



Rethinking Time in Distributed Systems

Can We Build Complex Systems Simply?

Paul Borrill

Stanford EE Computer Systems Colloquium

4:15PM, Wednesday, Nov 11, 2009

NEC Auditorium, Gates Computer Science Building B03

<http://ee380.stanford.edu>

Agenda

- **Motivation: Conquering Complexity**
 - *Complexity destroys Availability, Scaling & Recovery*
- **Computer Scientists Notion of Time?**
 - *History & State of Affairs*
- **The World is Massively Asynchronous**
 - *Humans think Sequentially; Programmers are Humans*
- **The Need to Rethink Distributed Systems**
 - *Why we need a new Theory of Infrastructure (TOI)*
- **A Way Forward?**
 - *Re-examine the Axiom of Time in Computer Science*

Part I: Motivation

January 28, 1986



September 11, 2001



Failures, Disasters, Attacks:
may be inevitable ...

*But what's stopping our
systems from recovering
quickly?*

The Computer Industry 2009

- **The processor industry:** In a *Concurrency Crisis*
 - Gets worse with each generation of processor (the number of cores doubles each generation instead of the performance of each core)
 - *No-one has thought about the software (John Hennessy)*
- **The storage industry:** In a *Complexity Crisis*
 - We “have to” scale-out, because “scale-up” systems are impossible to make sufficiently reliable
 - *No-one has thought about the software (Paul Borrill)*

Might there be a common cause behind these problems?

Complexity Crisis in IT

Complexity is holding our industry back ... a lot of what is bought and paid for doesn't get implemented because of complexity. Ray Lane, ex COO, Oracle

This is the sort of industry trauma that precedes a major paradigm shift. Pip Coburn, UBS Warburg.

IT has lost control over systems: the state of the art really is reactive firefighting. Harrick Vin, Professor, Computer Science University of Texas at Austin.

IT complexity acts as a significant tax on IT value. Bob Zukis, Pricewaterhouse Coopers.

Complexity leads to brittleness and high costs. Frank Modruson, CIO, Accenture

CIOs are standing in the path of a fire hose, spewing complexity, and many are getting soaked. Salen Granian, CIO

The problem is that we keep going through the same problems over and over and over again. Peter Neumann SRI

The IT industry is long past denial. IT today is in a state that we should be ashamed of; it's embarrassing. Greg Papadopoulos, CTO, Sun Microsystems.

That's why they've failed in the past; they don't look at the lifecycle; all of a sudden, they hit the [complexity] wall. Robin Ford, CIO, Wal-Mart

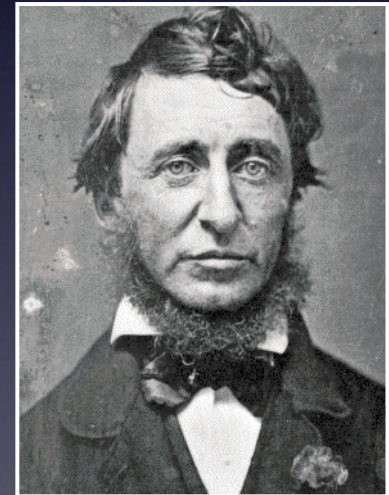
Every IT manager, system administrator, and developer is fighting against the monster of computing complexity. Mike Shapiro, Sun Microsystems

Complexity has arisen from evolution, of: Operating systems, apps, workload types, volume & users. Systems must constantly adapt to changes. Harro Vin, Prof, CS, UT

Each additional technology wasn't a tipping point; it was just one more thing. But at some point you realize you've reached a tipping point. It's easy to end up with unnecessary complexity due to technological and business-process diversity. Ray Dury, CIO, Fifth Third Bancorp

Everyone knows about it, why doesn't someone do something about it?

“Men have
become tools of
their tools”



