Perspectives on Complex Systems Software Engineering

40th Anniversary Distinguished Alumnus Lecture

Department of Computer Science
College of Engineering, Virginia Tech

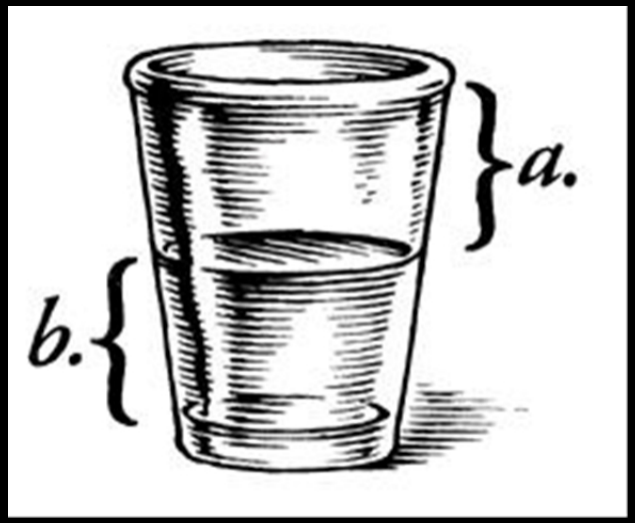
12 November 2010

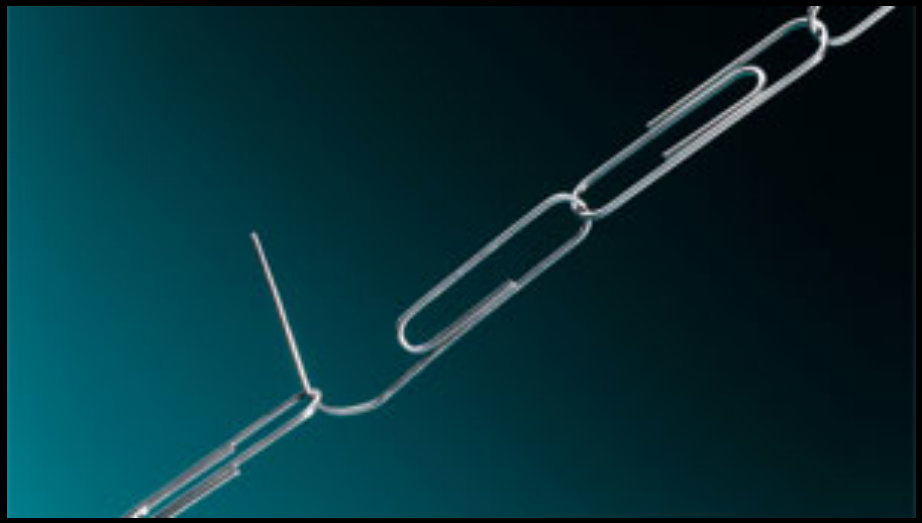
Dedicated to Robert Keith Lavender
and the memory of Polly Wright Lavender

Robert Gregory Lavender
VTCS M.S. 1988, Ph.D. 1993
Vice President of Foundation Engineering
Network Software & Systems Technology Group
Cisco Systems, Inc.

greg.lavender@cisco.com

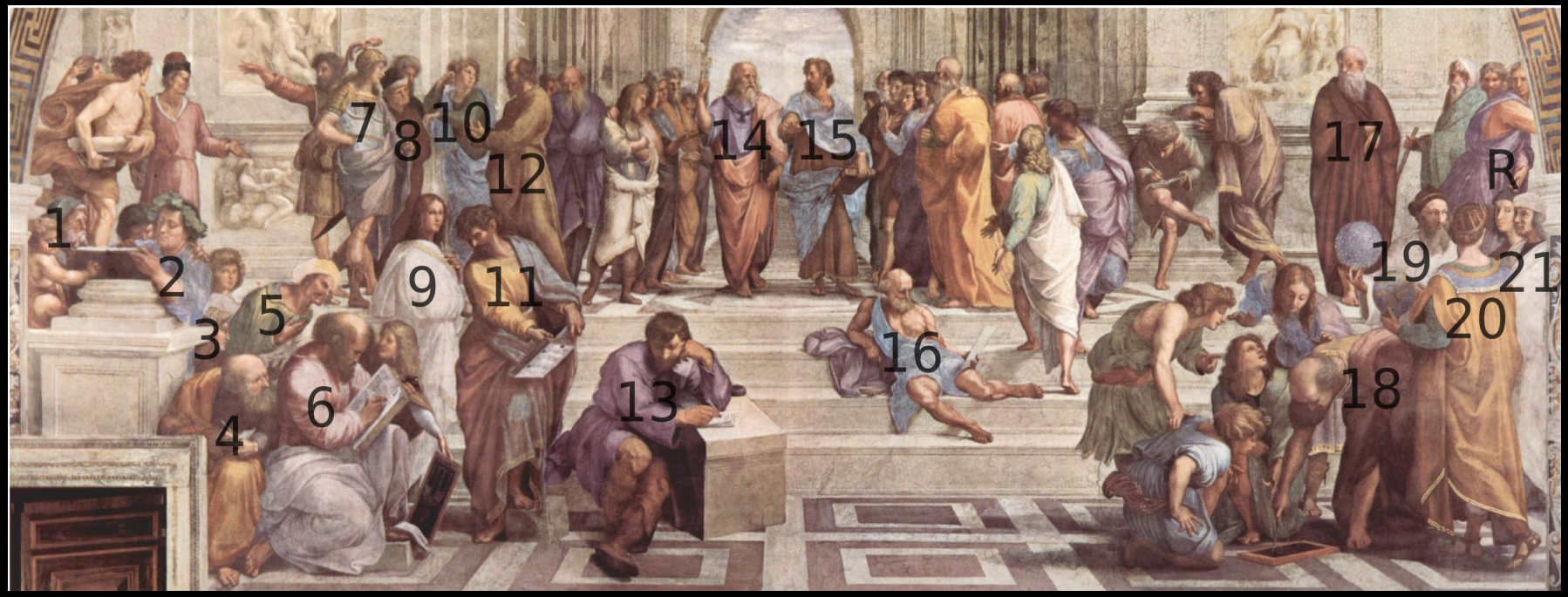
Perspective and Perseverance





Memoria – The Past

Allegory of the Abstract vs the Concrete



Typus Algoritmi

- Abu Adbdallah Muhammad ibn Musa Al-Khwarizmi
  - c. 780-850 ACE in Persia (Uzbekistan)
    - Scholar in the House of Wisdom in Baghdad
  - “Kitab al-jabr wa ‘l-muqabala”
    - “rules of equating & restoring”
    - Al-jabr => Algebra
    - Al-Khwarizm => Algorism
    - Algorism => Algorithm


```
-- Euclid's algorithm
gcd 0 0 = error
gcd x y = gcd' (abs x) (abs y)
    where
       gcd' x 0 = x
       gcd' x y = gcd' y (x `rem` y)
```

Margartia Philosophica

- *The Pearl of Wisdom*
  - An encyclopædia used as a popular university textbook in Renaissance Europe, published around 1500
  - Authored by Gregor Reisch, a monk and *Magister* who had a great reputation as an Oracle of knowledge and wisdom
- The Three faces of Philosophy
  - Memoria (the past)
  - Intelligentia (the present)
  - Providentia (the future)
- The Seven Liberal Arts
  - Logica, Rhetorica, Grammatica, Arithmetica (sitting using an abacus), Musica, Geometria, Astronomia

The Magister and the Disciple

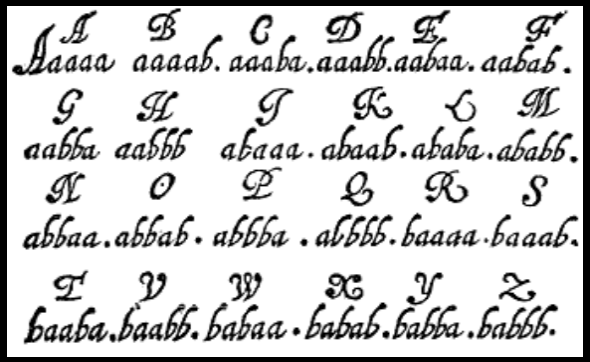
- The *Margarita Philosophica* contains twelve books:
  - Trivium:
    - Grammar – symbols & syntax
    - Dialectics – logic & semantics
    - Rhetoric – blogging ☺
  - Quadrivium: from Plato's *Republic*
    - Arithmetic – pure number
    - Music – number in time
    - Geometry – number in space
    - Astronomy – number in space & time
  - As well as:
    - Physics, Natural History, Physiology, Psychology, and Ethics

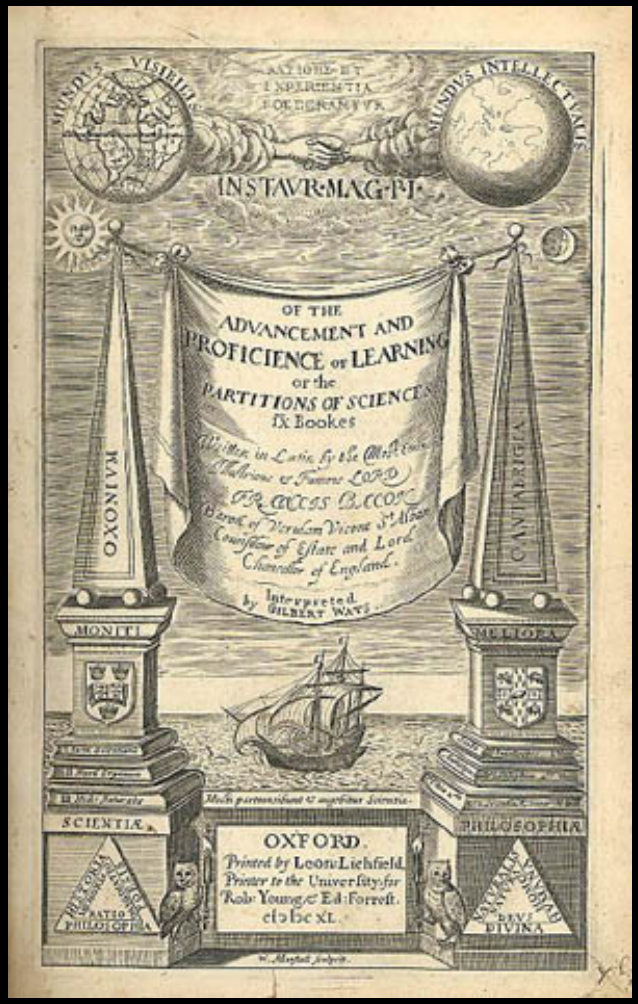
Boethius and the Quadrivium

Typus Arithmetica et Logica

Typus Scientifica Methodologia

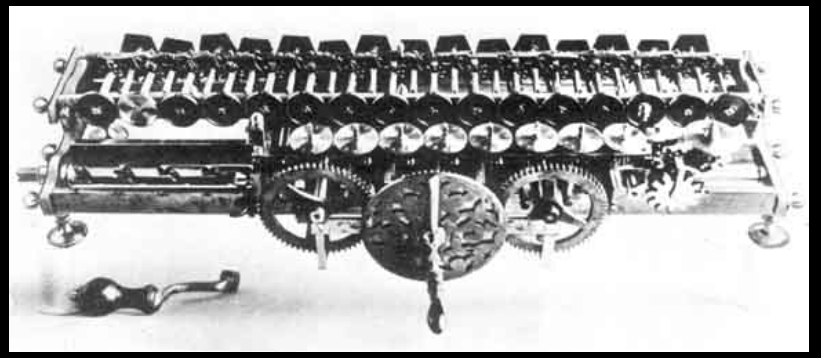
- Francis Bacon – Father of Empiricism
  - "For by this Art a way is opened, whereby a man may expresse and signifie the intentions of his minde"





Typus Mechanicus Universalis

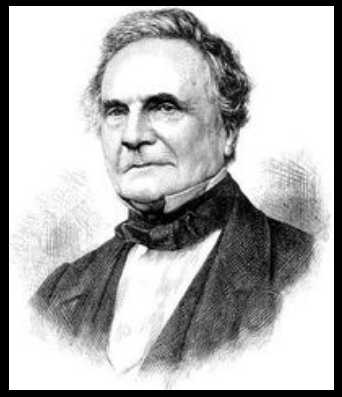
000	0	0
001	1	1
010	10	2
011	11	3
100	100	4
101	101	5
110	110	6
111	111	7

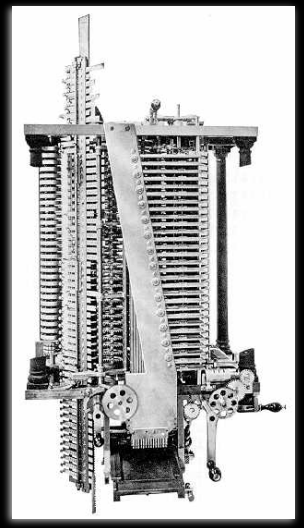


(*) See *History of Binary and Other Nondecimal Enumeration*, *Revised Edition*, Anton Glaser, Tomash Publishers, 1981. Out of print but available at: http://www.eipiphiny.org/books/history-of-binary.pdf

