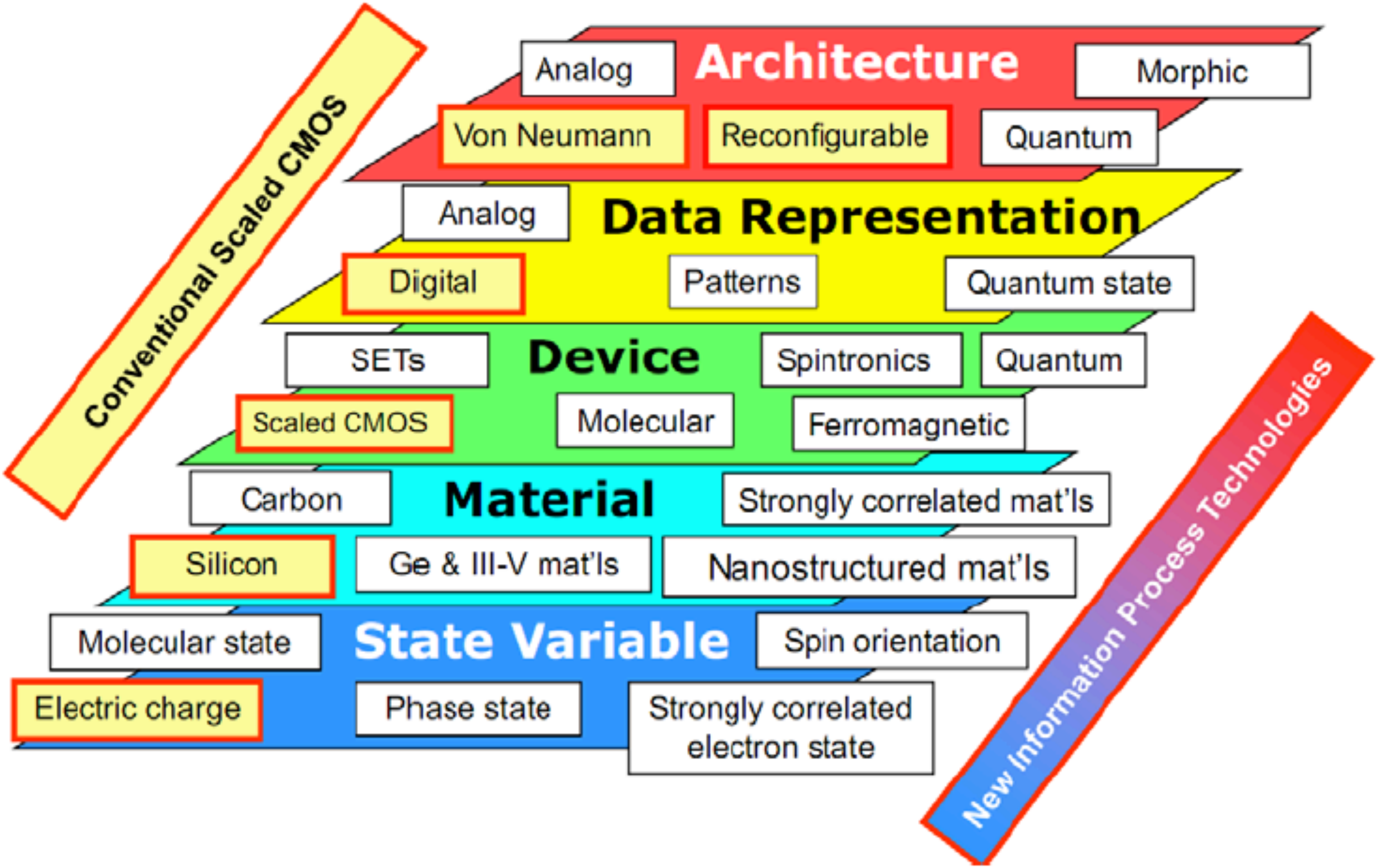
Caltech Cosmic Cube, 64 processors (8086/7)
3MFLOPS, 8MB RAM, 1982 (prototype)



Two choices:
Make it bigger, or figure out how to connect more than one smaller unit hopefully achieving both *speed* and *storage capacity* increases

Now Quantum:
How to design a quantum computer

A Taxonomy for Nano Information Processing Technologies



From ITRS Emerging Research Devices, 2009

Approaching Architecture

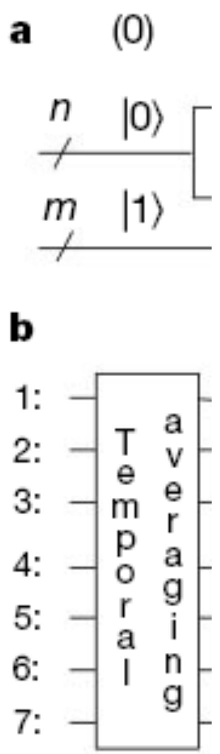
- Understanding workload
- Moving data
  - at application level
  - for QEC
- Managing resources
- Preparing for errors
  - Run-time changes to state
  - Non-functional components

Application Workload

- Shor's algorithm
  (*almost* beaten to death)
- QKD
  (definitely beaten to death)
- Quantum simulation
  (hmmm...)

Circuit for Shor's Algorithm

Time



Space (qubits)

Initialize

Compute

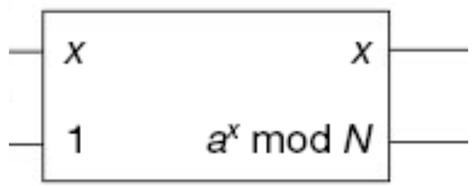
Measure

Vandersypen et al., Nature 414 (2001)

Circuit for Shor's Algorithm

Time

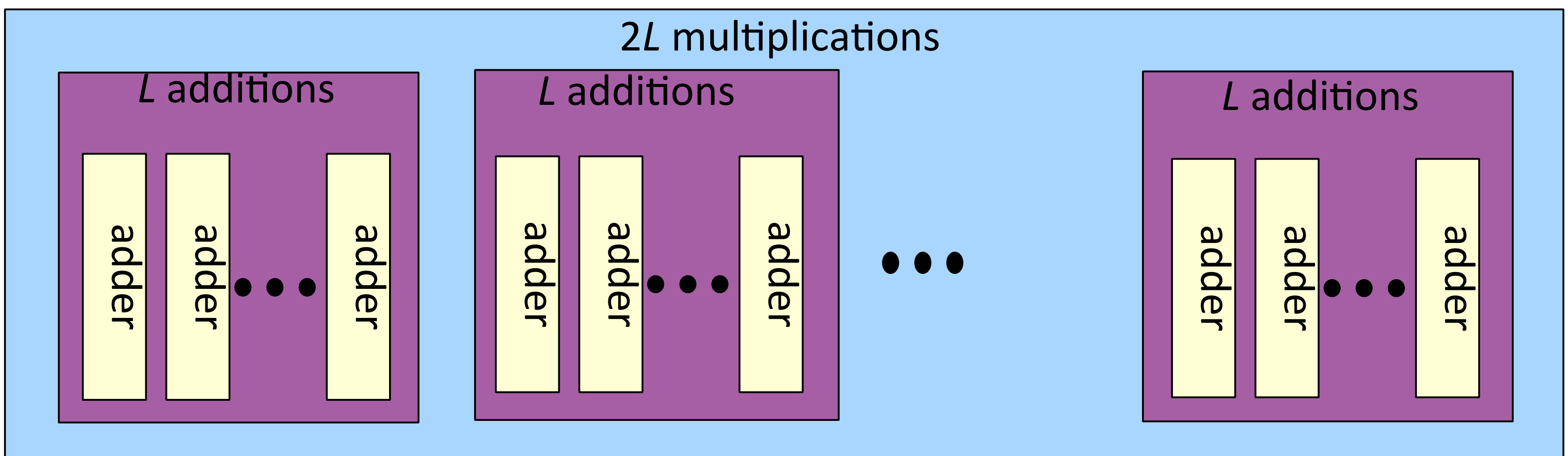
(2)



modular exponentiation

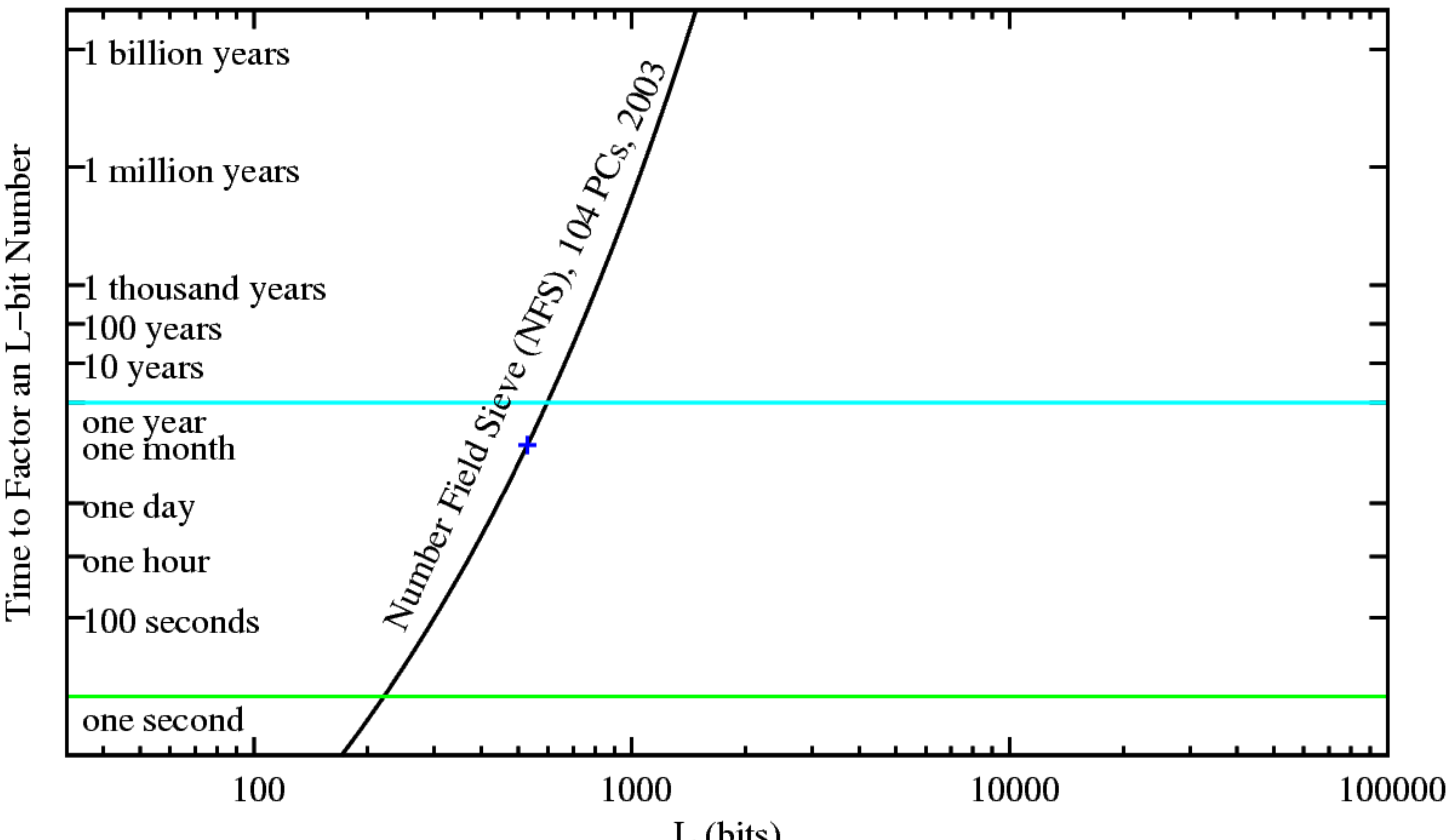
$2L$ multiplications
x L additions
x $O(L)$ ops/addition