Henry David Thoreau

The Perfect Storm

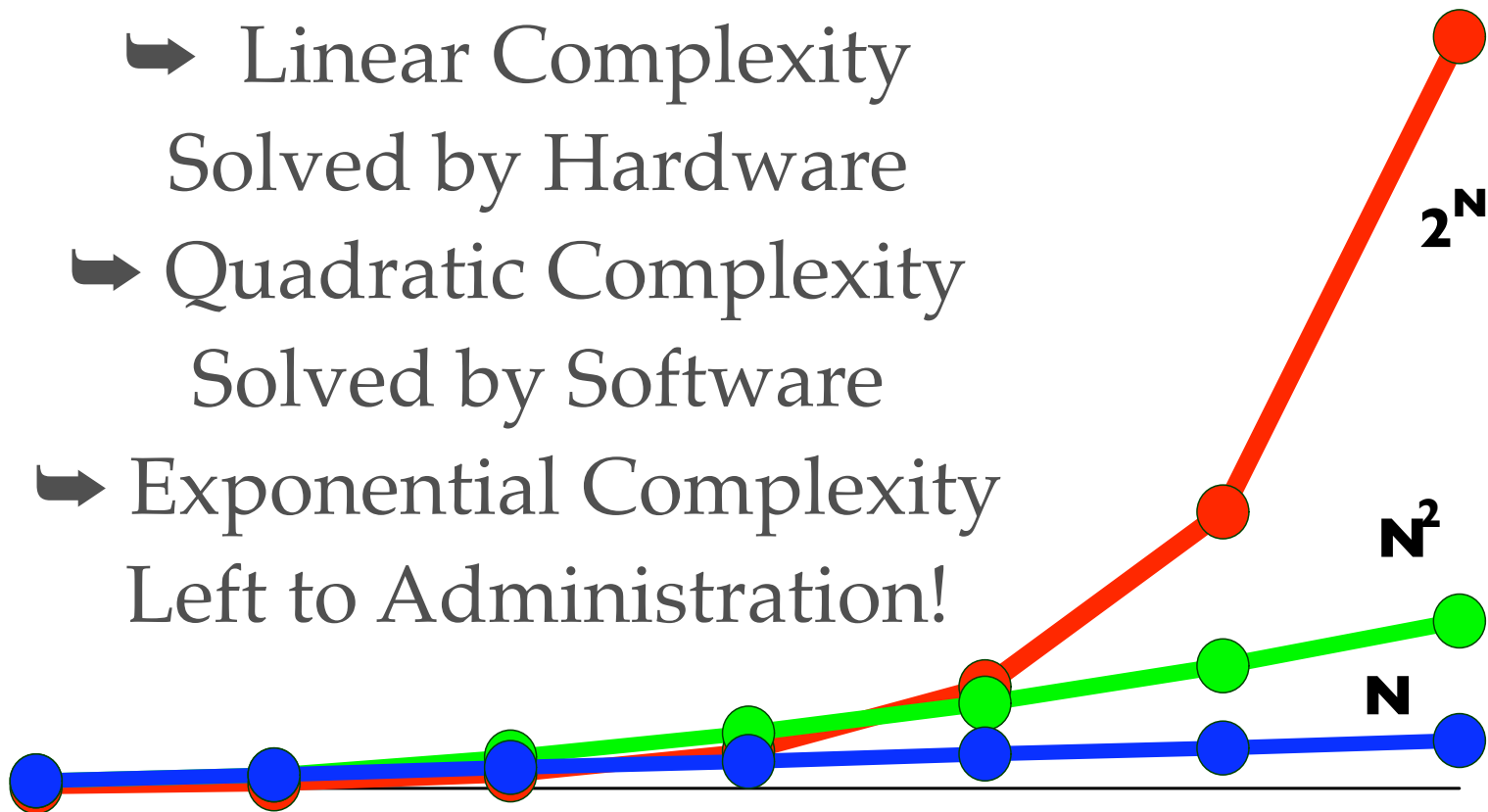
All happening at once:

- IT Complexity Exponentiating
- Baby Boomer (Experts) Retiring
- Fewer IT Graduates
- Fewer IT Immigrants
- Economic Downturn



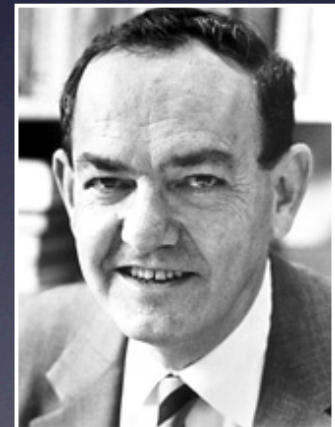
Complexity Scaling

- ➔ Linear Complexity
Solved by Hardware
- ➔ Quadratic Complexity
Solved by Software
- ➔ Exponential Complexity
Left to Administration!



“What information consumes is rather obvious: it consumes the attention of its recipients.

Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it”



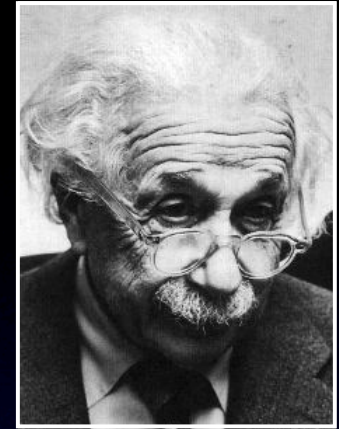
Herbert Simon

Part II Time & Causality

- Simultaneity is a Myth
- Causality is a Myth
- Time is not Continuous
- Time does not flow
- Time has no direction

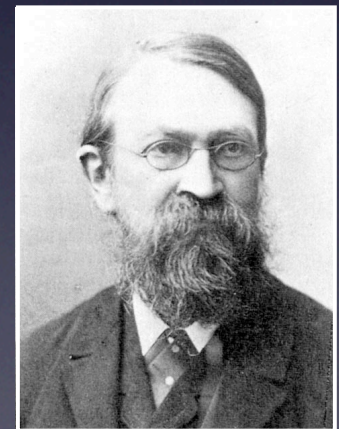
What is Time?

“A persistently stubborn illusion”



Albert Einstein

“It is utterly beyond our power to measure the changes of things by time” Quite the contrary, time is an abstraction at which we arrive through the changes of things”



Ernst Mach

Computer Scientists & Time ?

- A relationship with time is intrinsic to everything we do in computing, modifying and moving data
- The understanding of the concept of time among computer scientists appears far behind that of physicists and philosophers
- *If* fundamental flaws exist in the time axioms underlying the algorithms that govern access to and evolution of our data, then our systems will fail in unpredictable ways, and any number of undesirable characteristics may follow

Causality is a Myth

“Computer Scientists imagine that causation is one of the fundamental axioms or postulates of physics, yet, oddly enough, in real scientific disciplines such as special and general relativity, and quantum mechanics, the word “cause” never occurs. To me it seems that computer science ought not to assume such legislative functions, and that the reason why physics has ceased to look for causes is that in fact there are no such things. The law of causality, I believe, like much that passes muster among computer scientists, is a relic of a bygone age, surviving, like a belief in God, only because it is erroneously supposed to do no harm”

~Paul Borrill (with apologies to Bertrand Russell)

What do I mean by God?