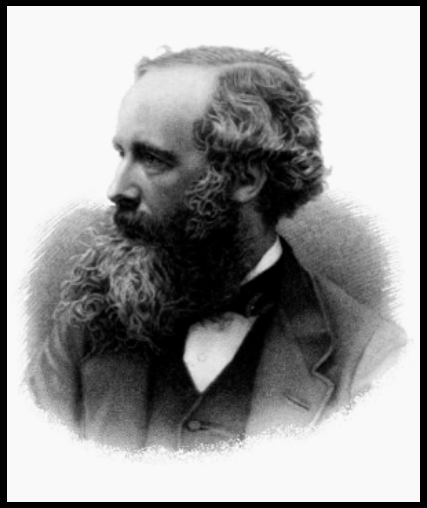
Typus Programmatica Mechanica

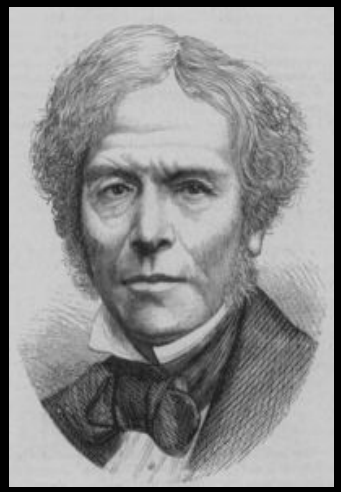
- Lady Ada Lovelace's mother insisted she study mathematics and science
  - Ada always loved poetry like her father, Lord Byron
  - She wrote to her mother saying: *if you can't give me poetry, can't you give me "poetical science?"*
- Ada learned of Babbage's machine at age 17 at a dinner party hosted by her mother
  - she was impressed by the universality of the idea and decided to devote herself to putting mathematics and technology into a human context
- In 1843, she presciently wrote a paper suggesting that Babbage's machine might be used to:
  - play music (158 years before the iPod)
  - produce graphics
  - lead to practical and scientific uses





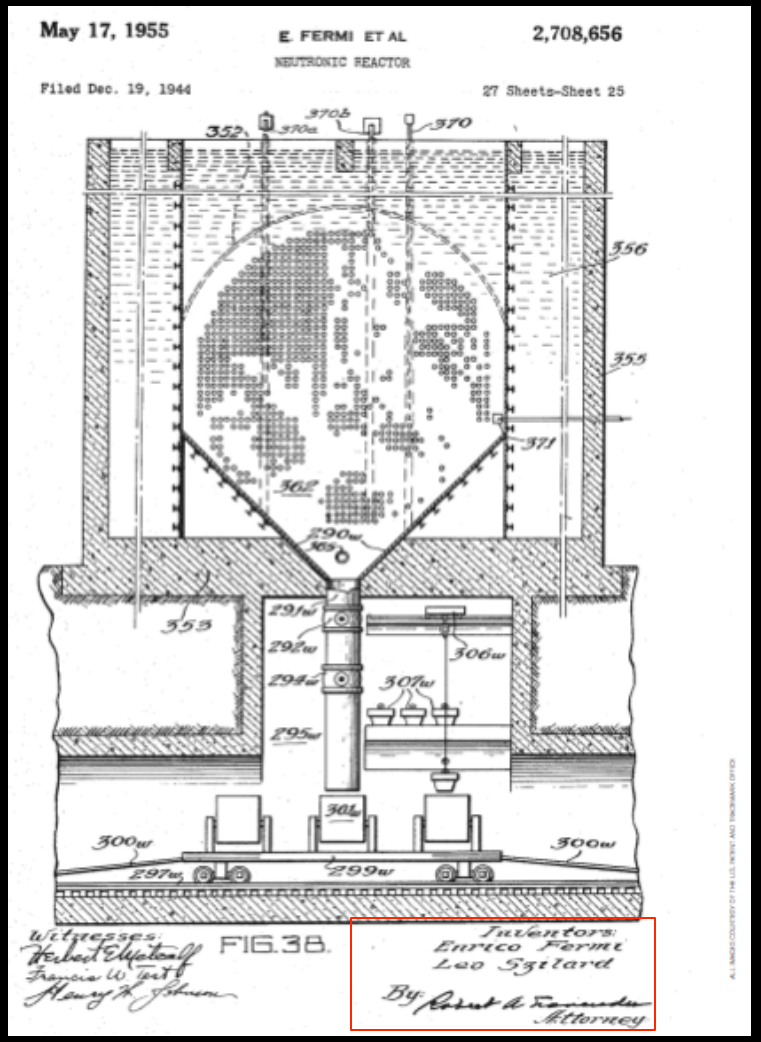
Typus Electro-Magneticus





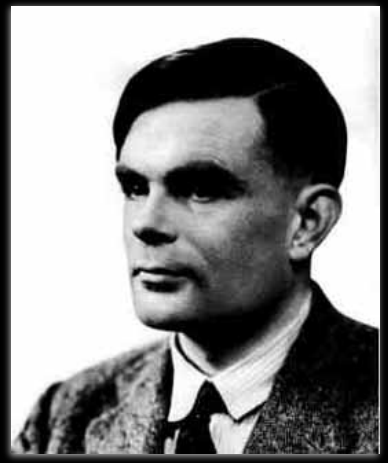
The Epoch of Scientific Power

- ~2,500 years of human cultural and economic history with rare epochs enabling the advancement of science
  - Fortunately, we live in one of those epochs
- 20th century geo-political and socio-economic forces
  - Accelerated the centuries long scientific process
  - WW II, the Cold War and the nuclear & space races
  - Human ingenuity driven by intellectual curiosity, but also fear, paranoia, ego and greed
- Vannevar Bush – OSR&D
  - *As We May Think* – essay published in 1945
  - Differential Analyzer, Memex, founded NSF
  - Atomic Bomb Patents – Alex Wellerstein@Harvard
    - http://www.people.fas.harvard.edu/~wellerst/atomic_patents/
    - http://www.npr.org/templates/story/story.php?storyId=89127786



Typus Universalis Computatis

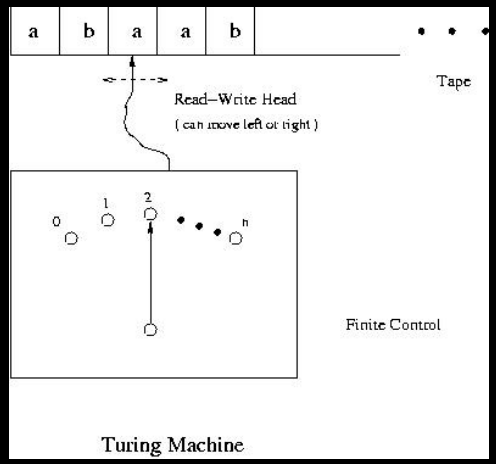


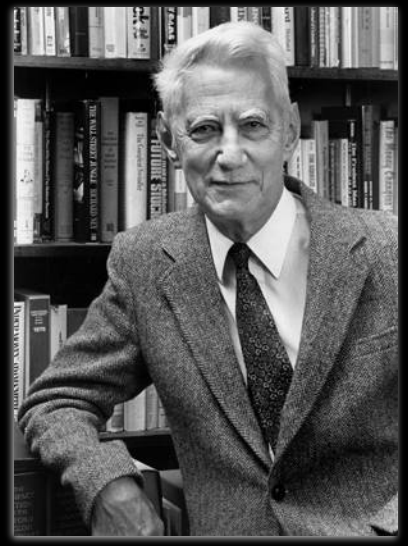
Untyped Lambda Calculus
(α-conversion, β-reduction)

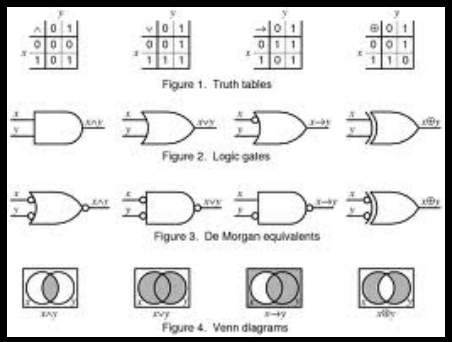
Exp ::= var | constant
| λ var '.' Exp -- function abstraction
| Exp Exp
| '(' Exp ')' -- function application

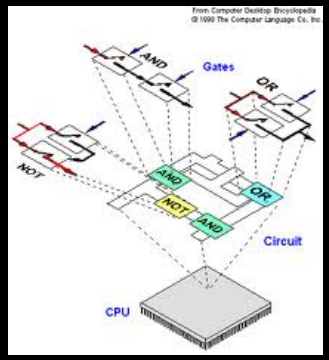
Y = λf.(λx.f (x x)) (λx.f (x x)) – lazy Y combinator