$O(L^3)$ operations



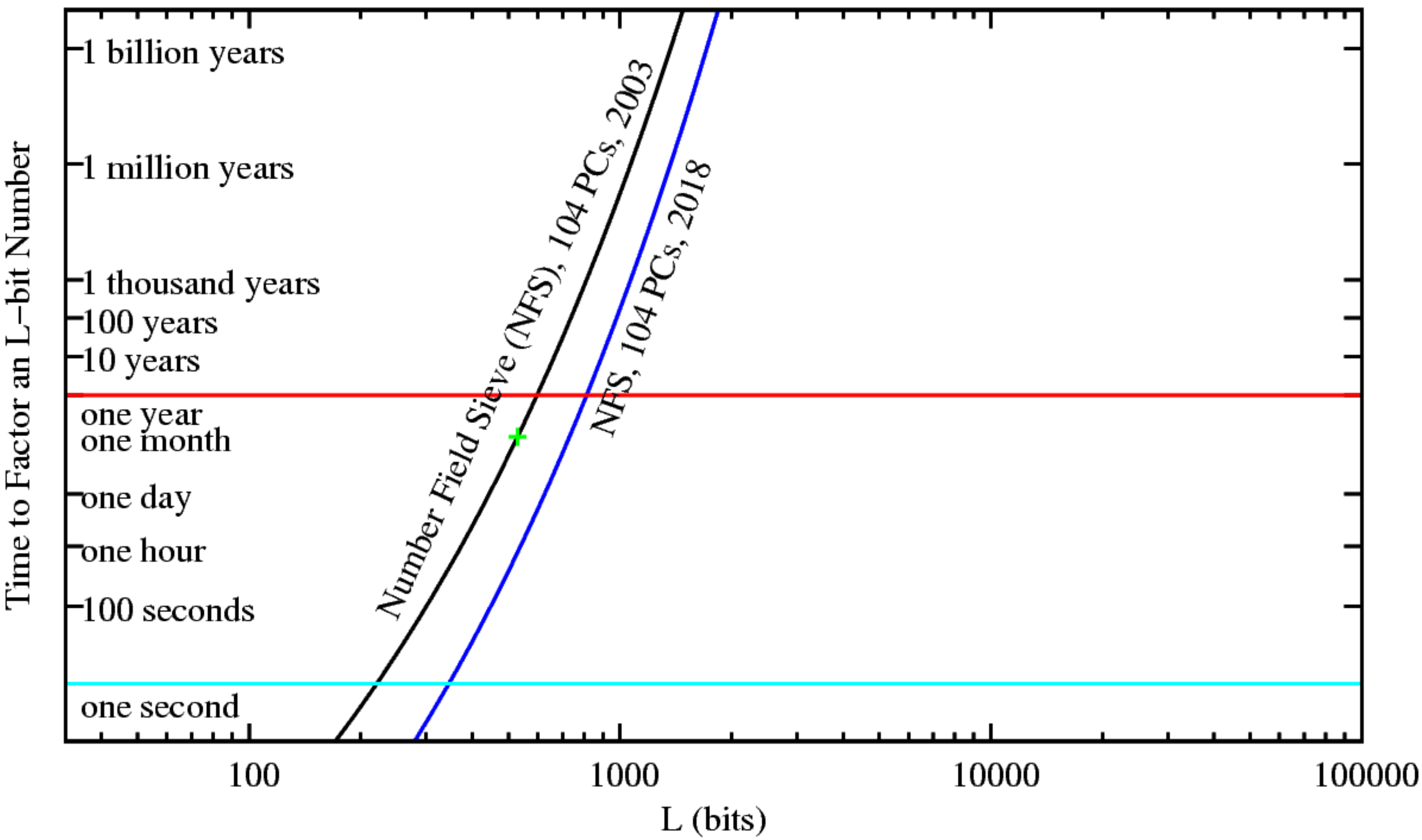
Vandersypen et al., Nature 414 (2001)