- It has nothing to do with religion:
- **GEV**=God's Eye View: the way programmers think.
c.f. Client-server, Linearizability
- **LOV**=Local Observer View: the way the world really works

*Related to the now obsolete context of “observer and system” in Quantum Theory. Otherwise known as event symmetry: we are all mutual observers**

* See: The internal description of a causal set: What the universe looks like from the inside. Fotini Markopoulou

Determinism is a Myth

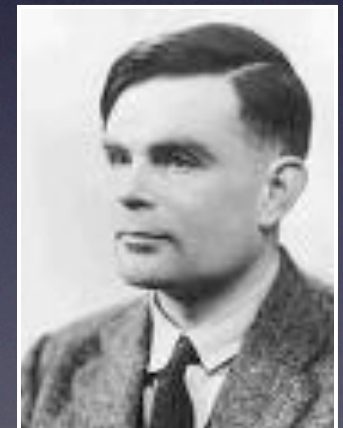
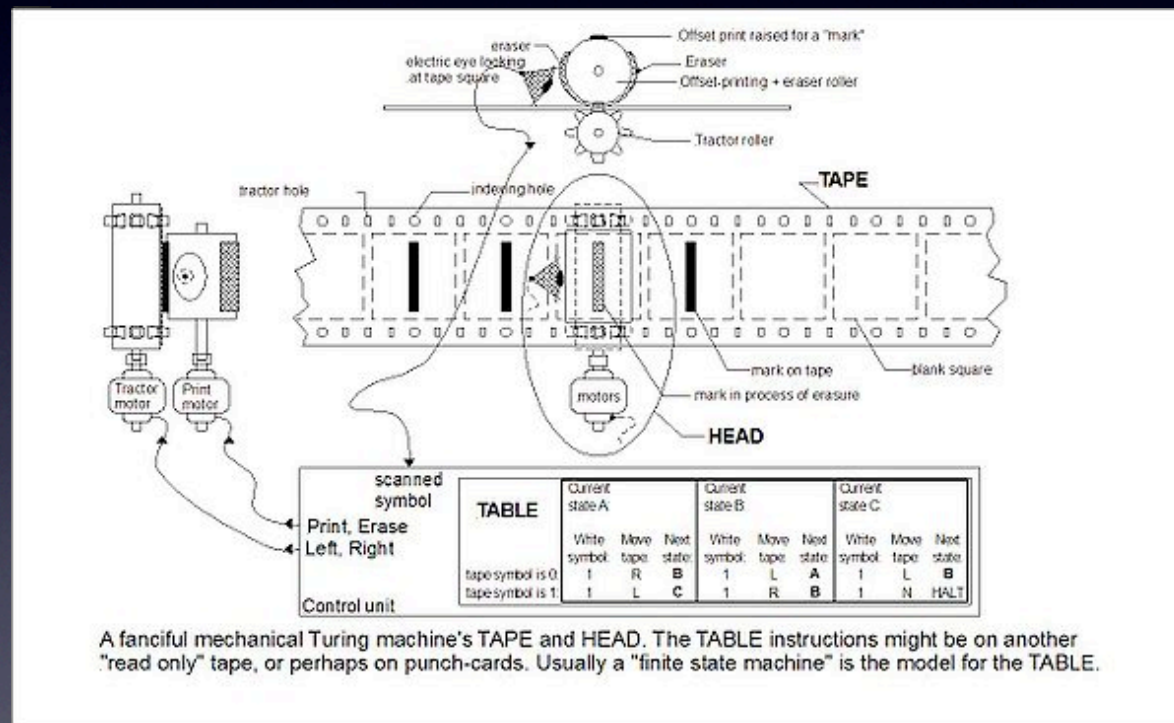
- *Distributed systems are fundamentally non-deterministic*
- *From the metastability of our logic gates to the decoherence of our communication links*
- *The problem goes deep into physics, and ultimately to the question of free will (FWF)**

*If we have free will, then so do elementary particles**

* See: John Conway & Simon Kochen. The Free Will Theorem.

Where do computer scientists get their notion of time?

A mathematically seductive idealization of a one dimensional object in \mathbb{R}



Alan Turing

Simultaneity is a Myth

Maurice Herlihy and Nir Shavit:

The Art of Multiprocessor Programming [2008]:

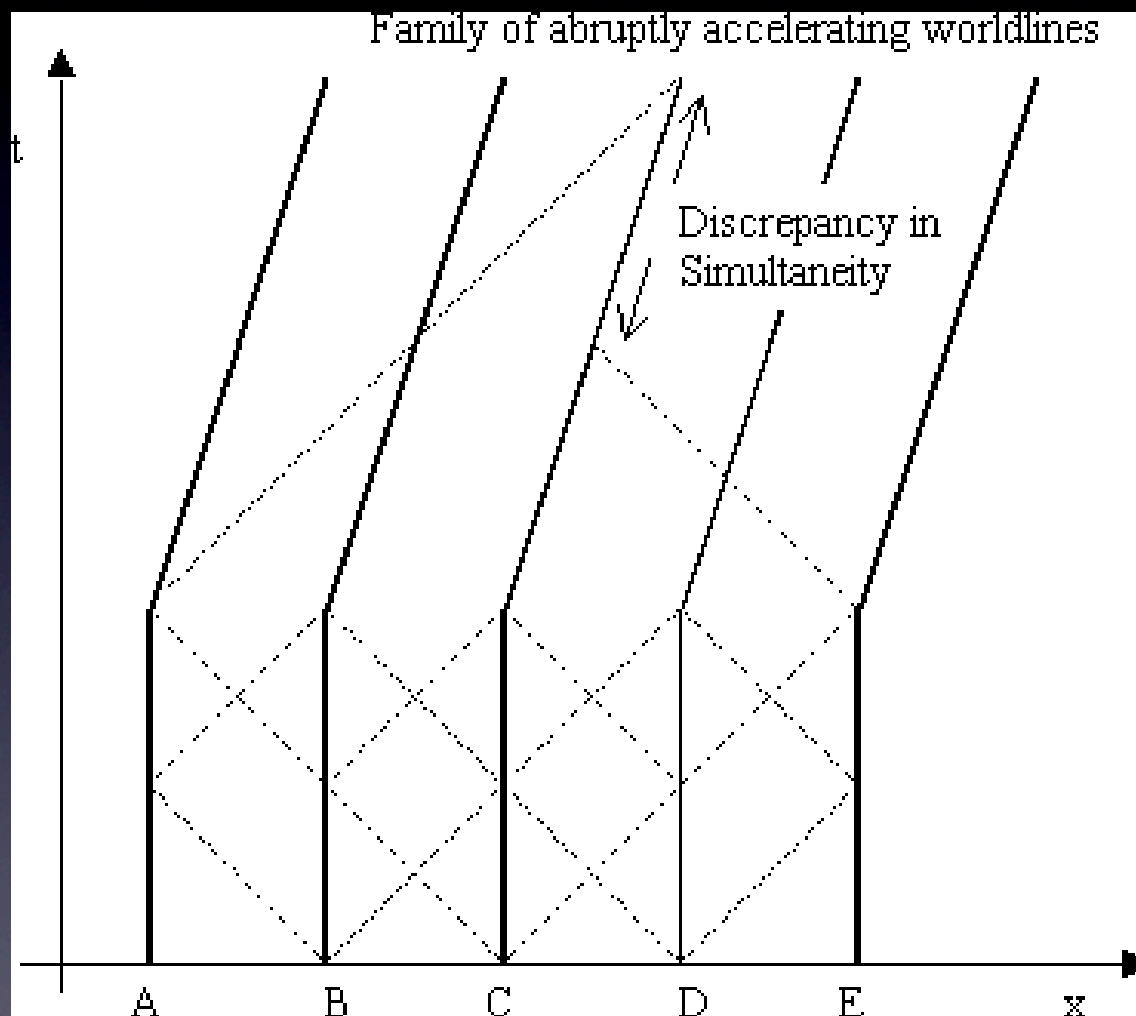
*"In 1689, Isaac Newton stated 'absolute, true and mathematical time, of itself and from its own nature, flows equably without relation to anything external.'
We endorse his notion of time"*

A notion of time proven incorrect over a hundred years ago ...

Simultaneity is a Myth

- In 1905 Einstein showed us that the concept of “now” is meaningless except for events occurring “here”
- In 1978, Leslie Lamport published “Time, Clocks and the Ordering of Events”, in which he defined the *happened before* relation
- Unfortunately, *happened before* is meaningless unless intimately associated with *happened where*. Lamport understood this, but many who read his paper don't
- In 2009, most Computer Scientists and programmers implicitly base their algorithms on absolute (Newtonian) Time, or use Lamport's timestamps as a crutch to sweep their issues with time under the rug