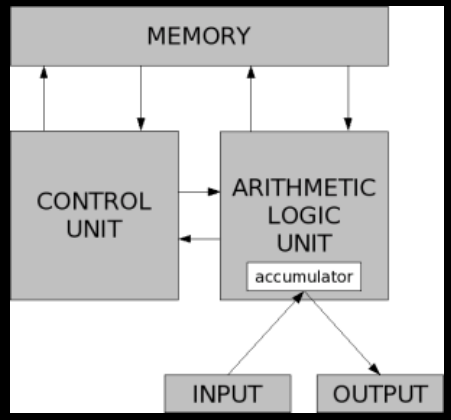


Typus Logica Digitalis Practicalis



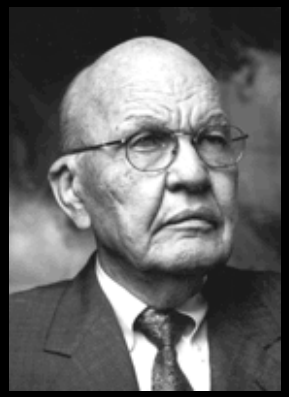


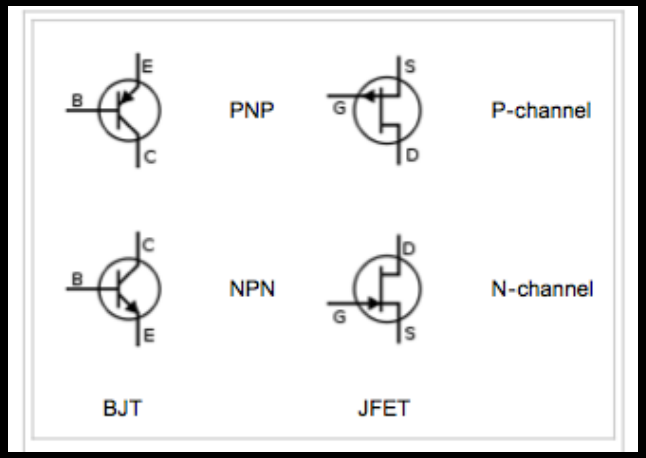


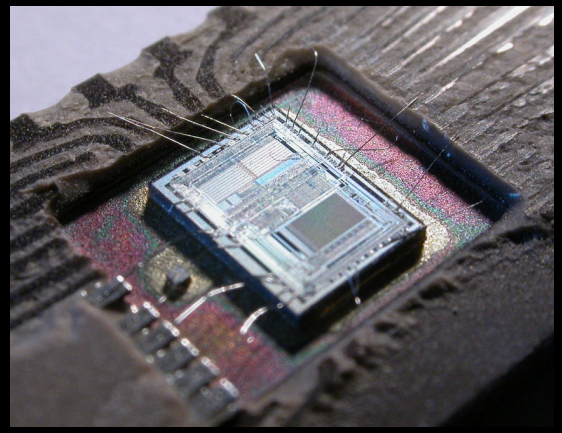


Typus Transistoris Micro-Circuitus

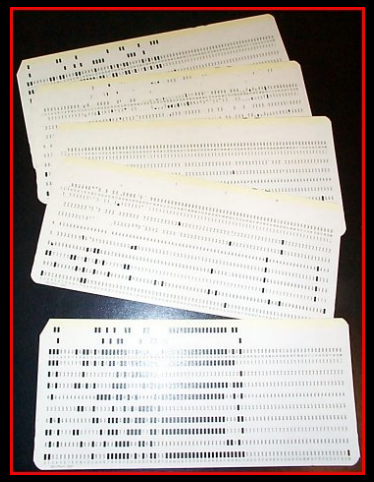


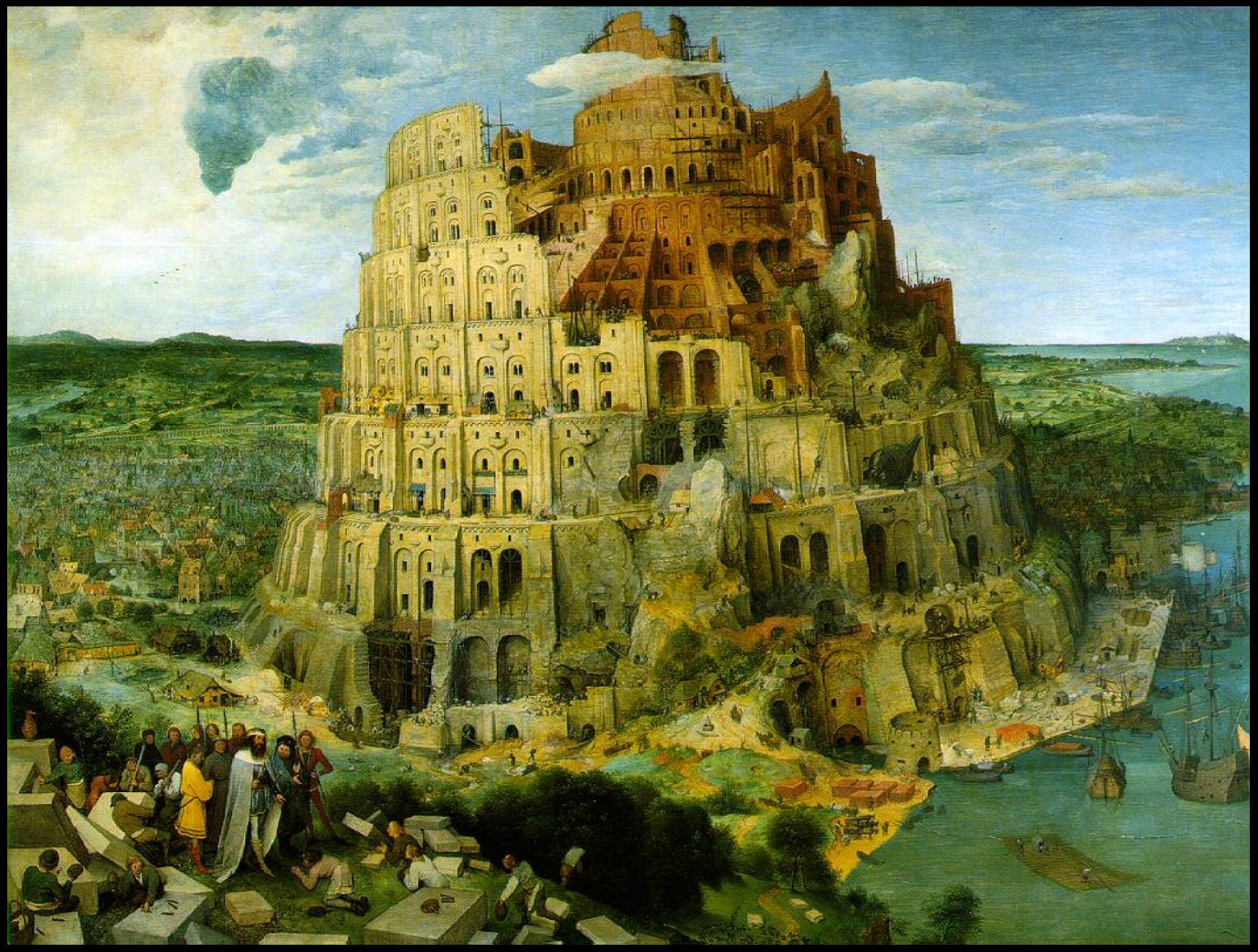





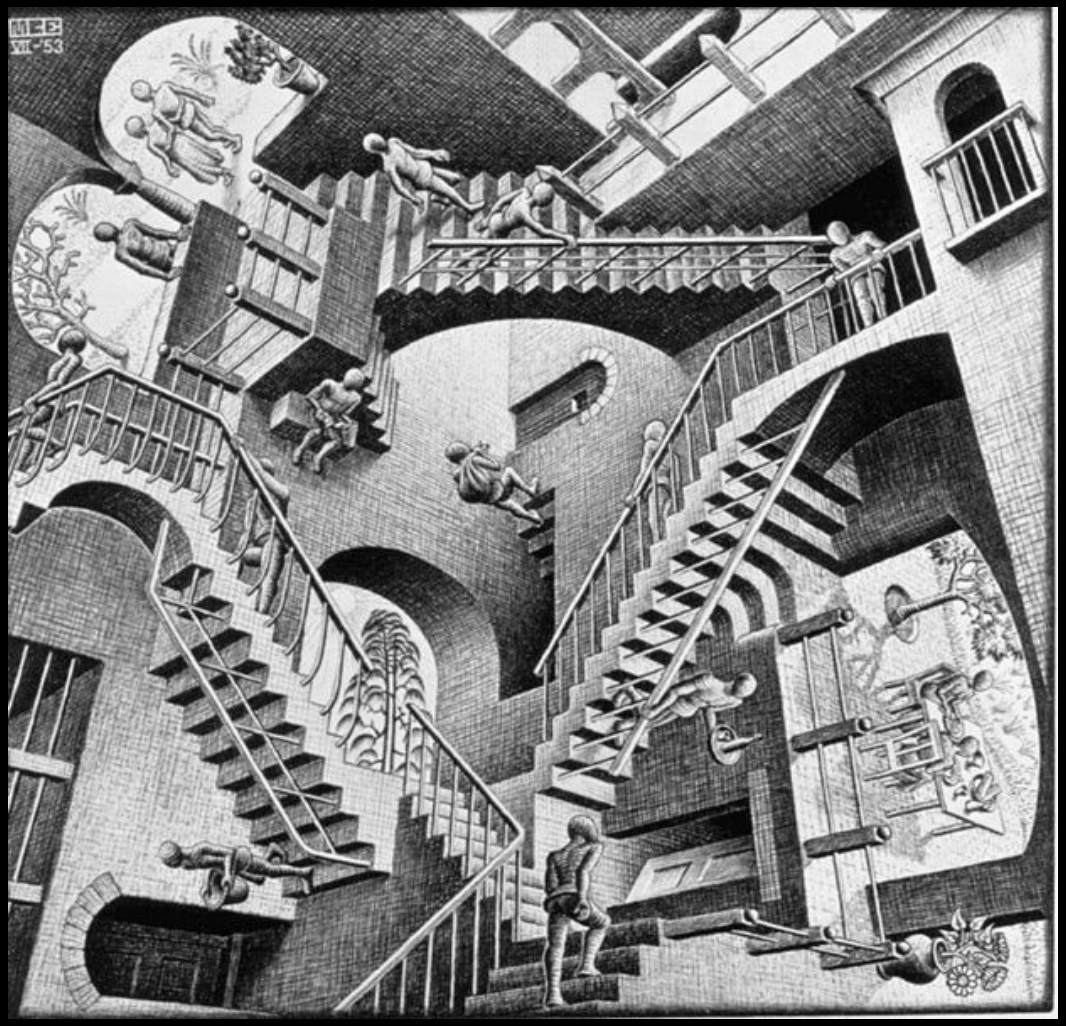
Typus Automatica Computatus



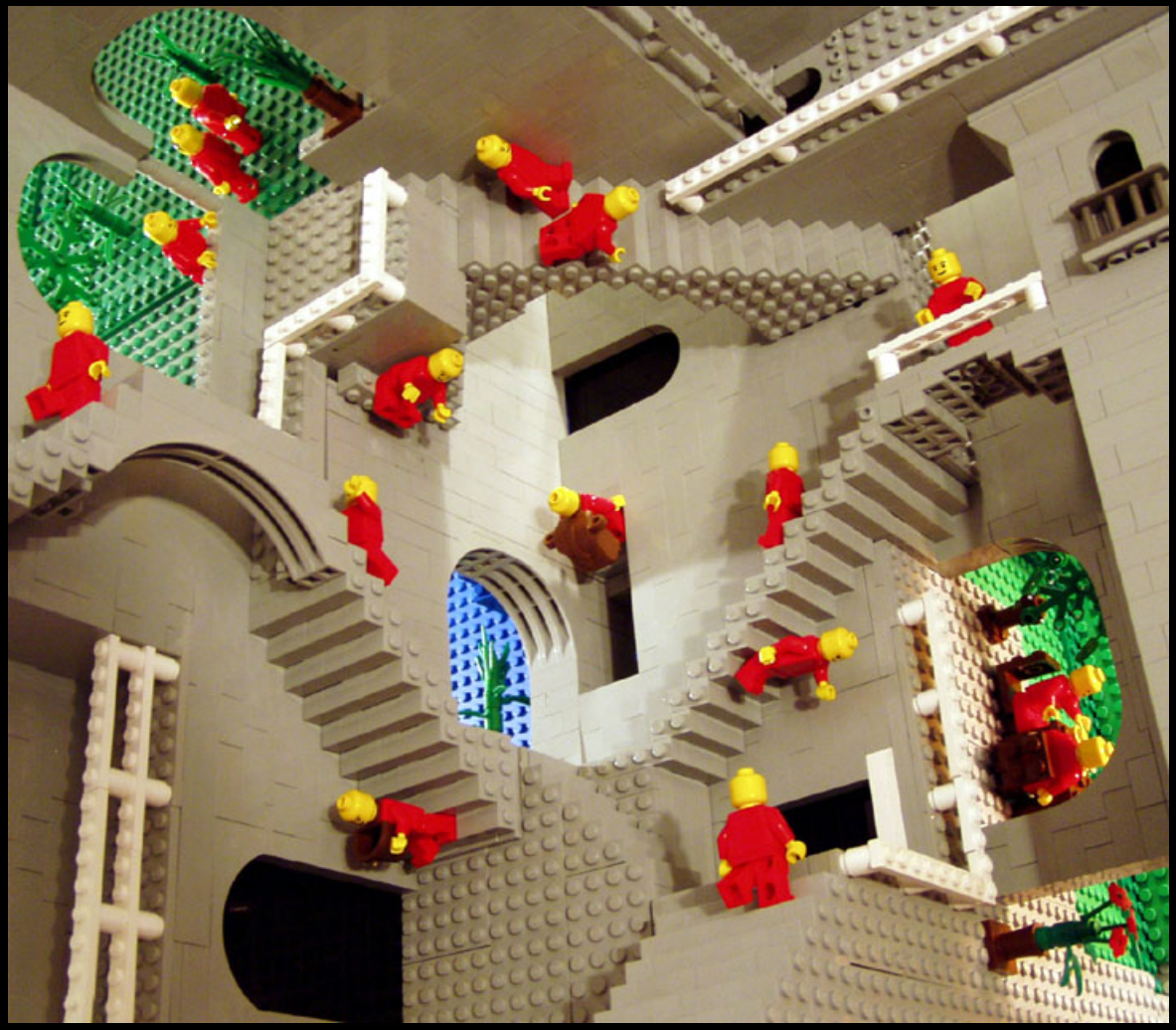
Intelligentia – The Present



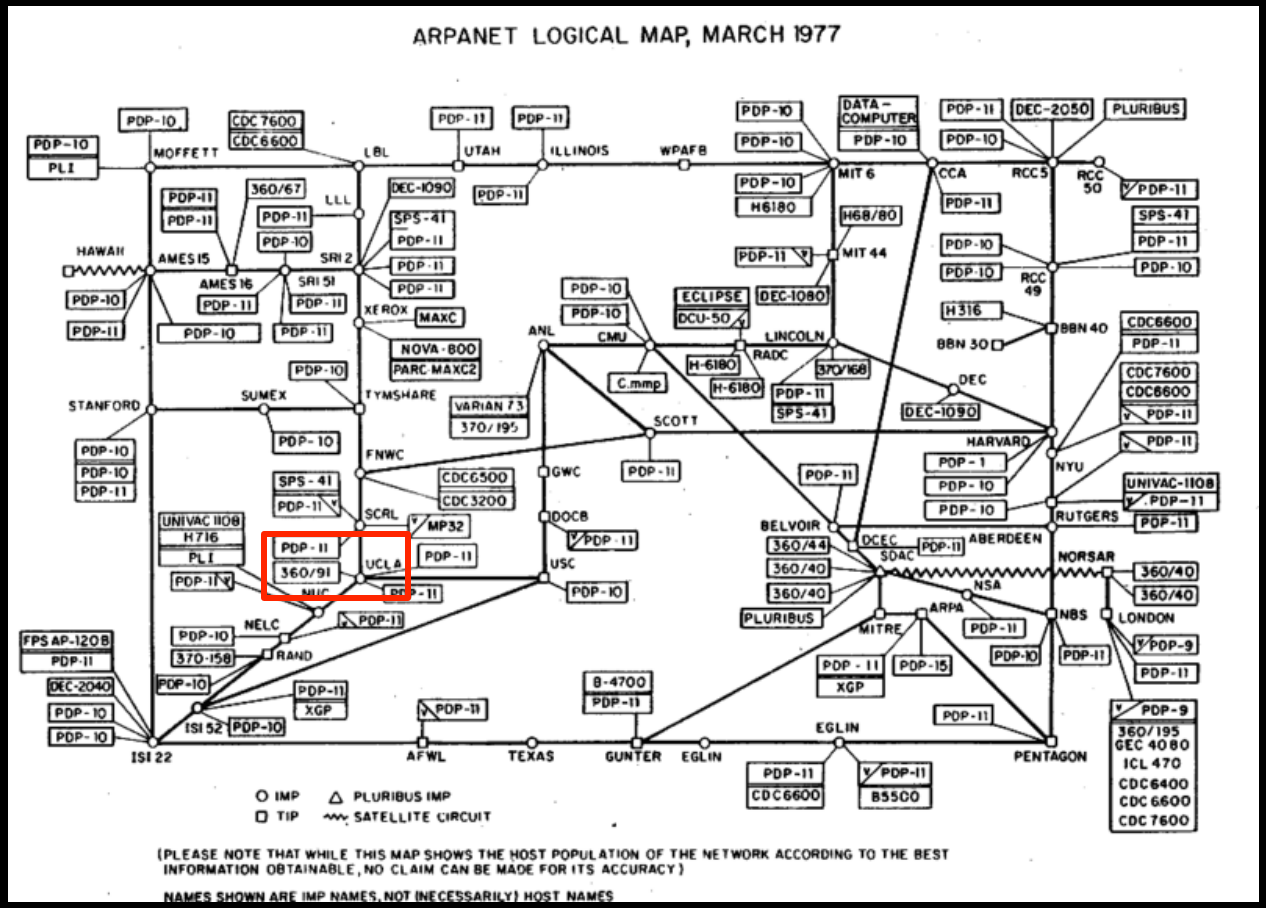
Orthogonal Layers of Abstraction



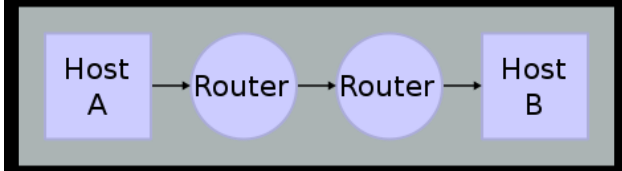
Over 40 Years We Remodeled Everything to be Object-Oriented