Factoring Large Numbers

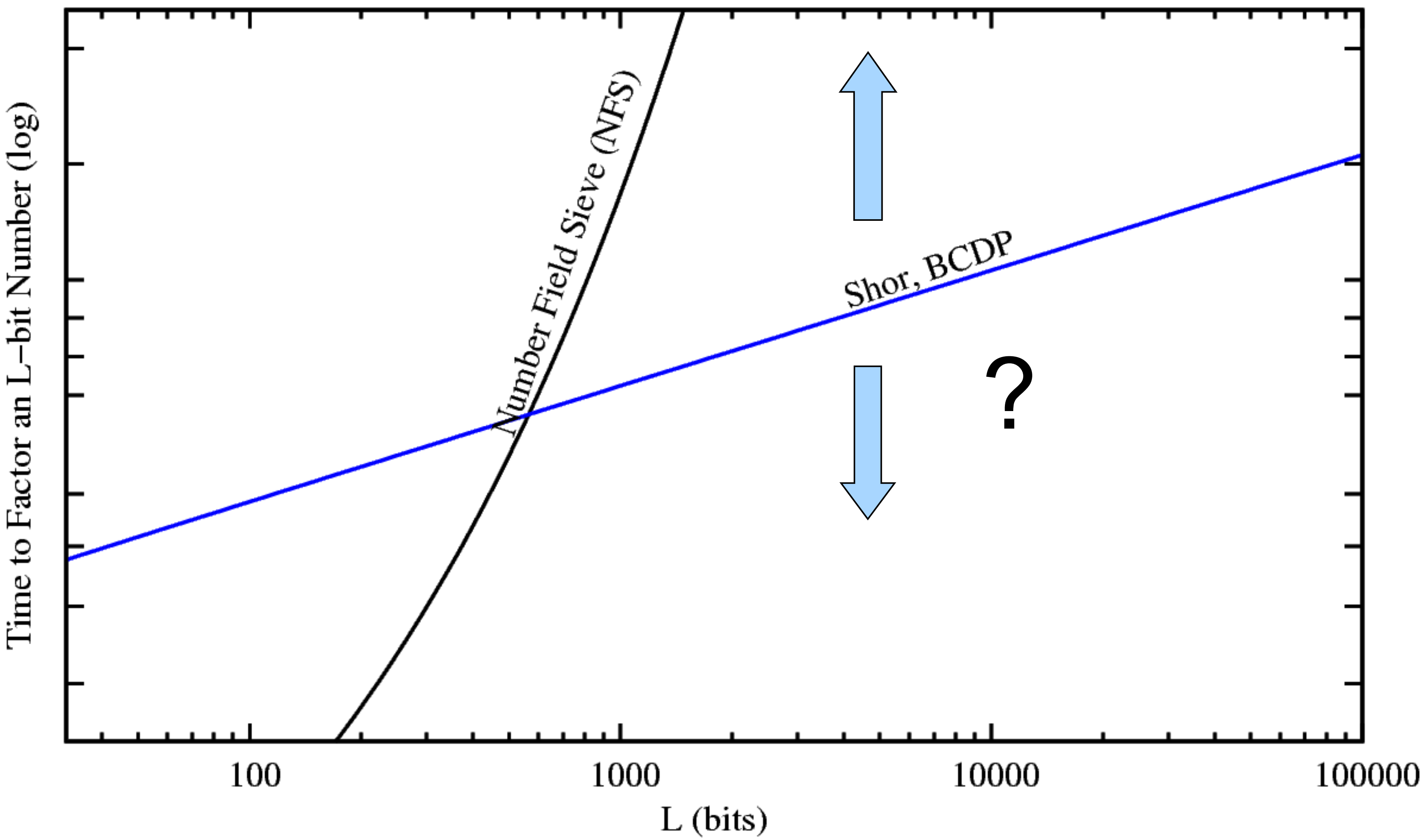


rdv, TDL, KMI, quant-ph/0507023

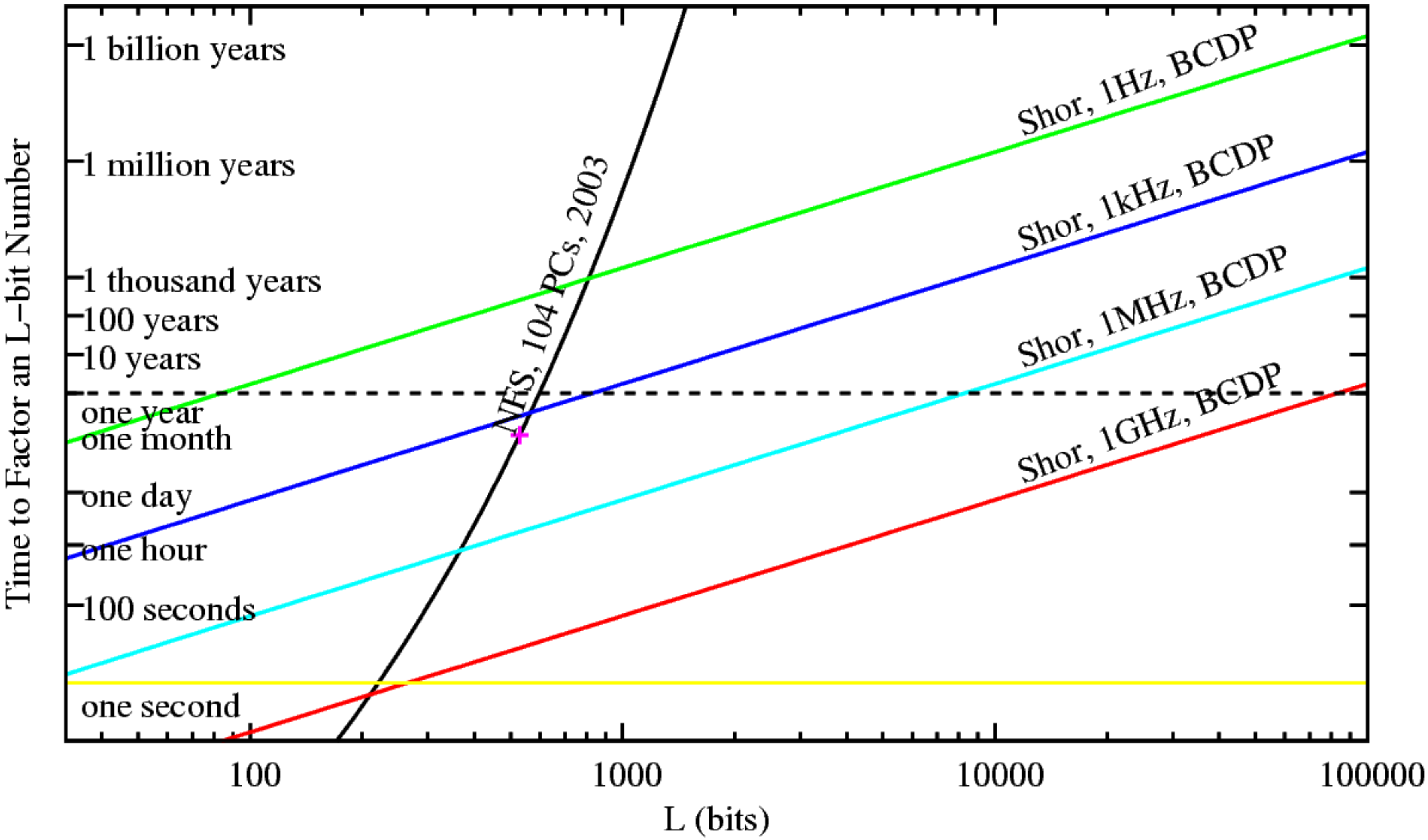
Factoring Large Numbers



Factoring Larger Numbers



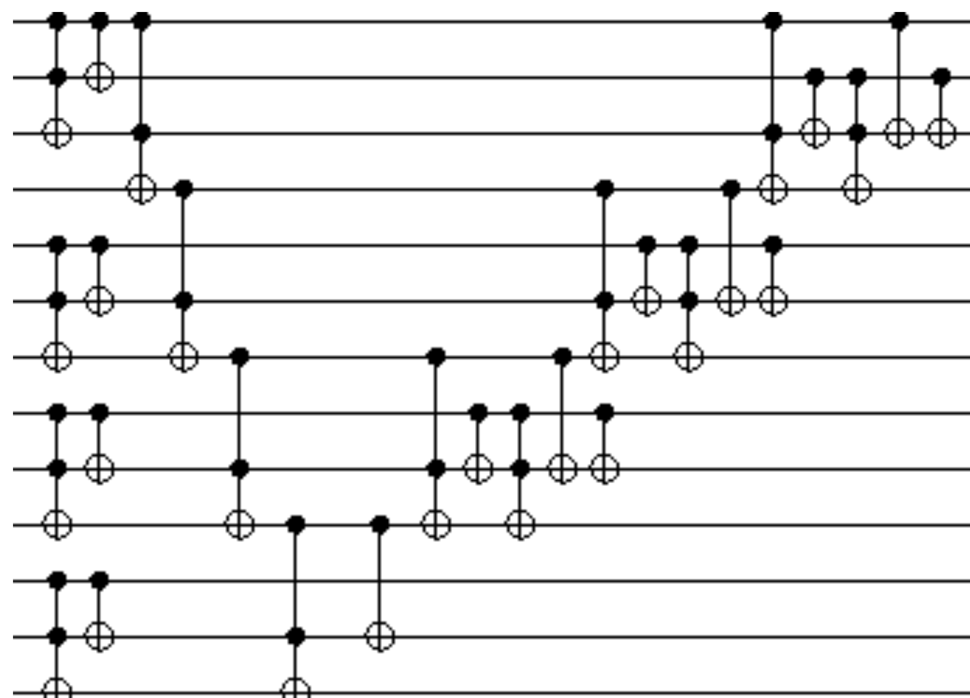
Factoring Larger Numbers



Moving Data

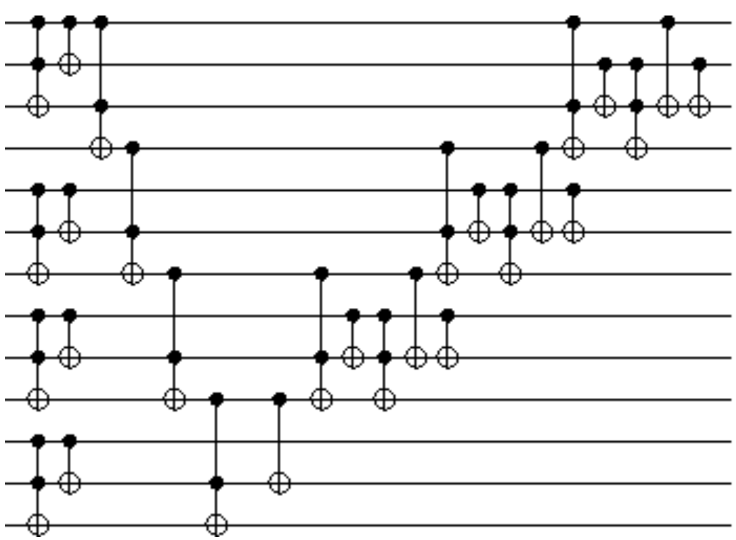
Quantum Addition

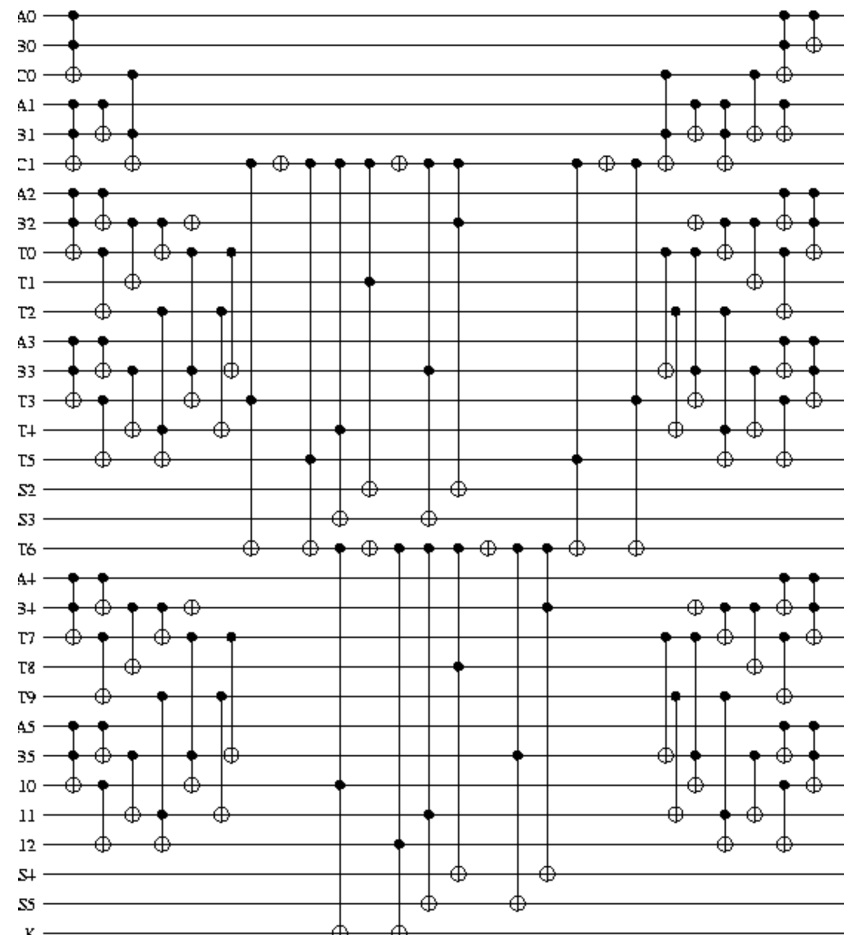
- Remember, done at *logical* level
- Classical-derived algorithms dependent on Toffoli gates
- Always $O(L)$ total gates
- Can be $O(\log L)$ time *if* long-distance gates available



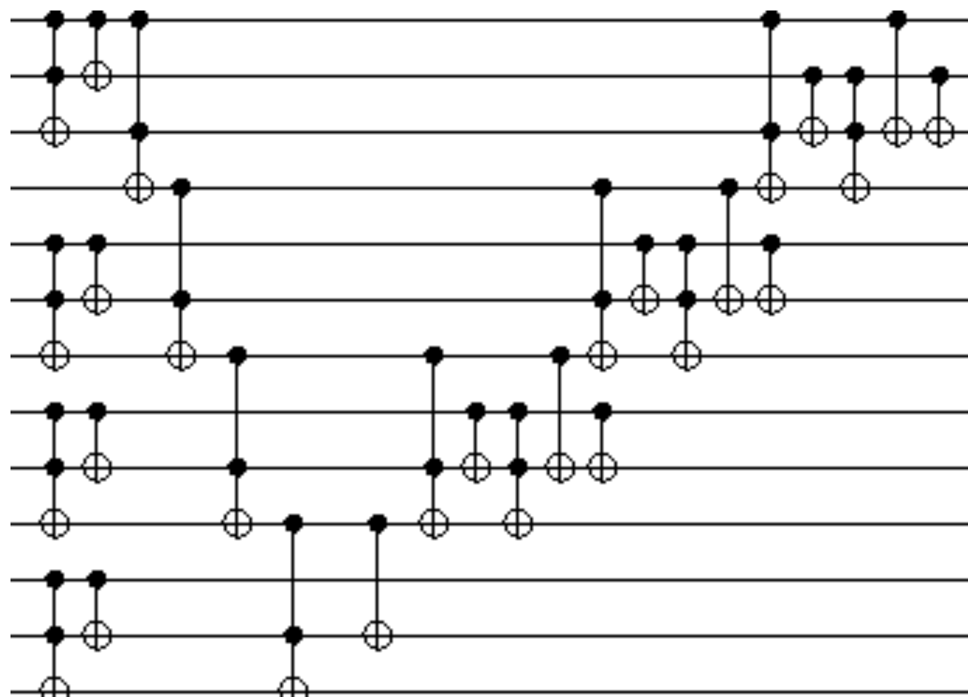
Quantum Addition

- Vedral et al., VBE ripple adder *PRA*, 1996
  $O(L)$ depth using nearest neighbor only
- Draper et al., *QIC*, 2006
  $O(\log L)$ depth
  carry-lookahead
  w/ long-distance gates
- See rdv & KMI, *PRA*