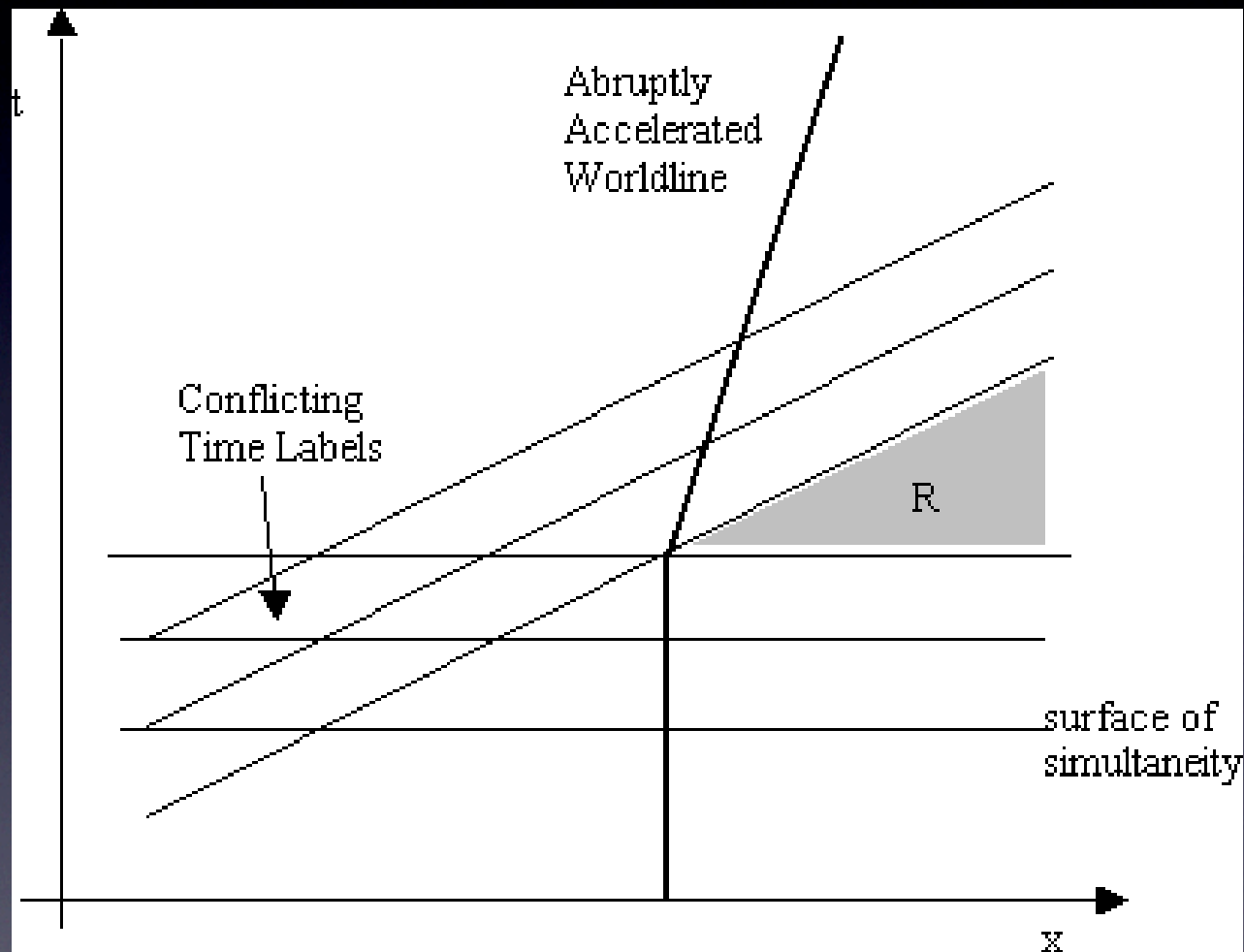
Breakdown in Simultaneity - I



Courtesy Kevin Brown

<http://www.mathpages.com/rr/s4-08/4-08.htm>

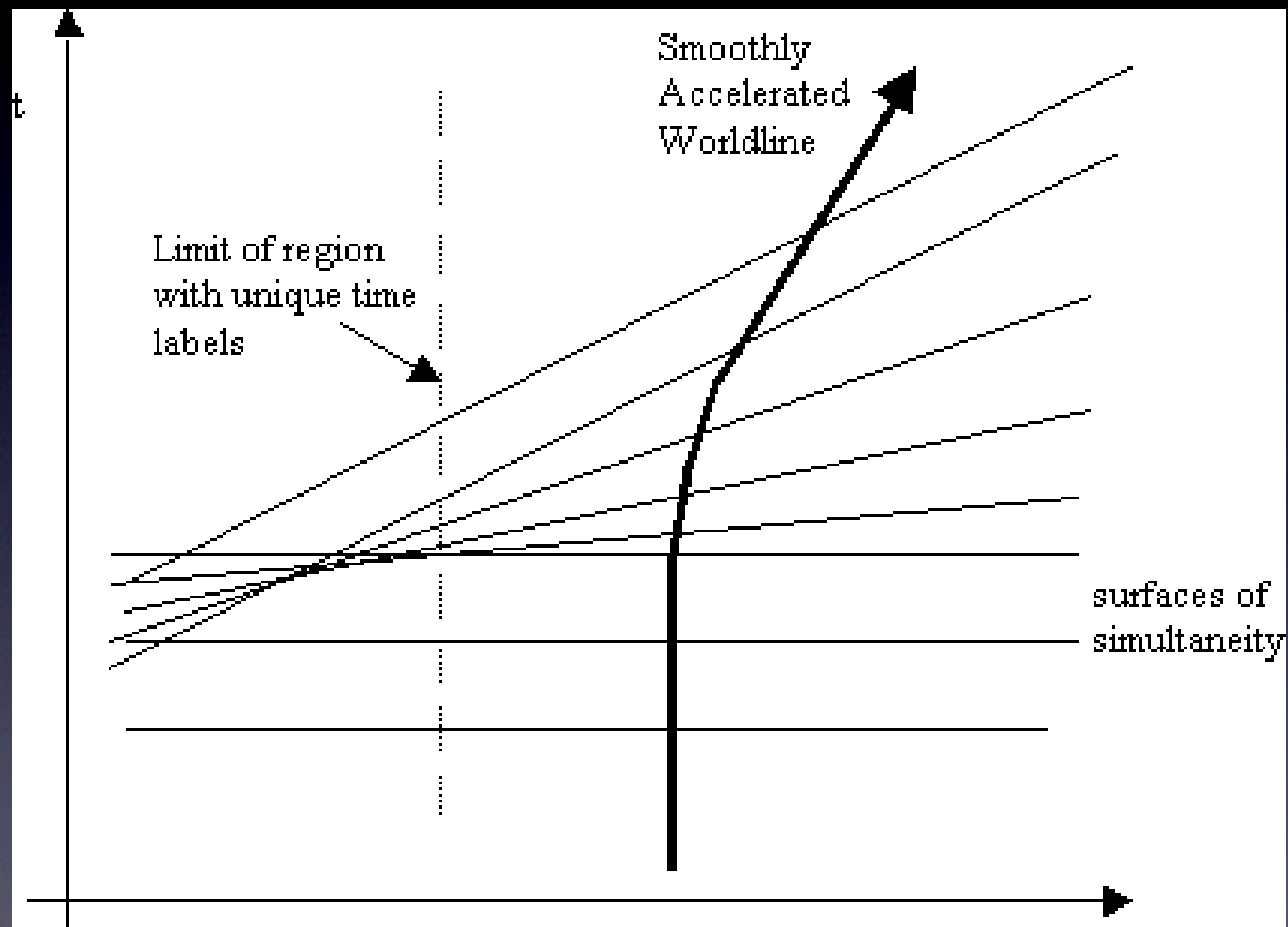
Breakdown in Simultaneity - 2



Courtesy Kevin Brown

<http://www.mathpages.com/rr/s4-08/4-08.htm>

Breakdown in Simultaneity - 3



Courtesy Kevin Brown

<http://www.mathpages.com/rr/s4-08/4-08.htm>

But wait - can't we assume an “inertial system”?