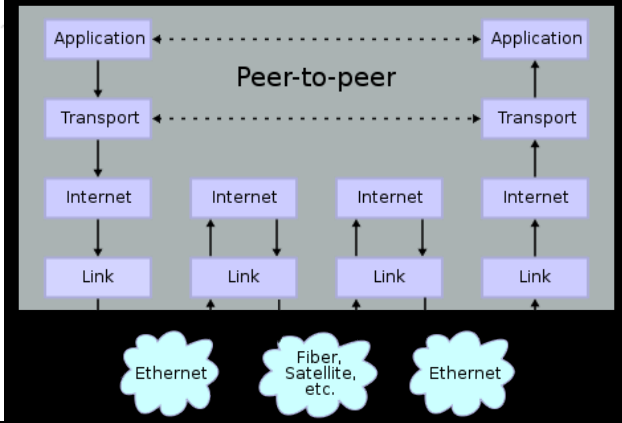


My Work Started Here in 1983



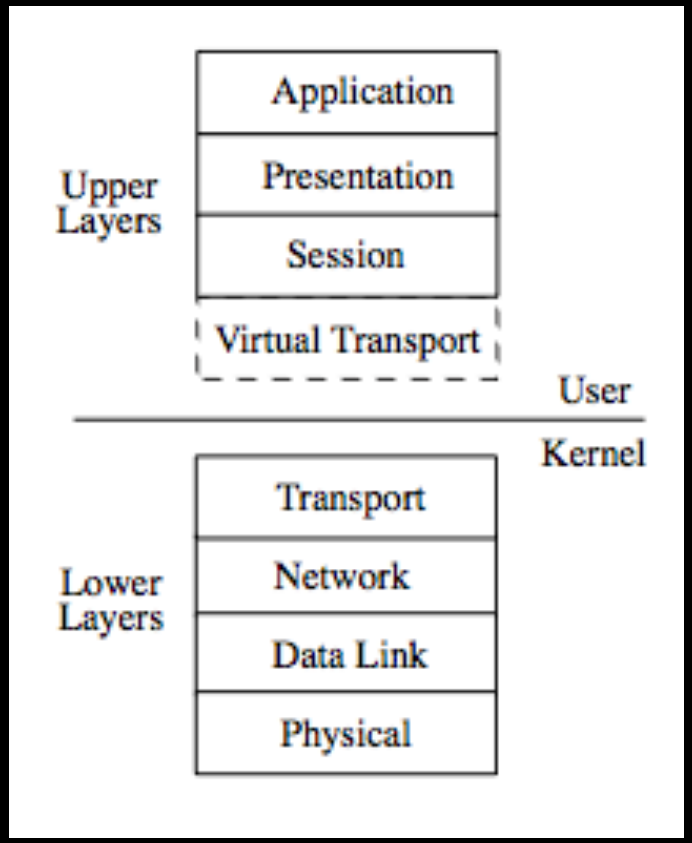
MIL-STD-1777
RFC 791 (IP)
RFC 793 (TCP)

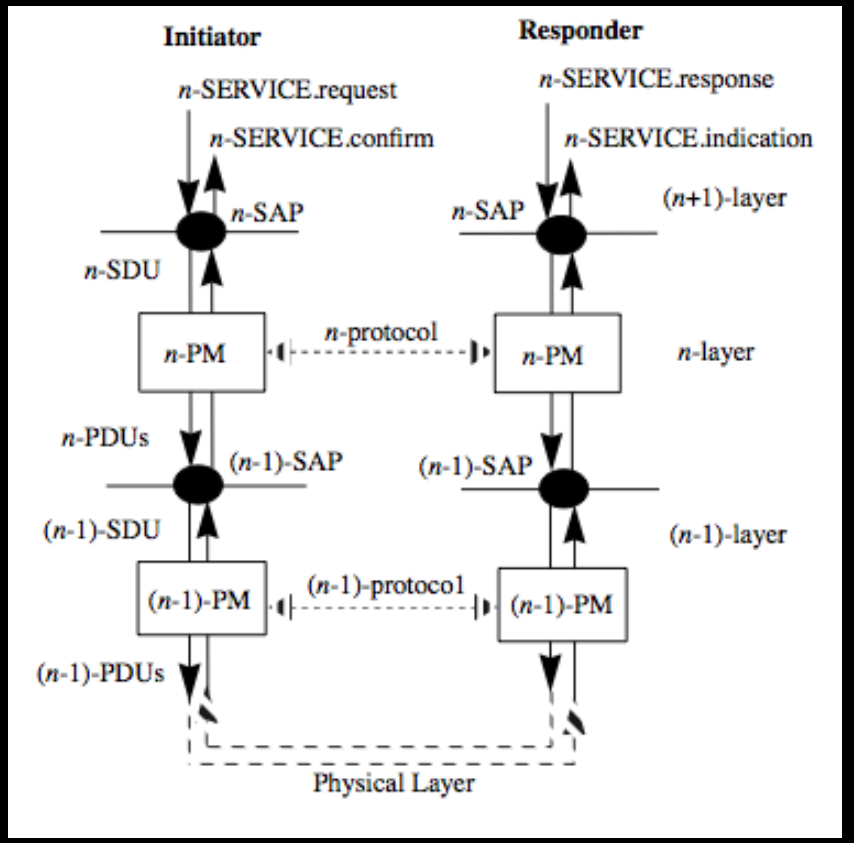




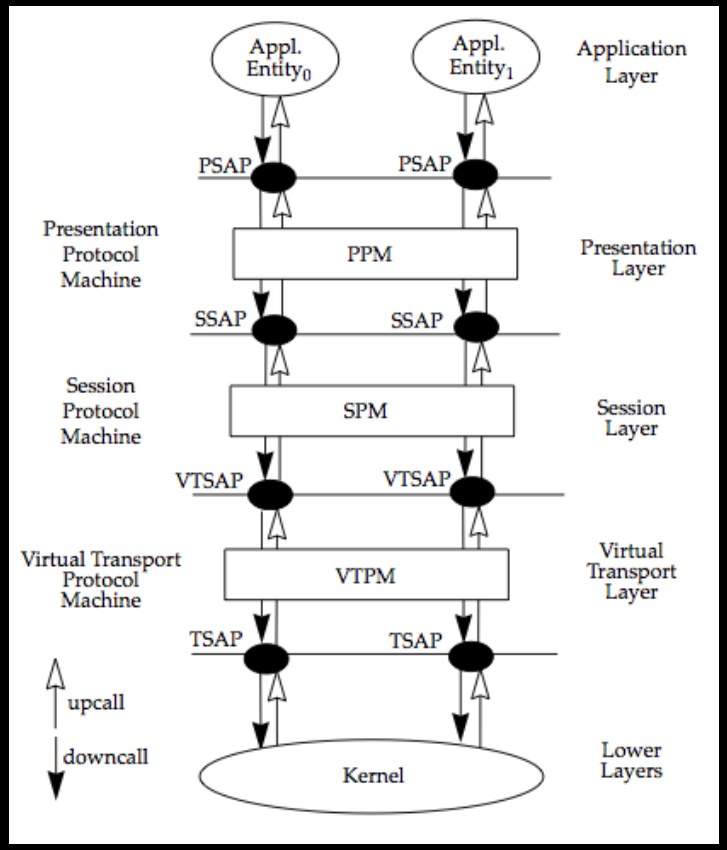
IBM System/370 Principles of Operation

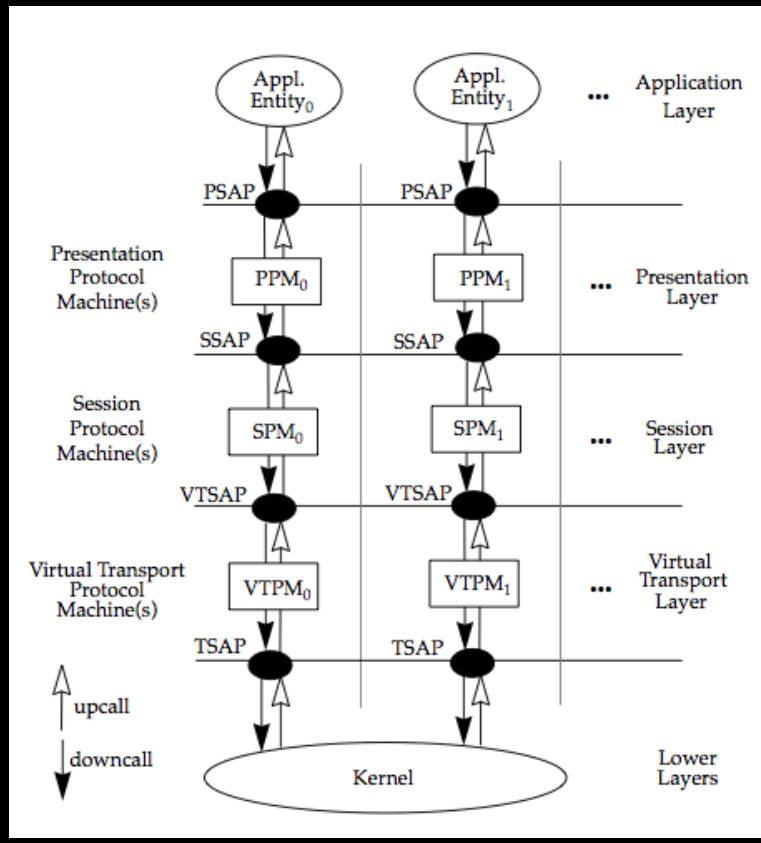
Virtualized User Space Network Stack (1990-1993)