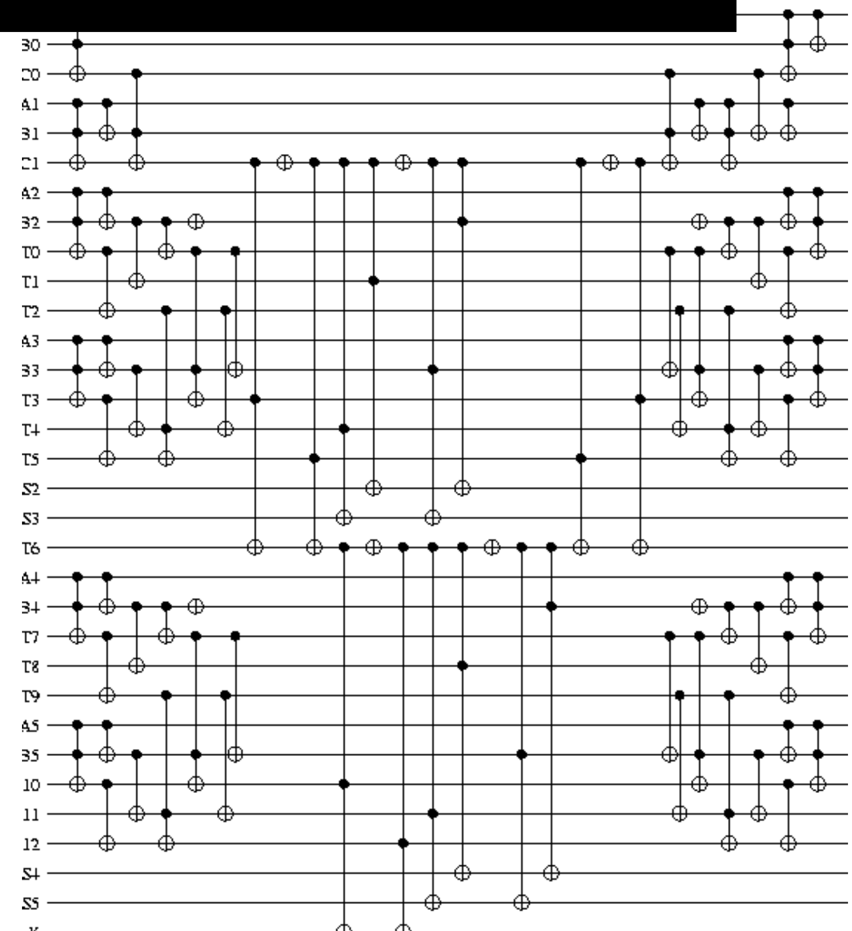




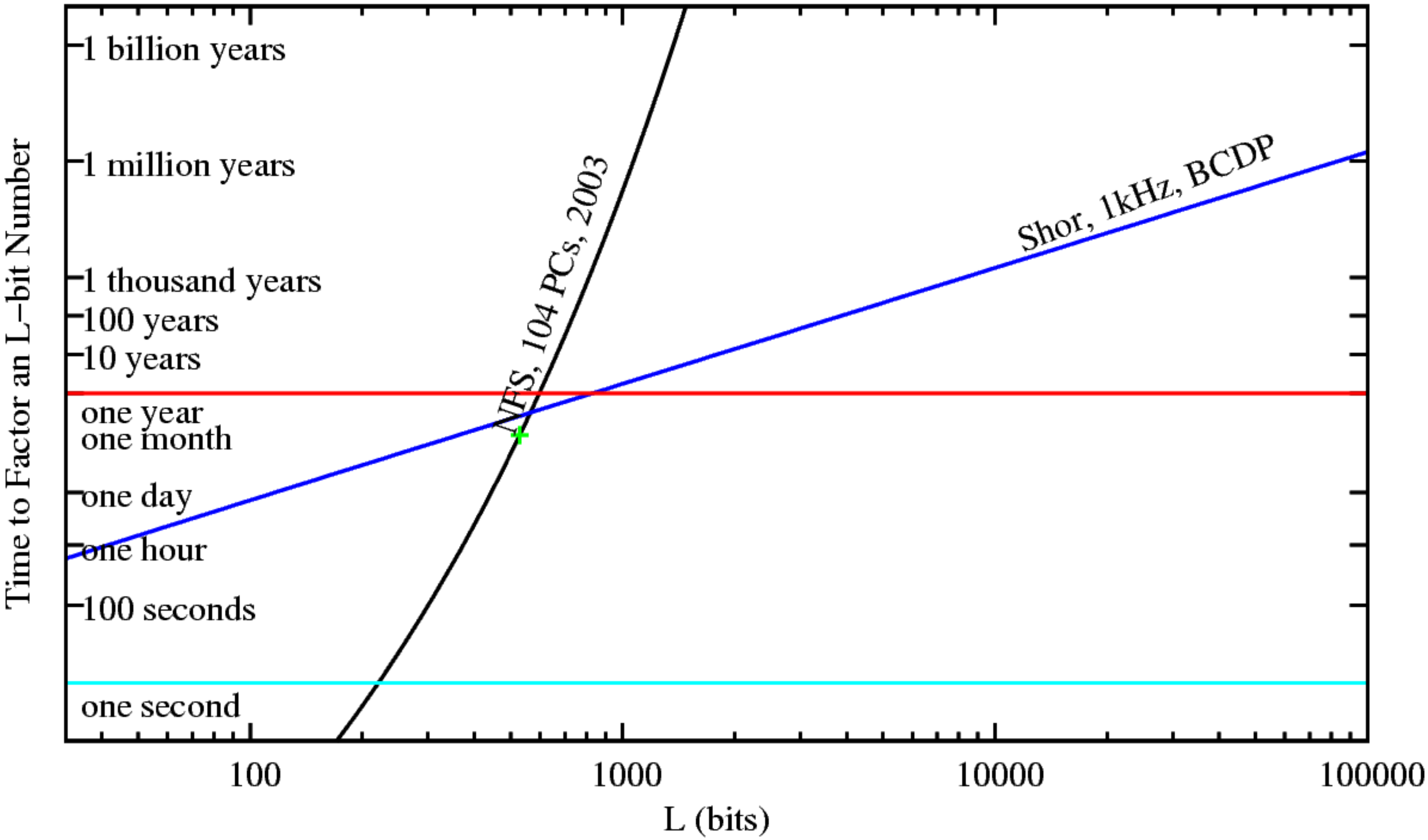
Quantum Addition



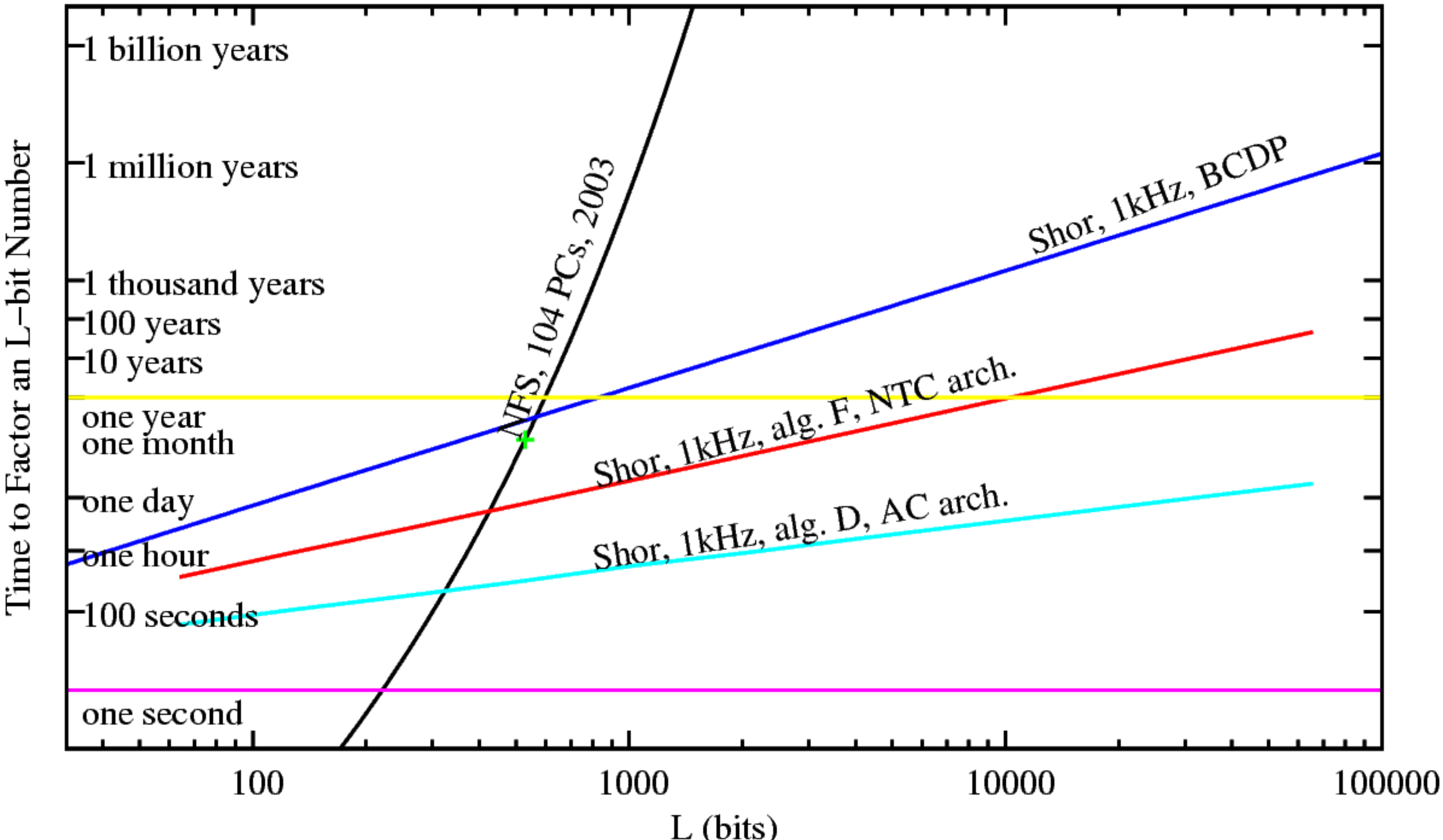
Quantum Addition



Factoring Larger Numbers



Factoring Larger Numbers



rdv, TDL, KMI, quant-ph/0507023

Impact of Connectivity on QEC

- Trade-off between interconnect and threshold
- Thresholds
  - unlimited range, unlimited qubits: ~ 10^{-2}