- **Our computers reside:**

- *On the Surface of a Rotating Sphere*
- *In a Gravitational Field*
- *Orbiting a Star*

- **Our Computers are connected:**

- Not with light signals in a vacuum, but with a stochastic latency distribution network
- Equivalence of Acceleration and variability of transmission delay in the propagation of packets
- Creating coherent time sources is “problematic”



Other difficulties with “time”

● Time is not a continuous background*

- Time is change. Events are unique in *spacetime*. There is no such thing as an indivisible instant.

● Time does not flow

- There is no more evidence for the existence of anything real between one event and another, than there is for an *aether* to support the propagation of electromagnetic waves through empty space

● In Physics, Time has no direction

- Time is intrinsically symmetric. We experience irreversible processes that capture “change” like a probability ratchet that prevents a wheel going backwards (the 2nd law)

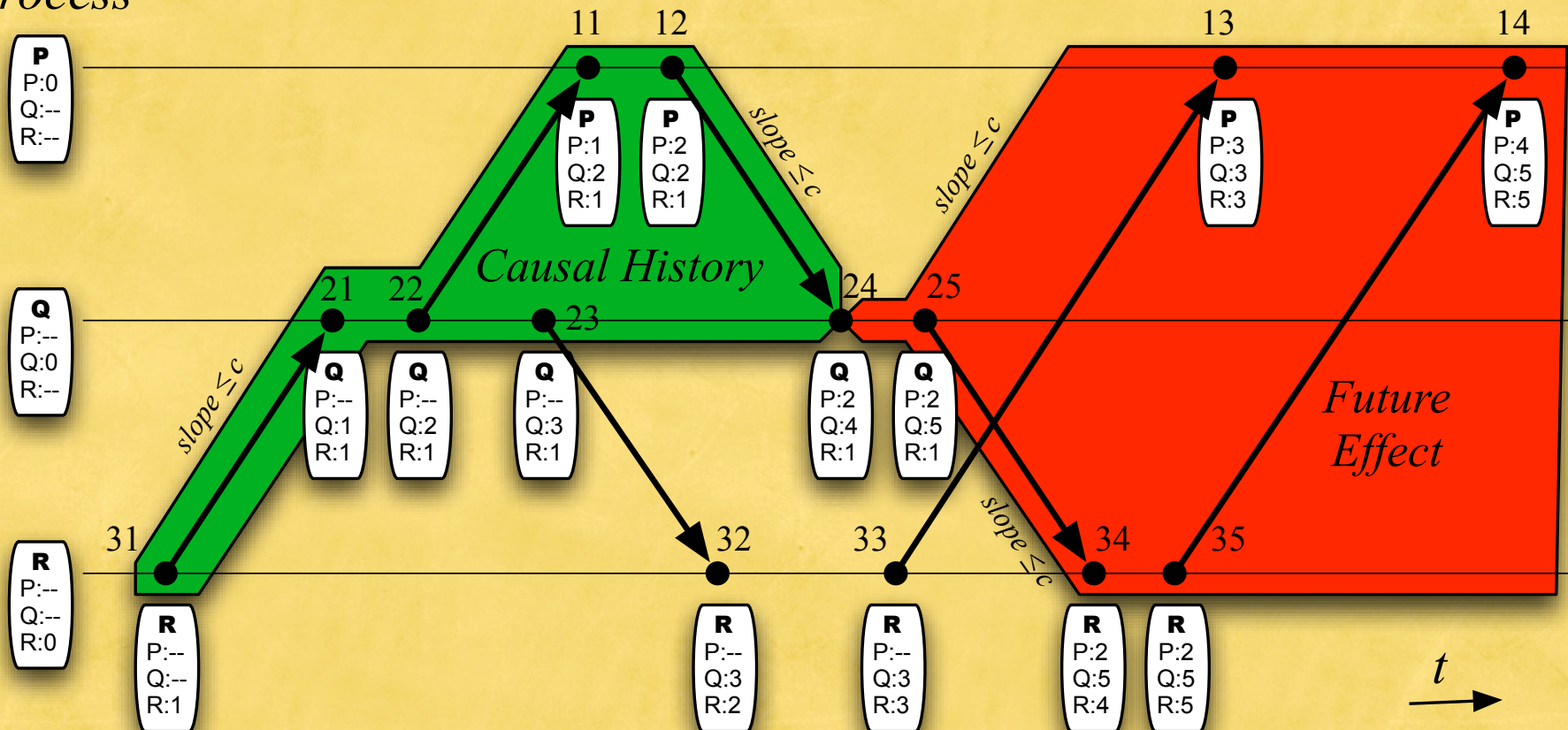
* See: The Case for Background Independence, Lee Smolin [ArXiv:hep-th/0507235v1 25-Jul 2005]

Leslie Lamport 1978

- Defined “**happened before**” *relation*: a **partial order**
- Defined “**logical timestamps**” which force an *arbitrary total order*, restricting the available concurrency of a system (i.e. the algorithm can proceed no faster than it would in a single processor)
- This “concurrency efficiency loss” gets worse as:
 - We add more nodes to a distributed system
 - These nodes become more spatially separated
 - Our processors and networks get faster
 - Our processors are comprised of more cores

