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

\[
gcd \ 0 \ 0 = \text{error}\\
gcd \ x \ y = gcd' \ (\text{abs} \ x) \ (\text{abs} \ y)\\
\text{where}\\
gcd' \ x \ 0 = x\\
gcd' \ x \ y = gcd' \ y \ (x \text{ `rem` } y)\]

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
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

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.people.fas.harvard.edu/~wellerst/atomic_patents/)
Untyped Lambda Calculus
(α-conversion, β-reduction)

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

\[ Y = \lambda f.(\lambda x.f (x x)) (\lambda x.f (x x)) \] -- 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

IBM System/370
Principles of Operation

MIL-STD-1777
RFC 791 (IP)
RFC 793 (TCP)
Virtualized User Space Network Stack (1990-1993)
Vertically Sliced Lock-Free PMs
Polymorphically Typed Composable Protocol Machines
Insights Gained from MS & PhD


- 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

Virtual Stacks

Services and Protocols

Compute Resources

CPU 1
VIRTUAL QUEUE

CPU 2
VIRTUAL QUEUE

CPU 'n'
VIRTUAL QUEUE

The MAC switches the MSI interrupt per Lane between interrupt and polling mode and controls the rate of packet arrival for the lane.

The VNICS are in the control path only. The data link layer is bypassed.

Memory Partition

Memory Partition

Memory Partition

NIC 1
Flow Classifier

NIC 2
Flow Classifier

Virtual NIC

Virtual NIC

Virtual NIC
Latest Multi-* System
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
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

<table>
<thead>
<tr>
<th>Catch Phrase</th>
<th>The Network Is the Computer</th>
<th>Objects</th>
<th>Legacy to the Web</th>
<th>The Computer Is the Network</th>
<th>Network of Embedded Things</th>
<th>Network of Things</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>100’s</td>
<td>1000’s</td>
<td>Millions</td>
<td>10xMillions</td>
<td>100xMillions</td>
<td>Billions</td>
</tr>
<tr>
<td>Leaf Protocol(s)</td>
<td>FTP, SMTP, POP</td>
<td>+Gopher +WAIS</td>
<td>+HTTP (+JVM)</td>
<td>+XML Portal</td>
<td>+RMI</td>
<td>Unknown</td>
</tr>
<tr>
<td>Directory(s)</td>
<td>Host tables</td>
<td>+DNS</td>
<td>+LDAP(*)</td>
<td>+Google +UDDI</td>
<td>+Jini</td>
<td>+?</td>
</tr>
<tr>
<td>Session</td>
<td>RPC, XDR</td>
<td>+CORBA</td>
<td>+SSL</td>
<td>+SOAP, XML</td>
<td>+RMI/Jini</td>
<td>+?</td>
</tr>
</tbody>
</table>

Client-server
3-tier
n-tier
n-way x-connect
N-way x-connect
Fractal x-connect
We Have a Scaling Problem Too

- IMS Research: 5 billion Internet Connected devices (Aug 2010)
- IMS: 22 Billion Connected Devices by 2020
- Ericsson: 50 billion devices by 2020
- Intel: 15 billion devices by 2015

IP address sharing,
Application gateways
Multi-level NATs...

Addresses in Global Routing Table
"Usable" Global IPv4 addresses
IPv4 Routing

References:
Intel: http://news.bbc.co.uk/2/hi/8272003.stm
by Mark Townsley & Ole Traan
October 2010
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**
  - $\lambda$-calculi, operational and action semantics

- **Concurrency models**
  - CCS, CSP, petri nets, $\pi$-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

Someone here hacked into your laptop, or snooped your email (userid, passwd) pair, or …

You were here safely at home/work

Then you went out for a coffee
Billions and Billions of Packets

Your unencrypted personal info
You Need a Really Good Map...
Wish I Could Help But I am Now Working at Layer 9
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!