    Knill, quant-ph/0410199
  - unlimited range, many qubits: ~ 10^{-3}–10^{-4}

    Steane, Phys. Rev. A 68, 042322 (2003)
  - 2D lattice, nearest neighbor (CSS): ~ 10^{-5}

    Svore, QIC 7, 297 (2007)
  - bilinear nearest neighbor: ~ 10^{-6}

    Stephens, QIC 8, 330 (2008)
  - linear nearest neighbor: ~ 10^{-8}

    Stephens, in preparation
  - surface code, 2D lattice nearest neighbor: ~ 1.4%

    Wang & Fowler, PRA 83, arXiv:1009.3686v1 [quant-ph]

Heterogeneous Interconnects

- Real systems will (almost certainly) have heterogeneous interconnects
- Individual device capacity limited
  - e.g., cavities 50 microns, chip size maybe 1 sq.cm.
  - must couple off-chip
- Even within device, homogeneity unlikely
  - work around classical control structures
- Multi-level error management necessary

Purification, followed by QEC

WIDE

KEIO UNIVERSITY CALAMVS GLADIO FORTIOR 1858

Aqua : Advancing Quantum Architecture

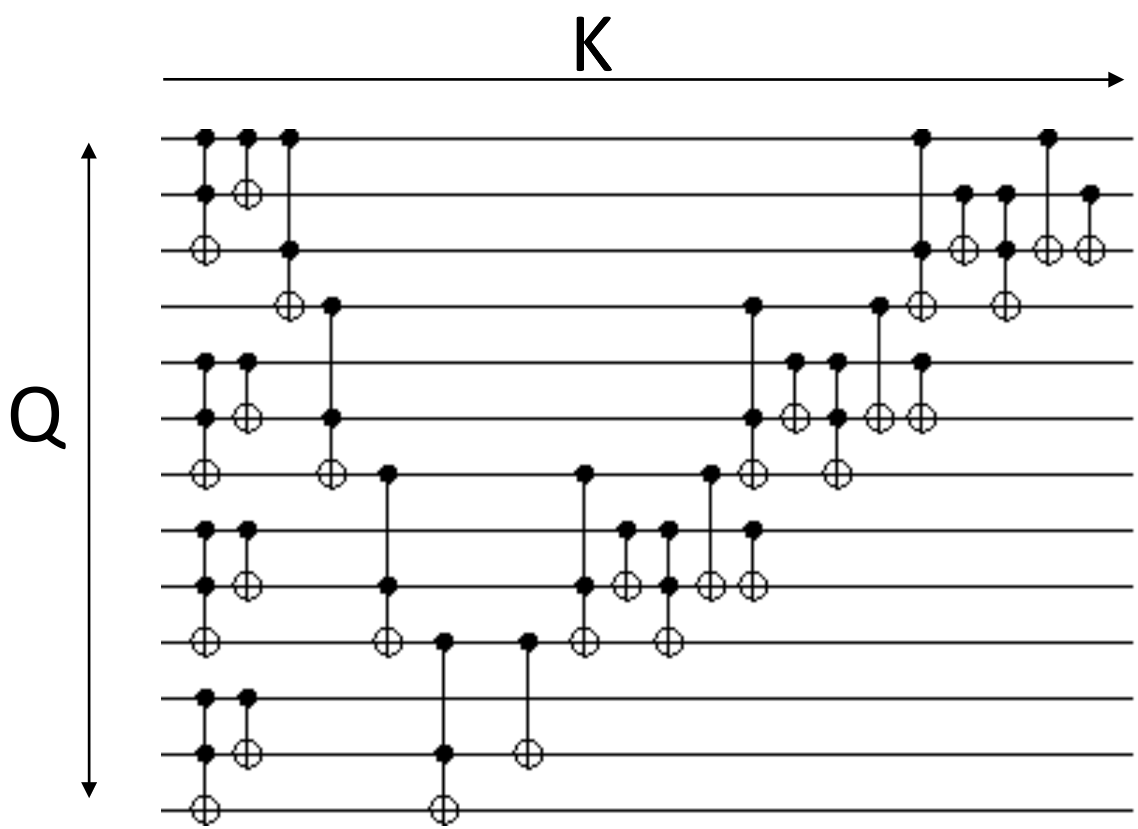
Phy & Log Connectivity

- Systems have both physical and logical topology
- In surface code,
  - – logical surface is fine-grained
  - – "Wires" move data quickly through machine, but still consume resources
  - – Performance/resource impact still poorly understood, but classical techniques valuable
- Impact of connectivity felt both at both micro and macro levels

KQ and What It Means For You

Steane's *KQ* Analysis for QEC

- Q is number of logical qubits
- K is application circuit depth
- Integral of qubits in use



KQ for Modular Exponentiation

- For n = 1024, varies from 2.4E11 to 2E14, depending on algorithm & system

algorithm	KQ
cVBE	$2n \times n \times 5 \times 3n \times 7n = 210n^4$
algo. **D**	$2l \times n \times 2 \times 4\log_2 n \times 5n \approx 40n^3 \log_2 n$
algo. **E**	$2l \times n \times 2 \times 4\log_2 n \times 3n \approx 24n^3 \log_2 n$
algo. **F**	$2l \times n \times 2 \times 2n \times 3n \approx 6n^4$
algo. **G**	$2l \times n \times 3 \times 2n \times 6n \approx 18n^4$

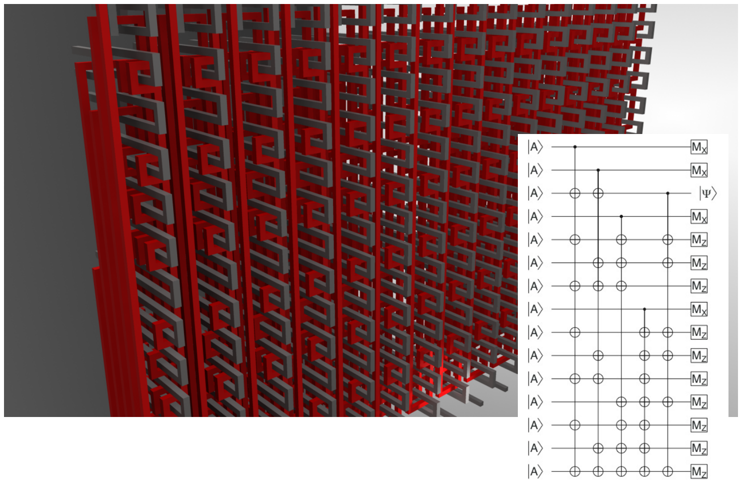
rdv, Ph.D. thesis, quant-ph/0607065

Using *KQ*

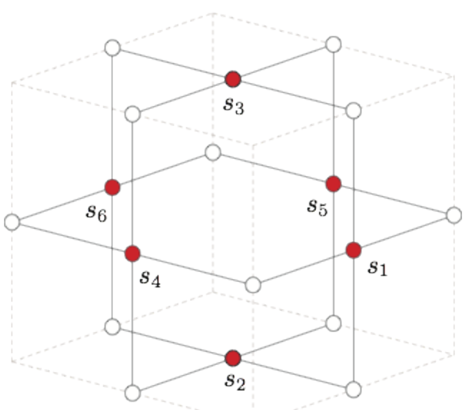
- Need Plog << 1/*KQ*
- Can engineer back to QEC requirements
  - For CSS codes, which codes & how many levels?
  - For surface code, separation distance & hole diameter?
- Steane, *PRA* 68, 042322 (2003).