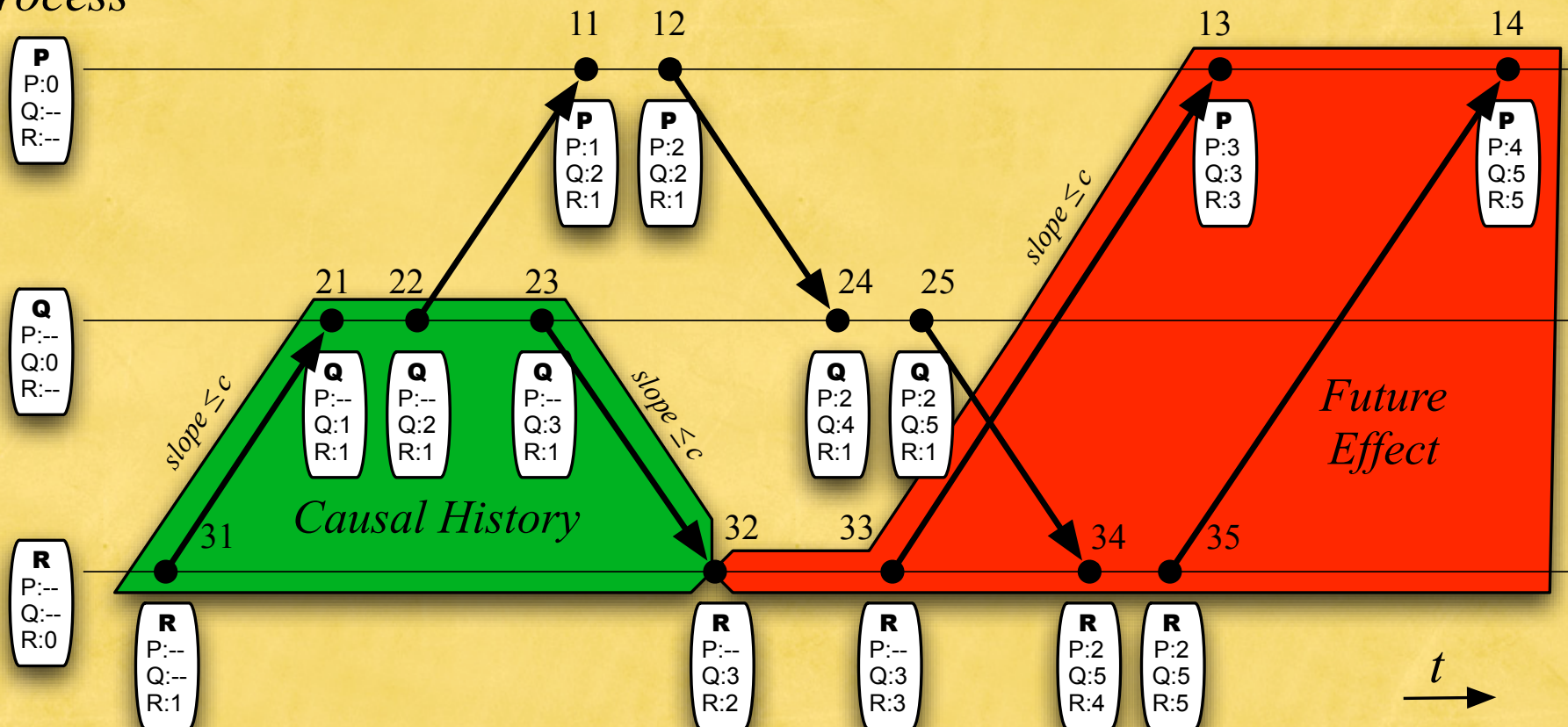
Time-stamps - Event 24

Process

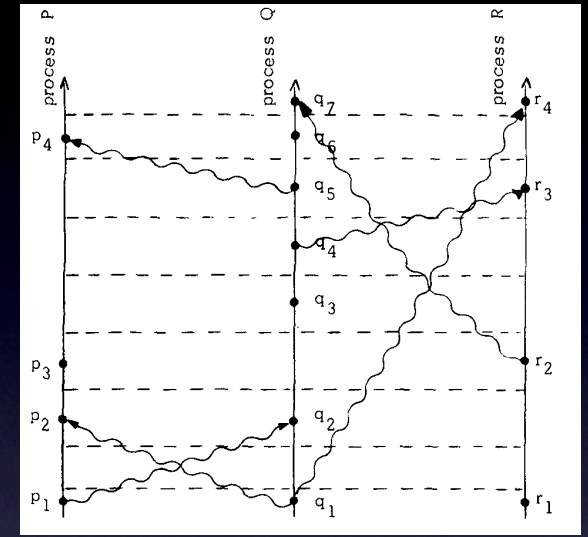
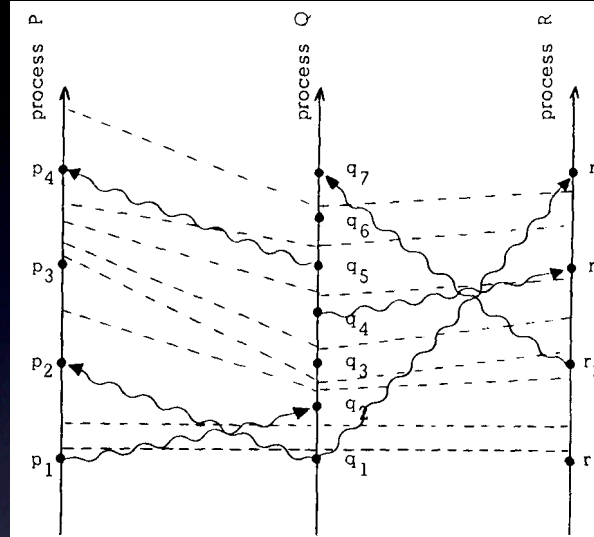
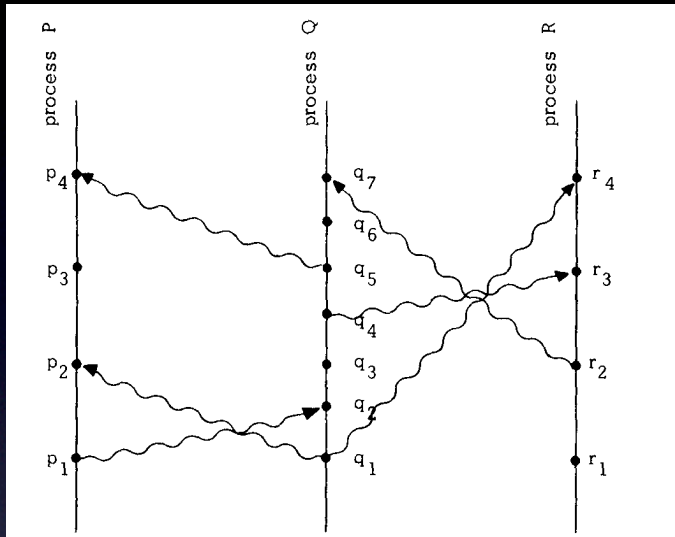