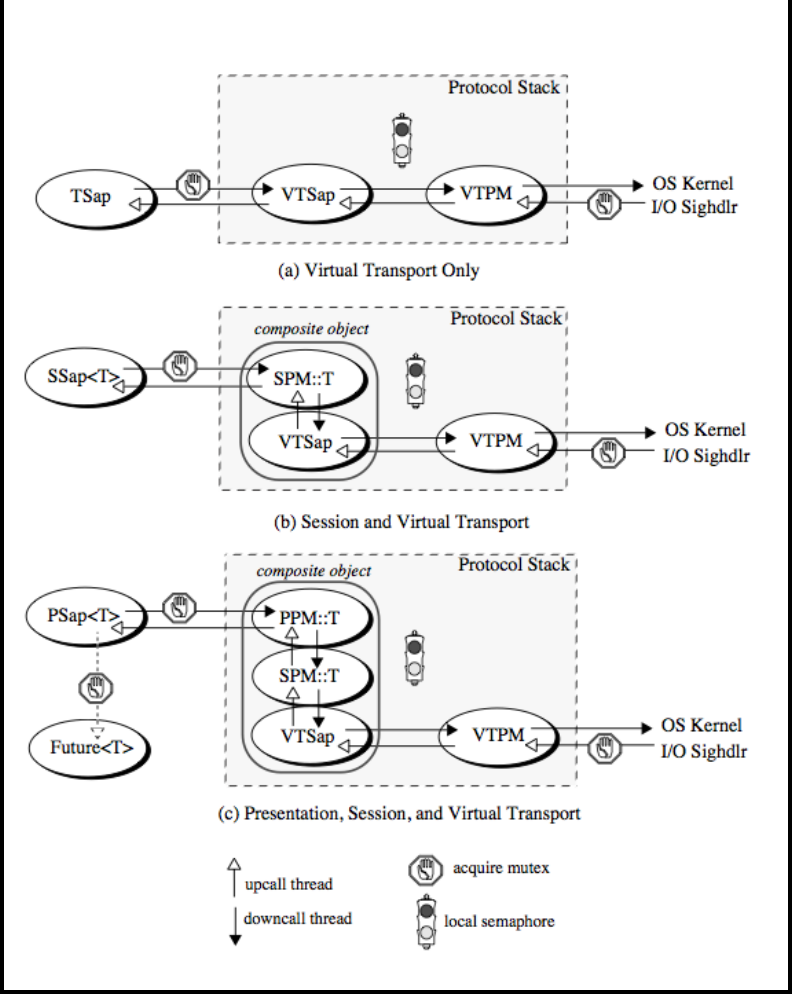


Vertically Sliced Lock-Free PMs





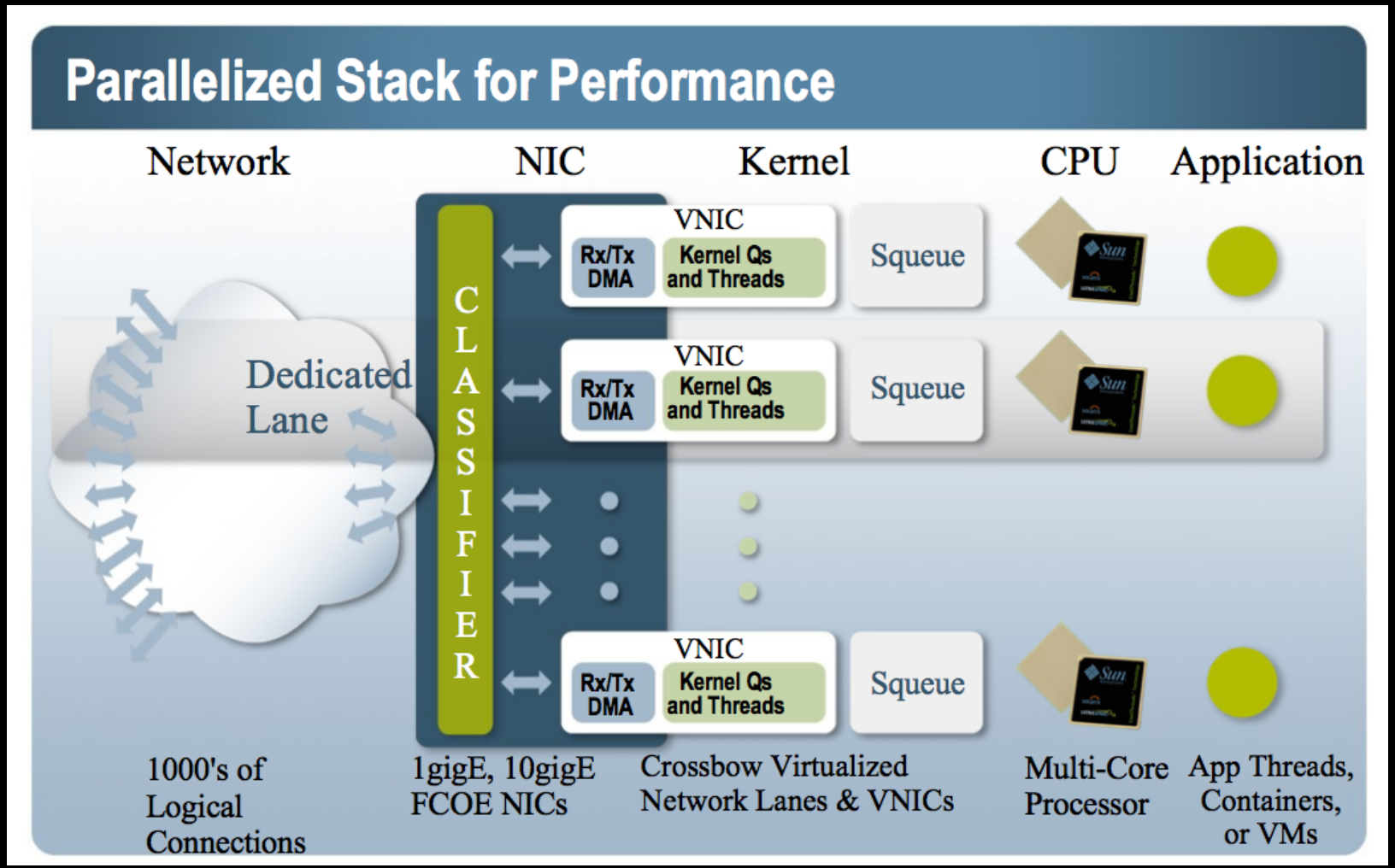
Polymorphically Typed Composable Protocol Machines



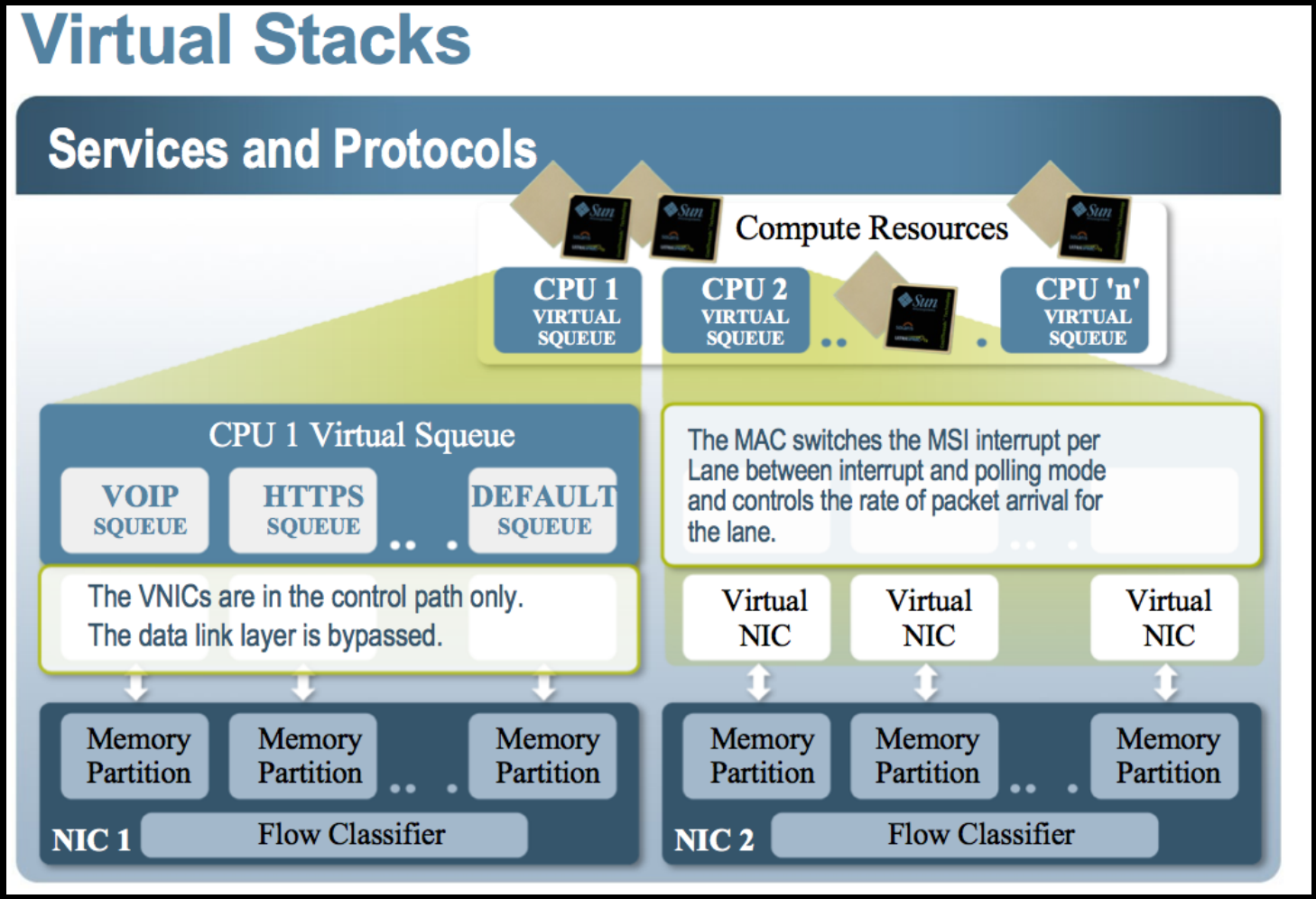
Insights Gained from MS & PhD

- 1988-1990: Virginia Tech, 1991-1993: MCC
  - Software engineering, theory of programming languages,and concurrency theory
    - OO and functional programming abstractions applied to networking
    - Hewitt/Agha's Actors, Milner's Calculus of Communicating Systems (CCS)
  - Protocol layering considered harmful IF taken literally at run-time
    - use syntactic protocol layering to maintain information hiding abstraction boundaries at compile-time but with minimal cost at run-time
  - Composable protocol state machines
    - integrated inter-layer packet processing with zero-copy and lazy evaluation
    - asynchronous interrupt driven virtual up-call dispatch
  - Efficient concurrent execution and asynchronous interrupt-driven communication
    - lightweight threads with lock-free semantics as much as possible
    - exploited Sun's new 2-CPU SPARC symmetric multi-processor (SMP) hardware for high performance (at the time)

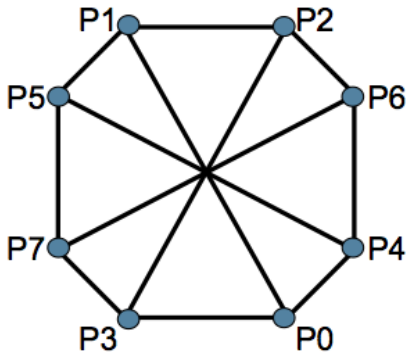
Virtualized Kernel Space Network Stack (2005-2010)