Putting it All Together: Some Quantum Computer Architectures

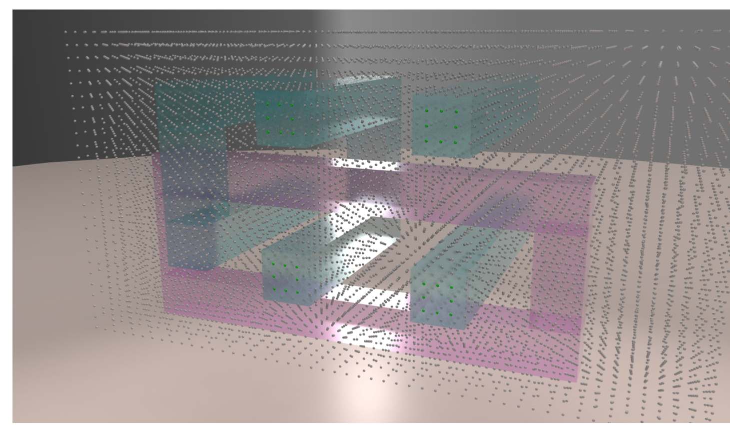
NII's optical 3-D surface code computer



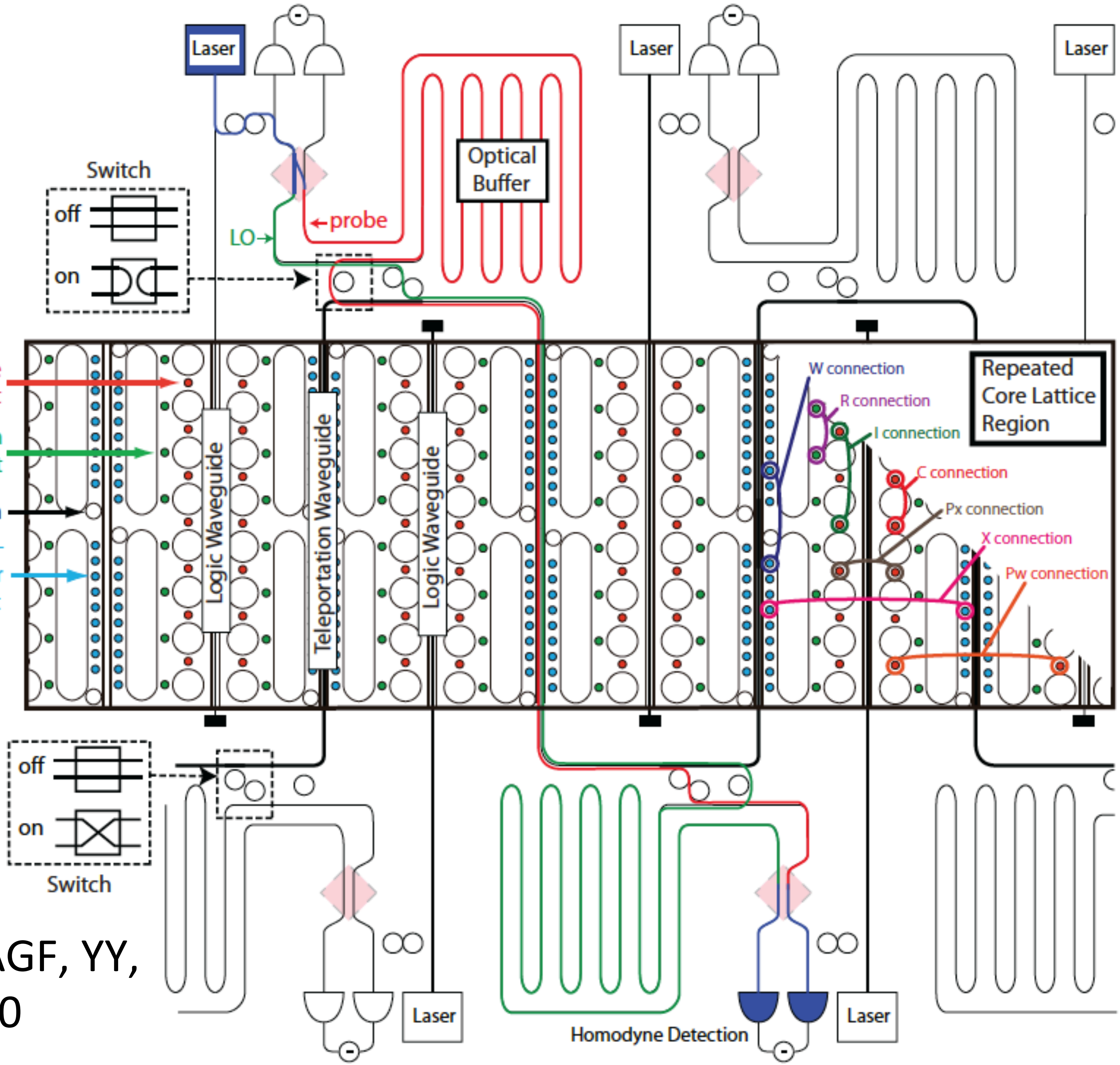
- Large fault-tolerant threshold, approximately 1%.
- Naturally nearest neighbour geometry
- Currently the standard for ALL modern quantum computing architectures.



- The Scheme requires a large 3D cluster state constructed from this unit cell element of 13 qubits.

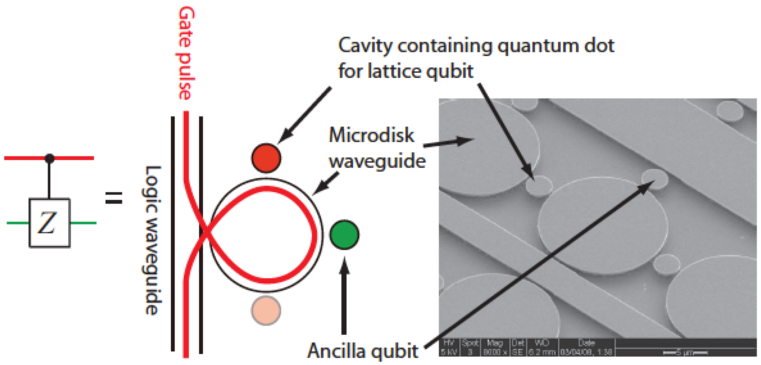


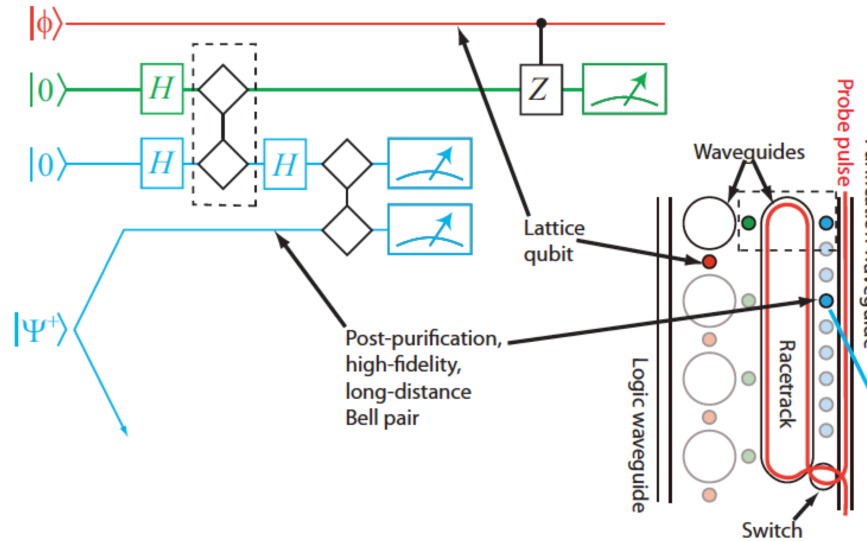
- Logic operations are performed via topological braiding, illustrated is an error protected CNOT gate.