Time-stamps - Event 32

Process



Lamport



The total ordering defined by the algorithm is somewhat arbitrary. It can produce anomalous behavior if it disagrees with the ordering perceived by the system's users. This can be prevented by the use of properly synchronized physical clocks. Our theorem showed how closely the clocks can be synchronized.

More Lamport

- “We can consider the tick lines to be the time coordinate lines of some Cartesian coordinate system on space-time”
- “Processes and messages are still represented by lines but tick lines become two dimensional surfaces”
- This is barely consistent with SR, it is not consistent with GR
- More importantly, there is no “Cartesian coordinate system” in current physics. if anything, we use relative (invariant) coordinate systems in GR
- ***Everything is relative, even when you think it isn't***

Part III Can we Build Complex Systems Simply?

- Continuing to build storage systems the way we do is *no longer a viable strategy*
- System requirements are inherently conflicting, diverse *and are often unknowable*
- Large systems are not designed, they are evolved, *and they fail constantly*
- We have no choice but to embrace *Commodity Reliability And Practices to make the economics work*
- We cannot employ a *God's-Eye View*, or *centralized control*, *if we expect our systems to scale*
- We must trade off abundant resources for scarce resources; *skilled people are the scarce resource*

A New Theory of Infrastructure?

- We need a cure; not an endless overlay of band-aids that mask failed architectural theories
- **The Curse of the God's Eye View (GEV)**
 - Time & Causality
 - Identity & Individuality
 - Persistence & Change
- These problems are not adequately appreciated in the computer science literature
- *GEV designers don't relieve us of complexity - they cause it!*

Approaches

- **Decentralize everything**: no more master nodes, master copies, master programs; to make **scalability possible**
- Employ **adaptive architectures which grow and evolve** as the organization's needs and challenges develop
- Build systems out of **independent, autonomous units** that can be cloned, distributed, and replaced at will
- **Eliminate diversity**: make elements **substitutable**, standardize hardware, software, configurations
- **Require systems to manage their own** configuration, healing, provisioning and migration, don't make slaves out of humans

A Theory of Exchanged Quantities (EQ)

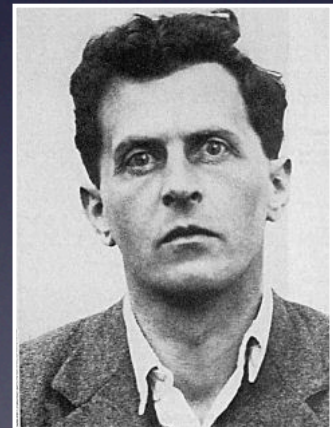
- Every interaction exchanges specified quantities
- Quantities may be conserved (e.g. locks, money transactions, minimum numbers of replicas)
- EQ overcomes many of the problems of time and causality, allowing all events to be processed between nodes rather than attempt to recreate a GEV for time or control
- Conserved quantities can be recovered (e.g. locks, lost replicas) and audited (e.g. money transactions)
- Corresponds with “safe assumptions” regarding time from physics and philosophy

Wittgenstein asked a friend: why do people always say it was natural for man to assume that the Sun went around the Earth rather than that the Earth was rotating?" His friend replied, "Well obviously because it just looks as though the Sun is going around the Earth!

Wittgenstein replied ... "Well what would it have looked like if it had looked as though the Earth was rotating?"



Richard Dawkins

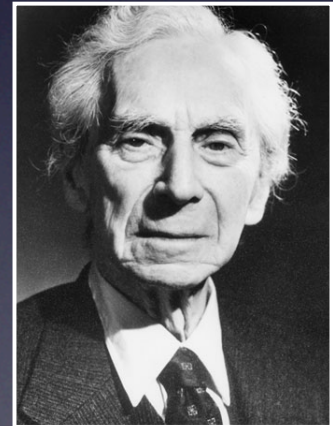


Ludwig Wittgenstein

God's Eye View

- So what would it look like, if we were not able to reach out as designers, and to be unable to directly control things as the number, connectivity and diversity of things scale?
- Maybe it would look exactly like what we are experiencing: the out of control complexity *robbing us of our productivity, preventing us being able to scale our systems, to exploit the available concurrency, or recover from perturbations*

“The ultimate goal of machine production – from which, it is true, we are as yet far removed – is a system in which everything uninteresting is done by machines and human beings are reserved for work involving variety and initiative”



Bertrand Russell



Thank You

Simplicity is the ultimate sophistication

~ Leonardo da Vinci

paul@replicus.com