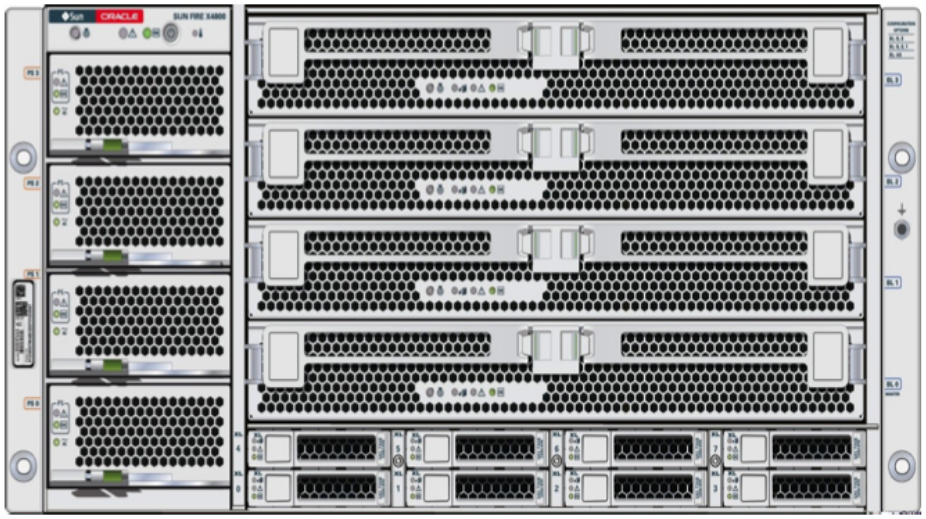
L2 HW Accelerated Multi-* Stack

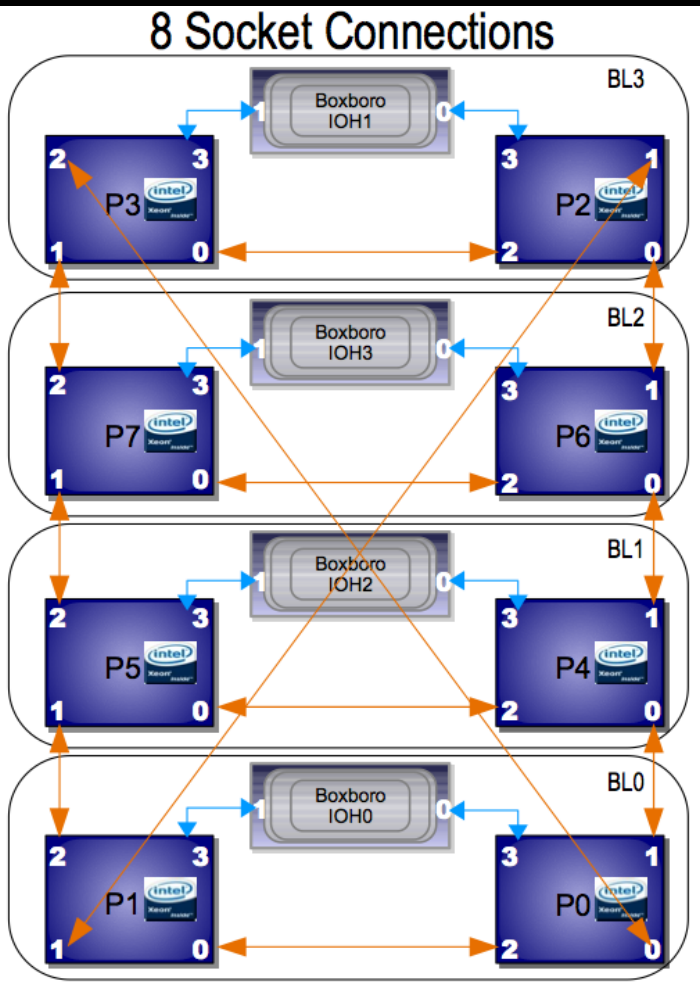


Latest Multi-* System

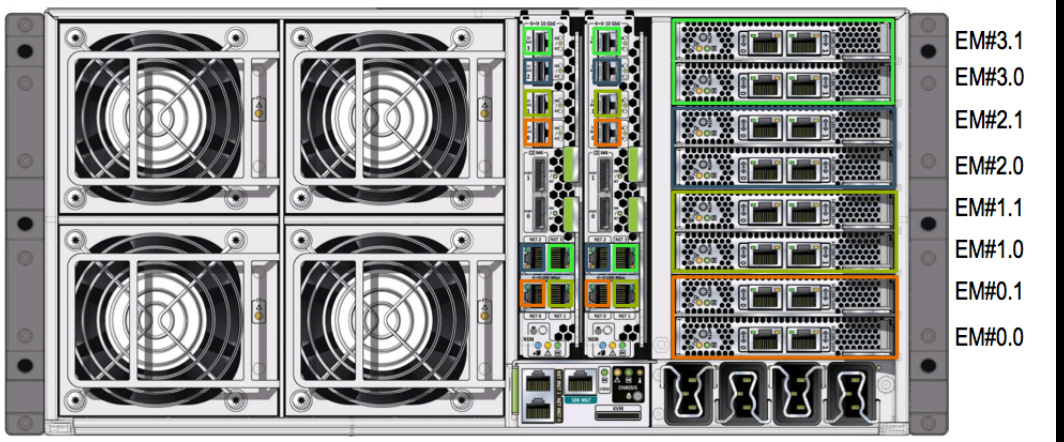


Topology





Networking I/O Smokes



Compute

- 8x Intel Nehalem-EX (128 vcpus, 24MB cache)
- 128x DDR3-1066 RDIMMs
- Up to 1 TB of memory (256GB/socket)

I/O

- 8x PCIe 2.0 EM slots (8x lane per EM)
- 8x 10GbE SFP+
- 8x GbE ports
- 8x 2.5" SAS/SATA/SSD drives

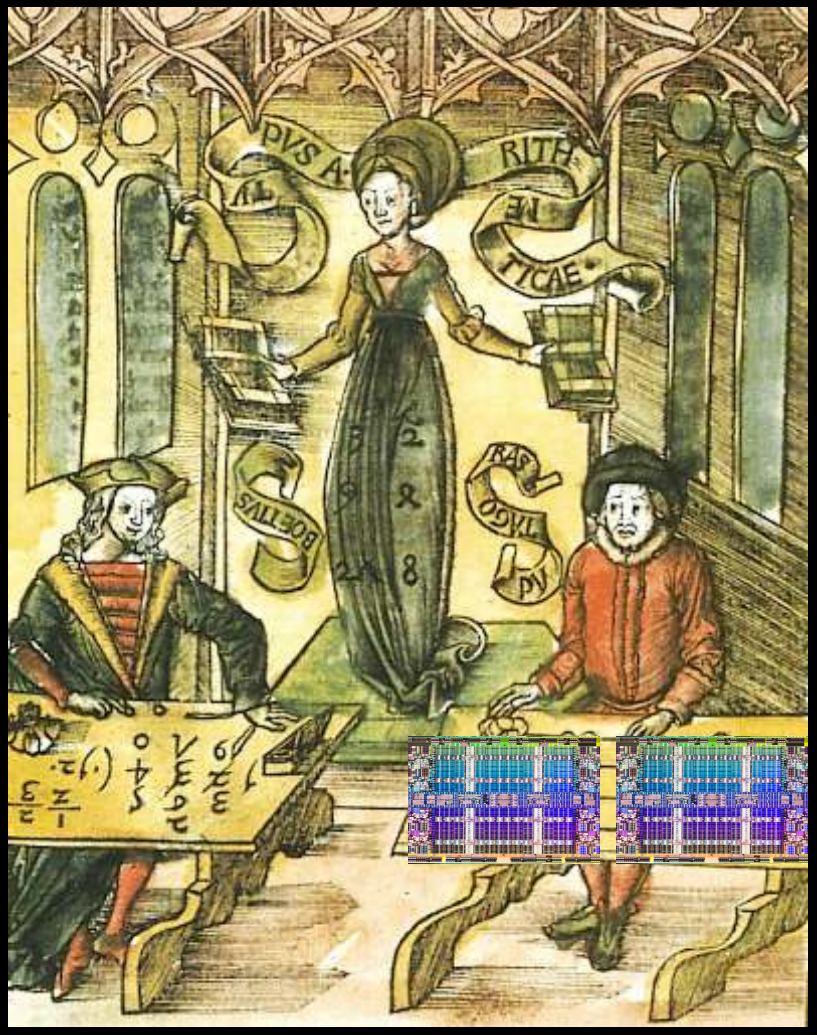
Availability

- Hot-plug disks
- HW RAID 0,1,5,6,10,50,60 (w/REM)
- Hot-swap redundant 2000W PSUs
- Hot-swap redundant fans
- Hot-swap I/O and SP

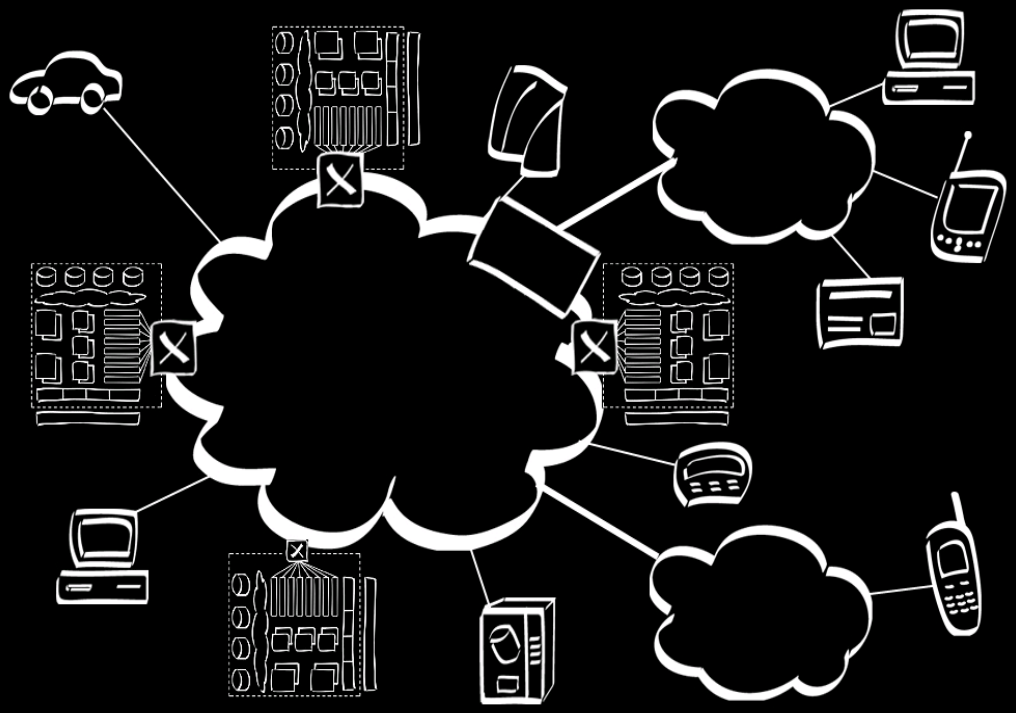
Management

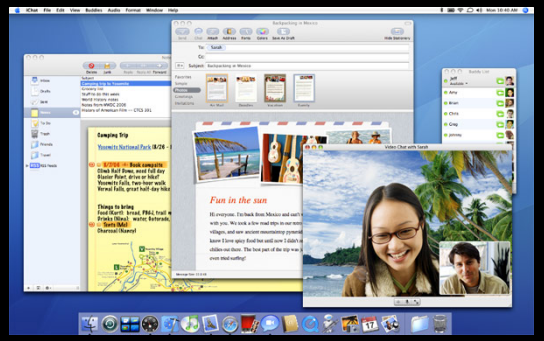
- Integrated LOM Service Processor 3.0
- Solaris, Linux, Windows or VMware