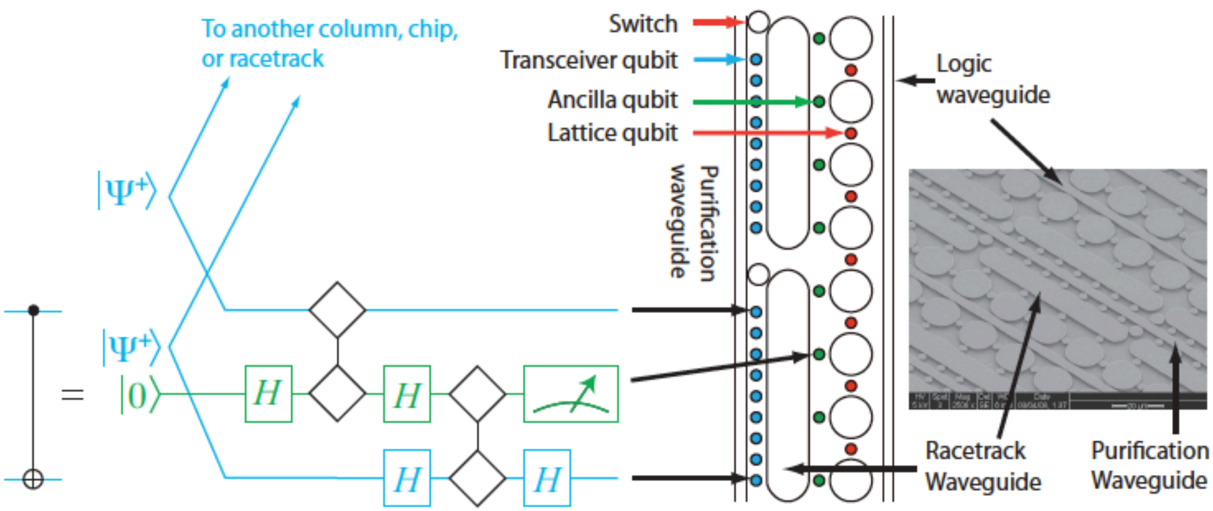


rdv, TDL, AGF, YY,
IJQI 8, 2010

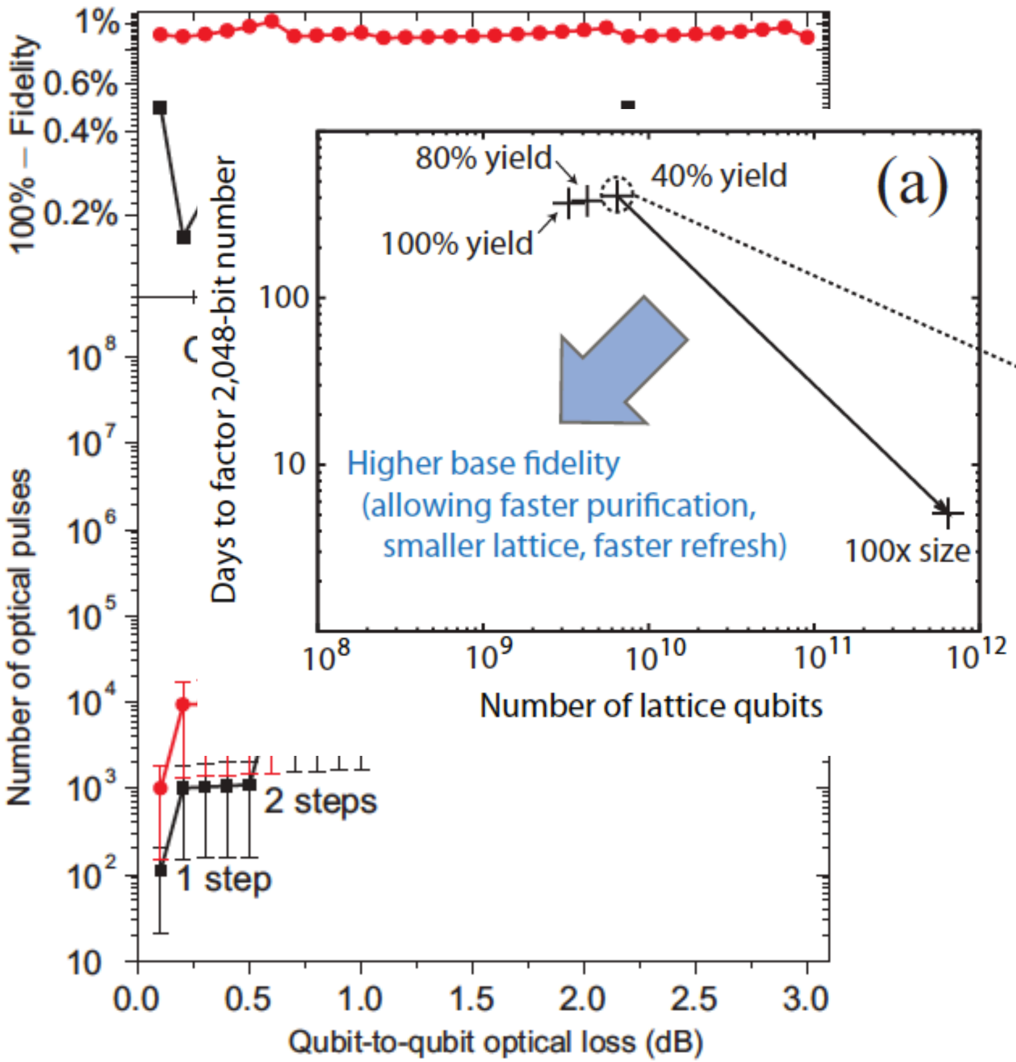
Gate Types

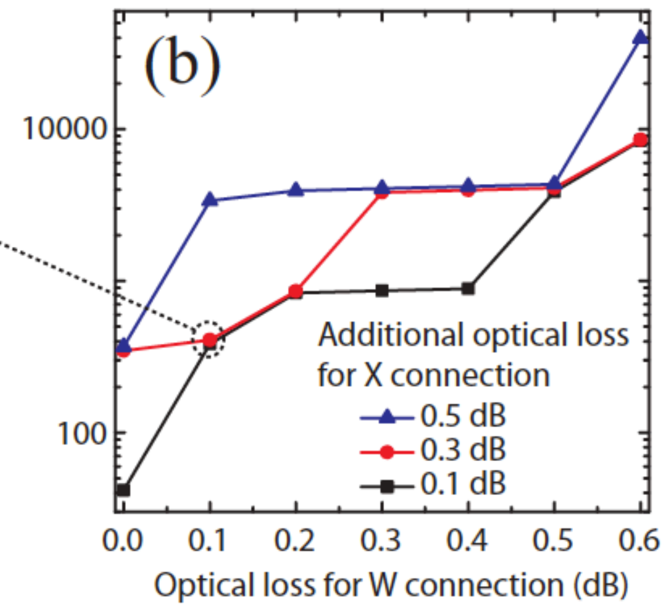

Gate pulse
Logic waveguide
Cavity containing quantum dot for lattice qubit
Microdisk waveguide
Ancilla qubit
Z
=


|ϕ⟩
|0⟩
|0⟩
|Ψ+⟩
H
H
H
Z
Lattice qubit
Post-purification, high-fidelity, long-distance Bell pair
Waveguides
Probe pulse
Purification waveguide
Logic waveguide
Racetrack
Switch


To another column, chip, or racetrack
Switch
Transceiver qubit
Ancilla qubit
Lattice qubit
Purification waveguide
Logic waveguide
|Ψ+⟩
|Ψ+⟩
|0⟩
=
H
H
H
H
Racetrack Waveguide
Purification Waveguide

Performance Depends on Interconnect





rdv, TDL, AGF, YY,
IJQI 8, 2010

Complete System

Large-scale system hardware:

▲ chip: 128 columns x 770 rows

▲ 64K chips

▲ 16M laser ports

▲ 16M measurement devices

▲ 6E9 physical lattice qubits

▲ 10GHz pulse rate

Functional requirements:

▲ adjusted gate error 0.2%

▲ local optical loss 0.02%

▲ working qubit yield of 40%

▲ memory lifetime 50 msec

Surface code:

▲ lattice refresh time 50 μsec

▲ lattice holes 14x14

▲ Toffoli gate time 50 msec

▲ 117K logical qubits:

12K application qubits

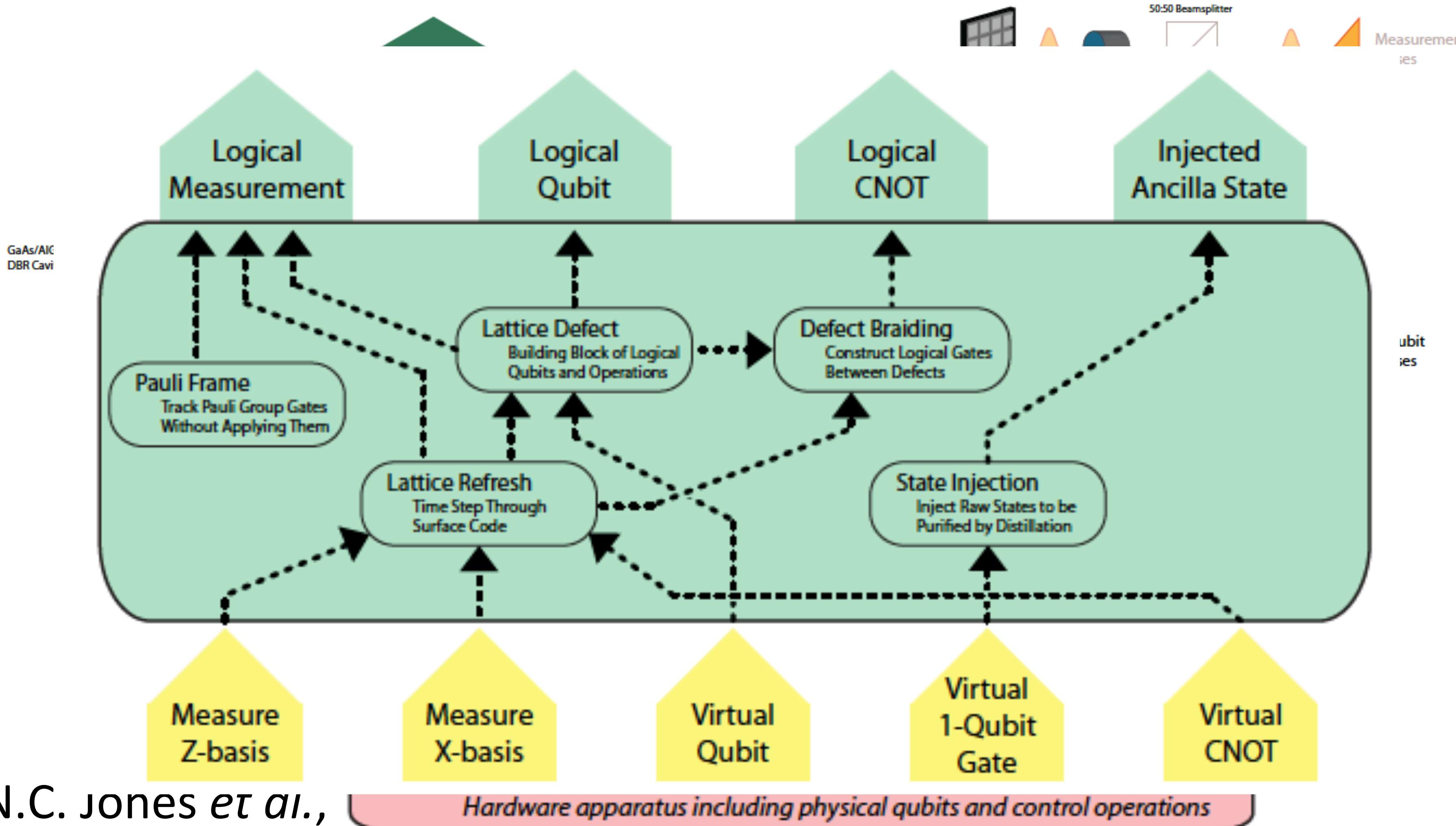
76K singular qubit “factories”

29K “wiring”

System Performance:

▲ 400 days to factor 2048-bit number

Evolution of Architecture



N.C. Jones *et al.*,
arXiv:1010.5022v1 [quant-ph]

PRL **106,** 050502 (2011) PHYSICAL REVIEW LETTERS week ending 4 FEBRUARY 2011

Does Adiabatic Quantum Optimization Fail for NP-Complete Problems?

Neil G. Dickson and M. H. S. Amin

D-Wave Systems, Inc., 100-4401 Still Creek Drive, Burnaby, British Columbia, V5C 6G9, Canada

(Received 13 October 2010; published 2 February 2011)

It has been recently argued that adiabatic quantum optimization would fail in solving NP-complete problems because of the occurrence of exponentially small gaps due to crossing of local minima of the final Hamiltonian with its global minimum near the end of the adiabatic evolution. Using perturbation expansion, we analytically show that for the NP-hard problem known as maximum independent set, there always exist adiabatic paths along which no such crossings occur. Therefore, in order to prove that adiabatic quantum optimization fails for any NP-complete problem, one must prove that it is impossible to find any such path in polynomial time.

DOI: 10.1103/PhysRevLett.106.050502 PACS numbers: 03.67.Ac, 03.67.Lx, 05.30.−d, 89.70.Eg