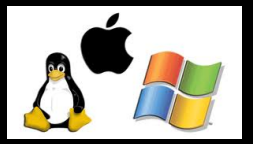
Pythagoras has a new Calculi Board



Providentia – The Future



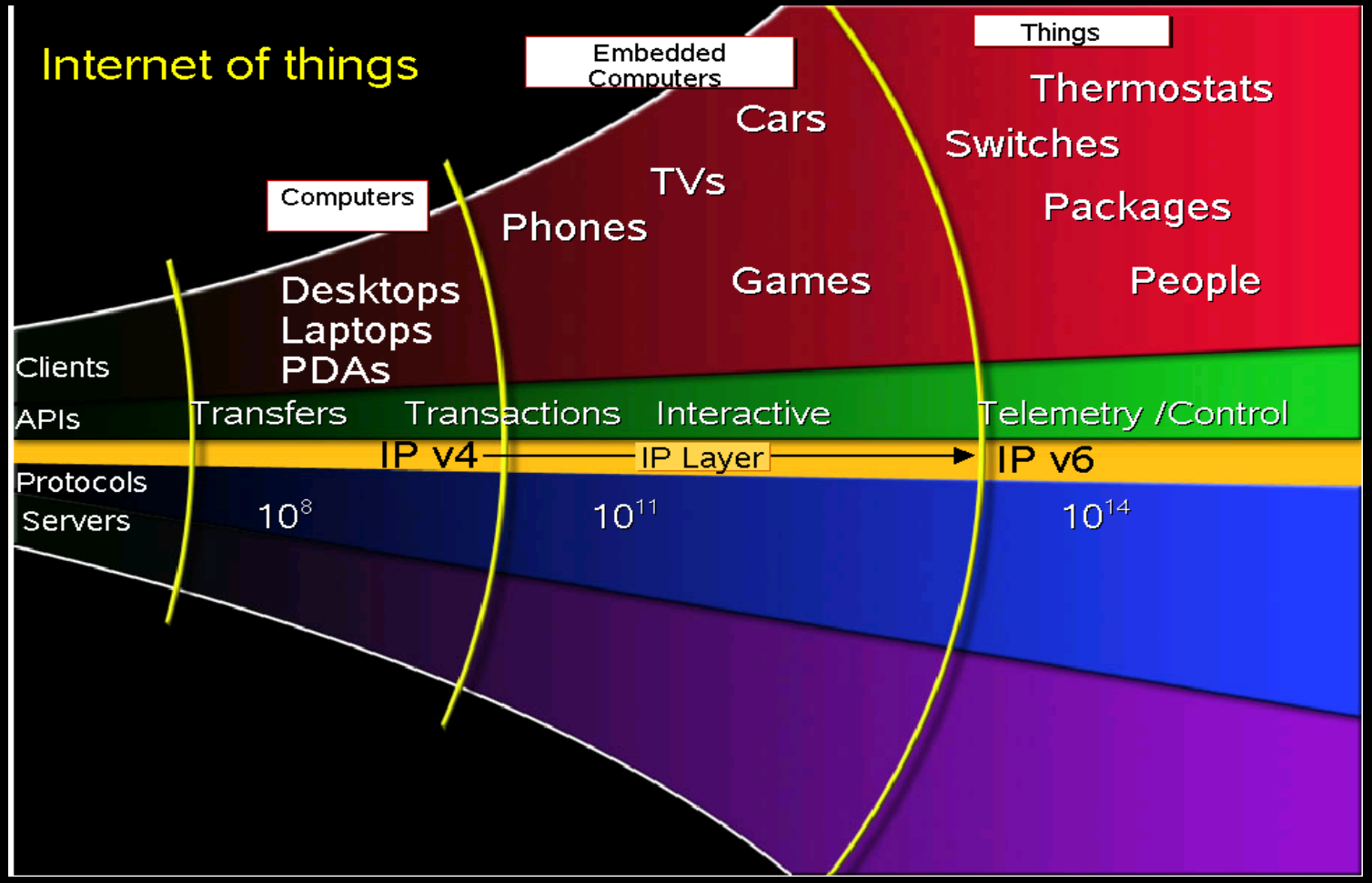




Chaos reigns within
Reflect, repent, and reboot
Order shall return

– "Haiku" system error message

Everything is Being Connected

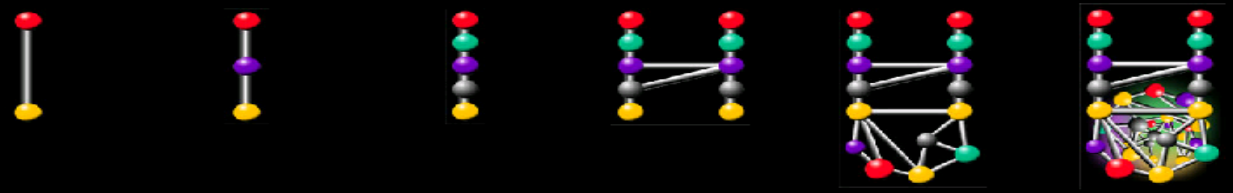


The Complexity is also Increasing

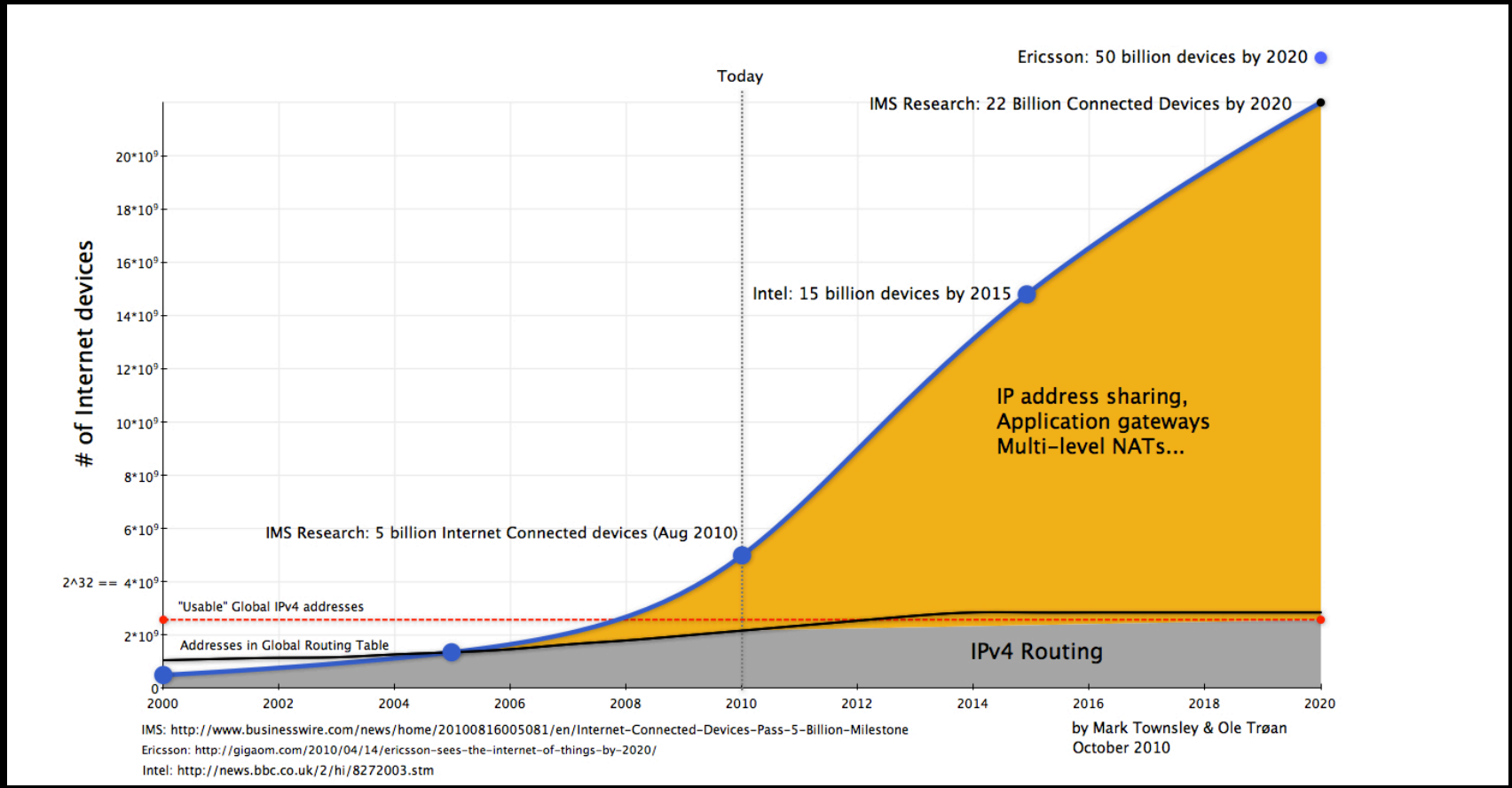
Fools ignore complexity. Pragmatists suffer it. Some can avoid it. Geniuses remove it. – Alan Perlis

Catch Phrase	The Network Is the Computer	Objects	Legacy to the Web	The Computer Is the Network	Network of Embedded Things	Network of Things
Scale	100's	1000's	Millions	10xMillions	100xMillions	Billions
When/Peak	1984/1987	1990/1993	1996/1999	1999/2003	1998/2004	2004/2007
Leaf Protocol(s)	FTP, SMTP, POP	+Gopher +WAIS	+HTTP (+JVM)	+XML Portal	+RMI	Unknown
Directory(s)	Host tables	+DNS	+LDAP(*)	+Google +UDDI	+Jini	+?
Session	RPC, XDR	+CORBA	+SSL	+SOAP, XML	+RMI/Jini	+?

Client-server
3-tier
n-tier
n-way x-connect
N-way x-connect
Fractal x-connect



We Have a Scaling Problem Too



Key System Properties

- Highly dynamic non-deterministic behavior
  - asynchronous inputs & outputs
  - causes unpredictable loads (storms)
  - load, latency, bandwidth, utilization
- Complex and unpredictable interactions
  - system extensibility, upgrades, patches, migrations,versioning
- Continuous diagnostics, fault isolation, and automatic recovery
  - In-service software and hardware upgrades
  - requires 99.999% reliability (<5 secs downtime/yr)
- End-to-End System Security
  - authentication, authorization, intrusion detection, etc.
- Dynamic Capacity Management and Power Efficiency

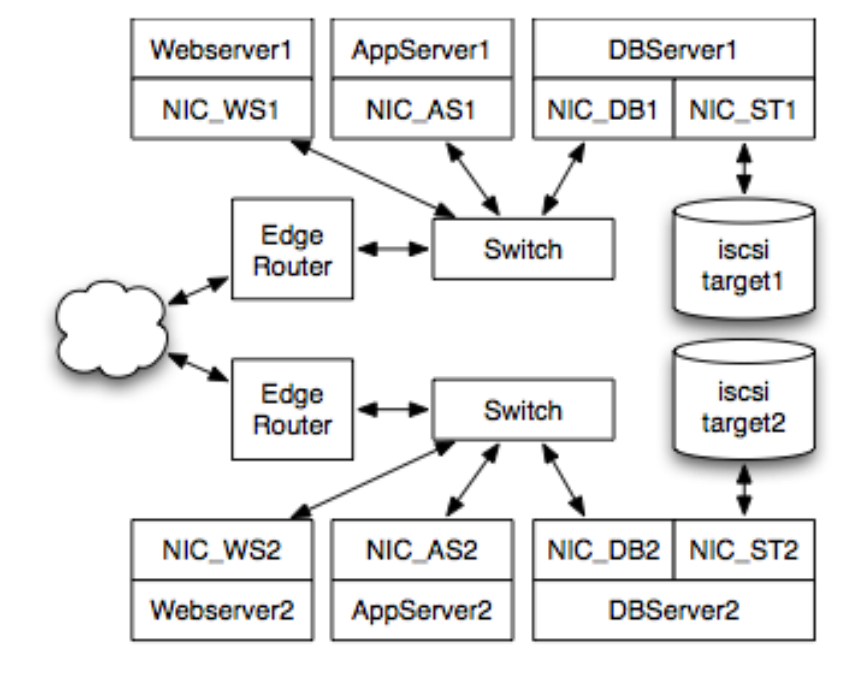
We Need Dynamic System Models

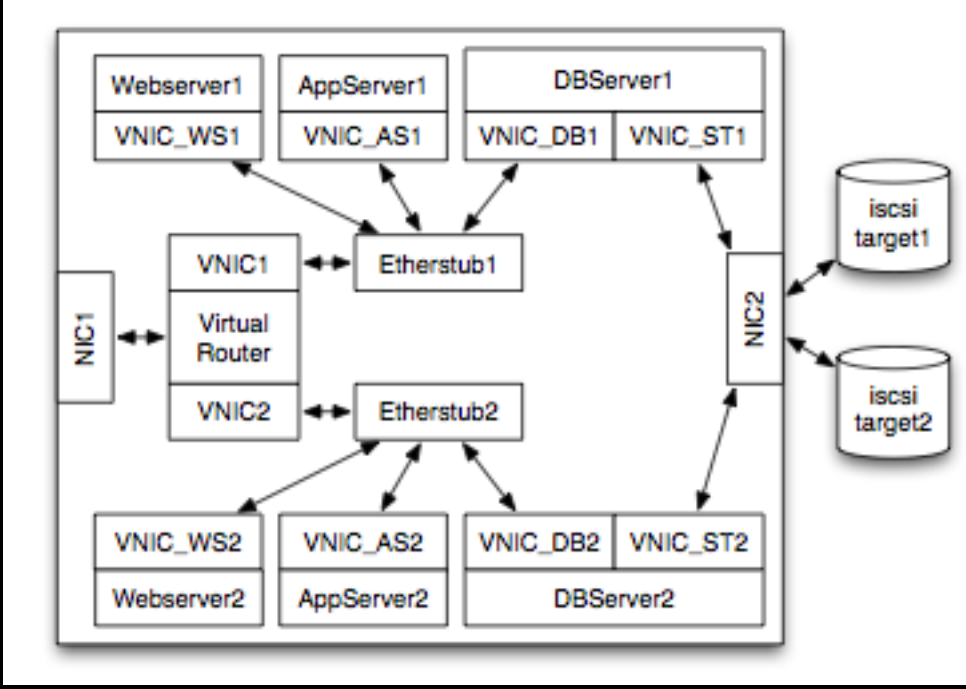
- Layered models
  - Layer 1-7 Networking, operating systems, file systems, etc.
  - have worked well for single systems
  - high degree of application diversity and complex interactions
- Algorithmic models
  - Automata, TMs, URMs, probabilistic, …
- Programming models
  - λ−calculi, operational and action semantics
- Concurrency models
  - CCS, CSP, petri nets, π-calculus, etc.
- Distributed computing models
  - publish-subscribe persistent message queuing
  - RPC
  - asynchronous message passing
  - grid/mesh models
  - …

A Challenge

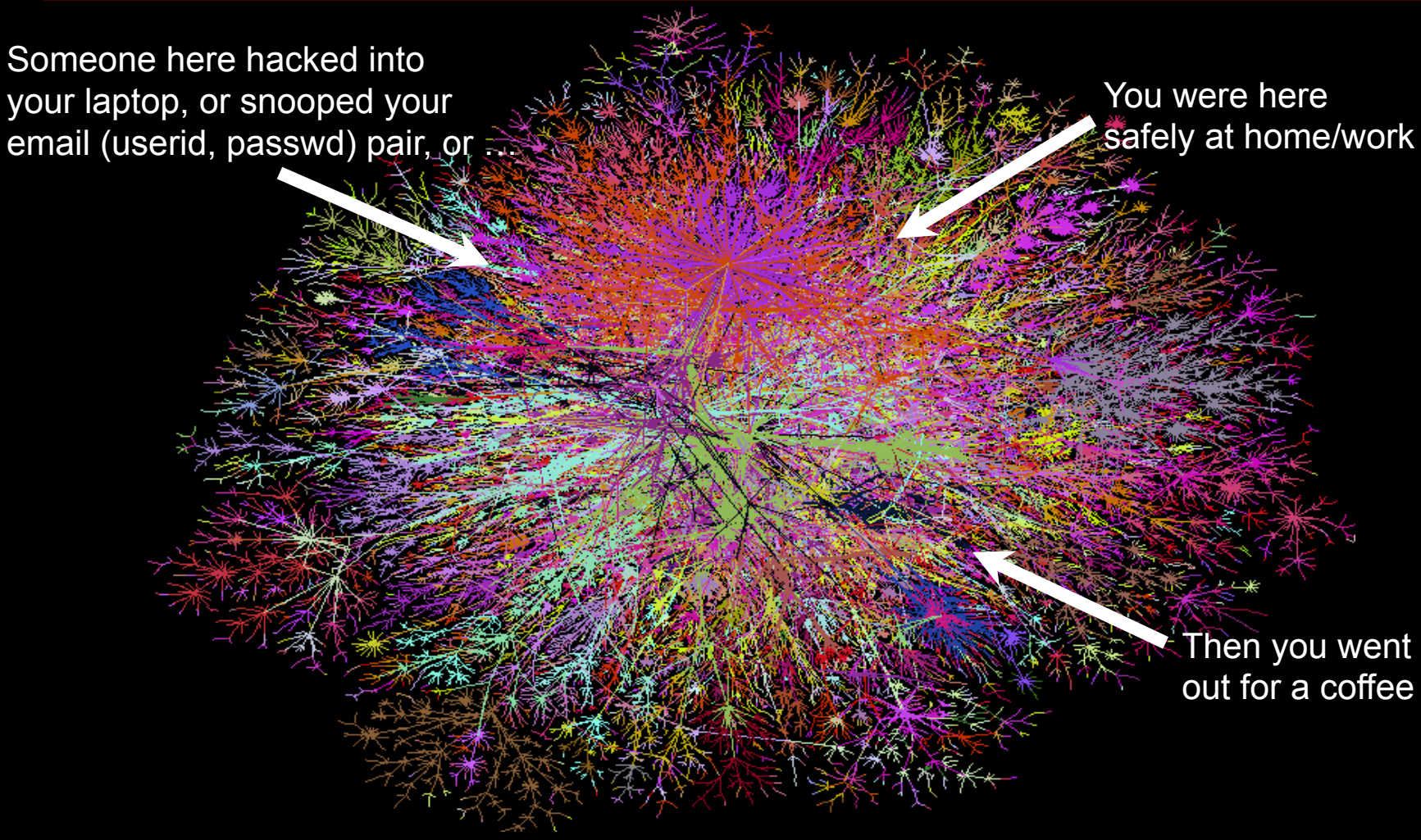
- Rethink our approach
  - new models of interaction of evolving components
  - Clear operational semantics instead of proliferation of new syntax
- Dynamic Systems Thinking
  - end-to-end system interaction models
  - Machine learning and dynamic optimization
  - Better use of simulation for capacity modelling
  - large-scale system testing infrastructure
    - this used to be what the Internet was for
    - virtual networking helps
- Systems are too rigid
  - need more dynamic adaptation
  - too much abstraction layering sometimes harmful
- Teach “systems thinking” to your students

From Physical to Virtual Networks

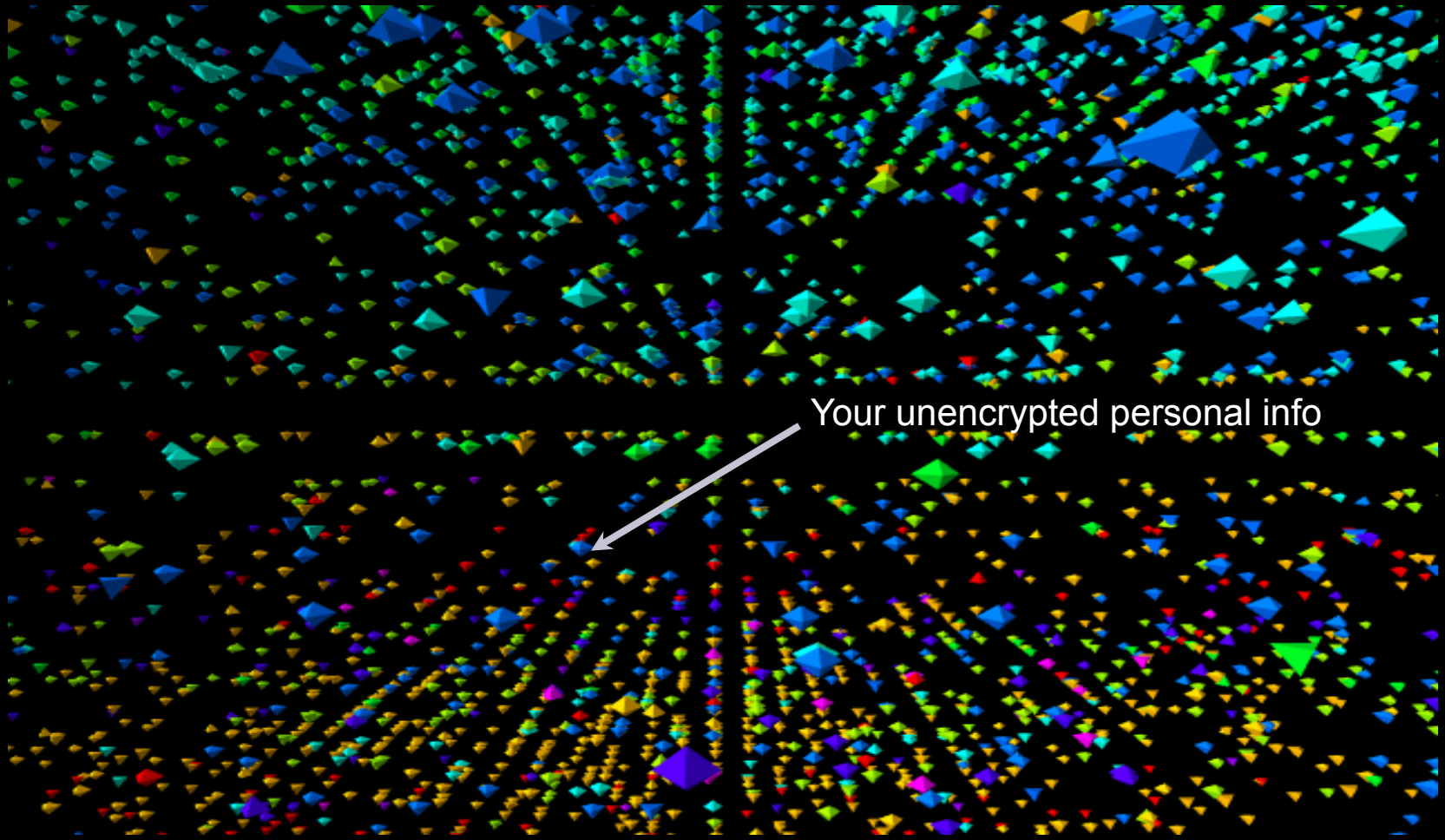




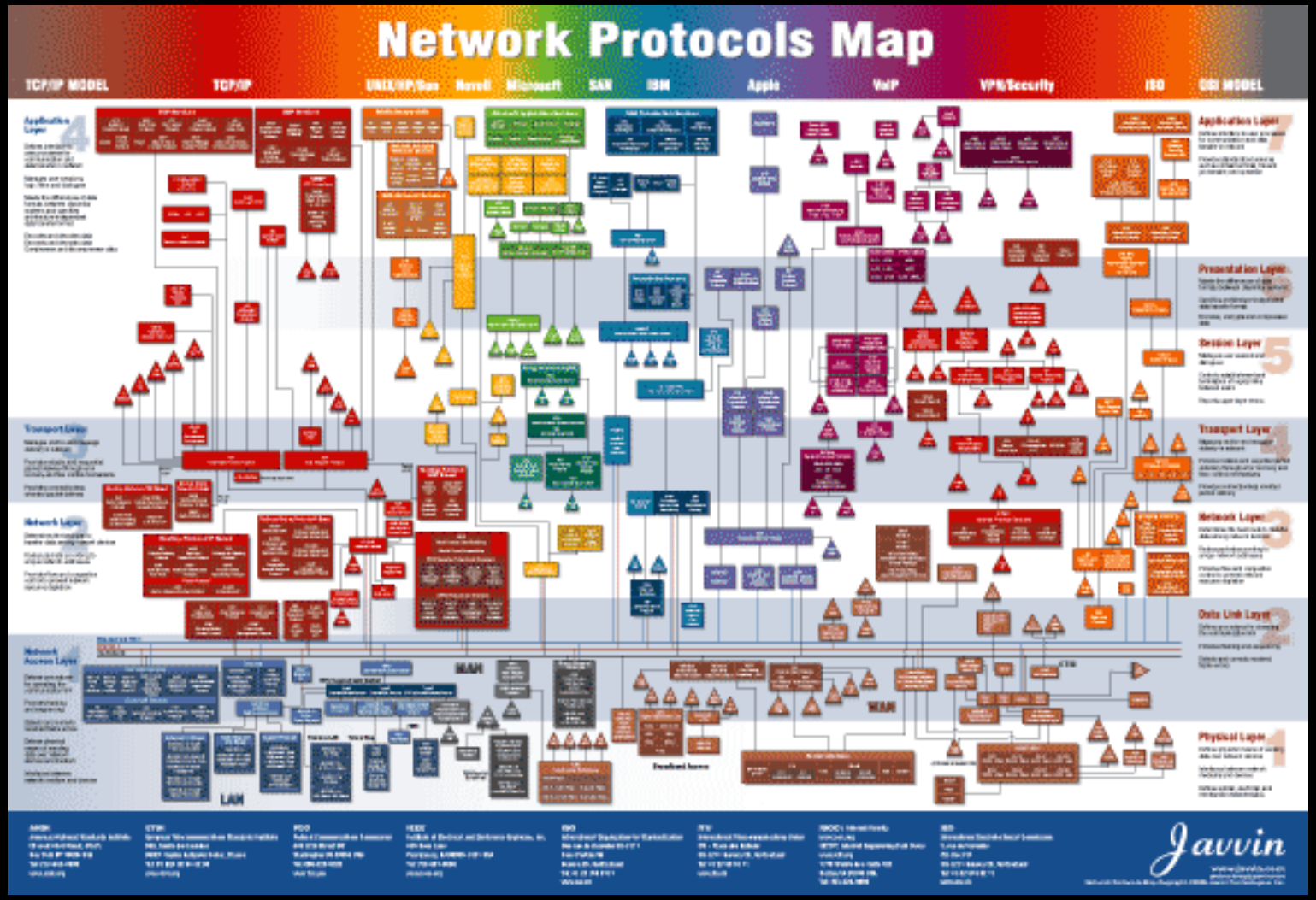
The Cloud is Expanding Rapidly



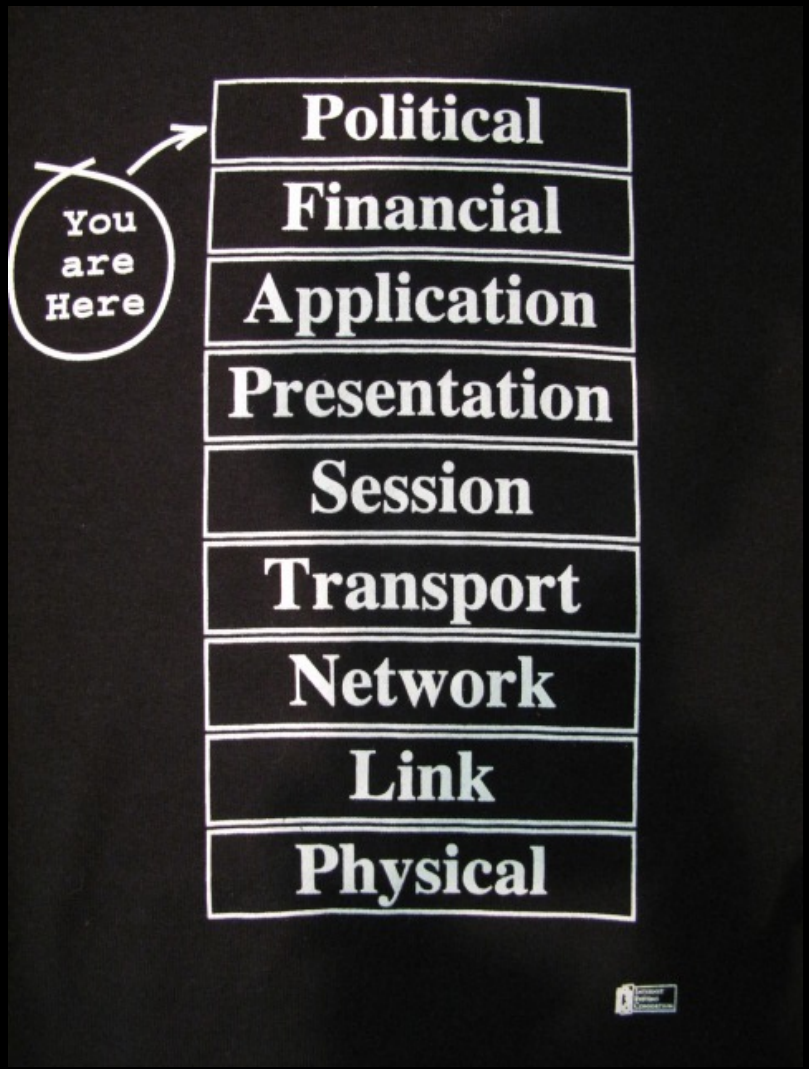
Billions and Billions of Packets



You Need a Really Good Map…



Wish I Could Help But I am Now Working at Layer 9



Questions?

The machine does not isolate us from the great problems of nature but plunges us more deeply into them

– Antoine de Saint-Exupéry (1900-1944)

Special Thanks to My Magisters

Trivia Question

- Who is this person?
  - Peter Naur
  - Typing the ALGOL 60 report
- Low Tech?
  - Quantum mechanical Turing machine poorly hidden under hair
  - Photonic magnification input device
  - Large multi-page flat-panel display
  - Wireless keyboard
  - Portable high density storage device
  - 1st generation ergonomic chair
    - not patented!