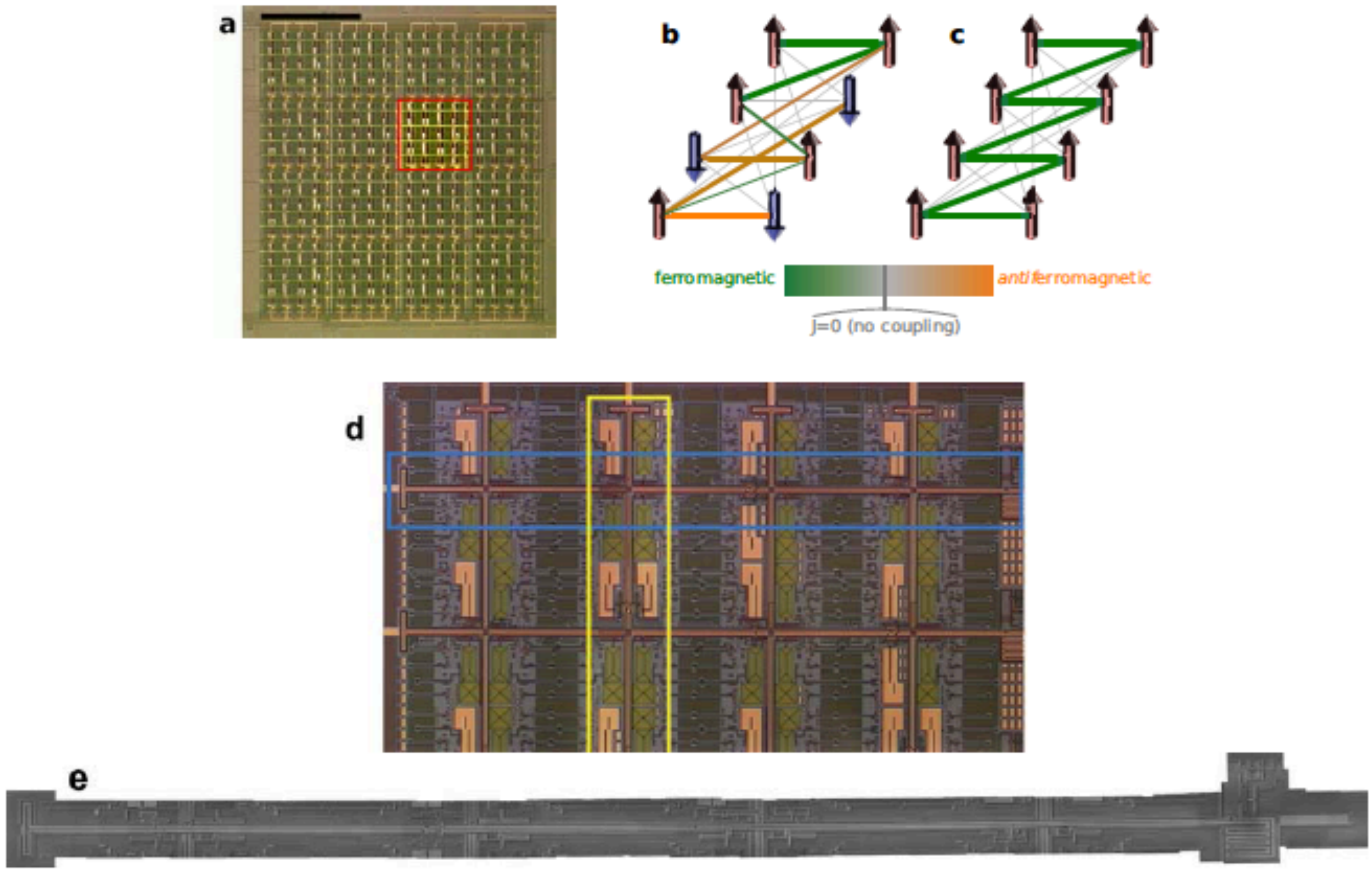
Adiabatic quantum optimization (AQO) was originally proposed [1] as a possible means for solving NP-complete problems faster than classical computation. In AQO, the Hamiltonian of the system is evolved from an initial form, H_B, whose ground state defines the initial state of the system, to a final Hamiltonian H_P, whose ground state is the optimal solution to an optimization problem. To ensure

Moreover, the possibility of avoiding small gaps by changing adiabatic path was again pointed out by Farhi *et al.* [11], and the fact that one problem can be mapped into many different Hamiltonians with different gap behavior was mentioned by Choi [12]. Those arguments, however, were based on numerical calculations for small problems, therefore inconclusive for large scales.

D-Wave Processor

Pics of chip

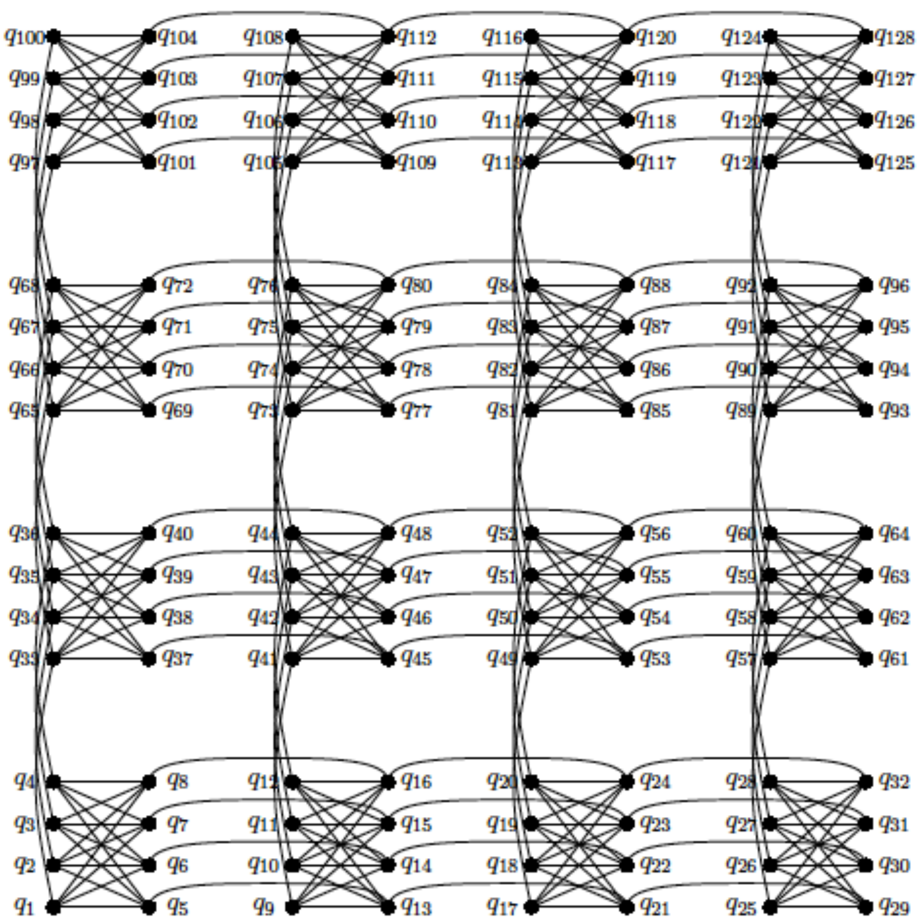
Functional qubits & couplings



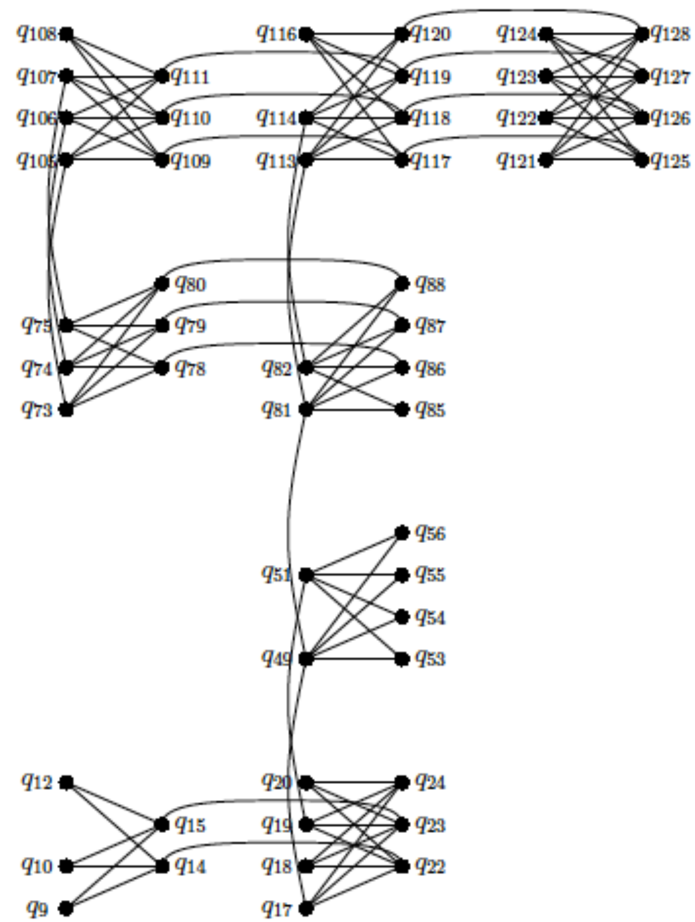
Johnson *et al.*, Nature 2011, online supplementary

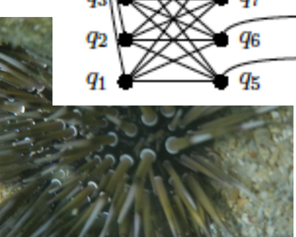
D-Wave Processor

Fabbed structure



Functional qubits & couplings





"The Ising Model", Bian, Aug. 2010

What is D–Wave Doing Right?

- Focusing on control for medium-scale systems
  – 1632 control signals needed for 128 flux qubits
  – Too many for external control
  – Programmable on-chip non-volatile memory(!) holds time-independent bias
  – Share external signal current for time-dependent
  – Reduced to 83 external signals
- Heterogeneous interconnect
- Defect-tolerant
  – 40% yield in previous slide, now 75%?
- Probably patenting refrigeration, packaging, self-test and diagnostics, dynamic control, noise suppression/magnetic shielding, programming tools…
- All of this requires a large, well-funded team

Harris *et al.*, Phys. Rev. B **82**, 024511 (2010) & matching slides

Conclusions

Summary

- Surface code is appealing
  - High threshold
  - Nearest-neighbor only
  - Resource requirements still high
- Architecture is critical
  - Hardware-software co-design
  - Heterogeneous interconnects necessary
- Any statement about "QC will do X in seconds," needs to be put in context of a specific algorithm, architecture, and technology!

AQUA: Advancing Quantum Architecture

When will a paper be published in *Science* or *Nature* in which the point is the results calculated using a quantum computer, rather than the machine itself?

That is, when will a quantum computer *do* science, rather than *be* science?

Key References

- See 14-page handout
- Fowler *et al.*, *PRA* 80, 052312 (2009)
- Ladd et al., "Quantum computers", *Nature* 2010
- Spiller et al., *Contemp. Phys.*, 2005
- rdv & Oskin, *JETC* 2, 2006
- Steane, *PRA* 68, 042322 (2003) (& others)
- Proc. Int. Symp. Comp. Arch. papers of Oskin, Chong, Kubiatowicz, rdv
- Papers of Fowler & Devitt
- Simulation: Kendon & Munro,
  Clark, Brown et al., arXiv:0810.5626
- ITRS

Key Aqua References

- *IJQI*, 2010
- Quantum multicomputer architecture:
  JETC 3(4): quant-ph/0607160
  my thesis: quant-ph/0607065
- Workload:
  PRA 71, 052320: quant-ph/0408006
  MS+S2006: quant-ph/0507023
- Purification scheduling:
  IEEE/ACM Trans. on Networking, Aug. 2009:
  quant-ph:0705.4128