Three sides of a coin:
Why it's easy/impossible/tricky to get your ideas into routers

PFLDnet Feb. 2007

Larry Dunn

Manager, Advanced Architecture

ldunn@cisco.com

Agenda

- Problem Statement
- Larry's personal experience
- What others say
- Implications for PFLDnet
- Tiny Cisco research plug
- Summary

Disclaimer: these are my thoughts/opinions; certainly *not* any sort of "Cisco position"

Problem Statement

- **“I’ve got a great idea/method for TCP xxx (or QoS method yyy, or security tool zzz, or routing protocol rrr).**

  What does it take to this idea into routers?”

Larry’s Personal Experience

Personal Experience… XCP

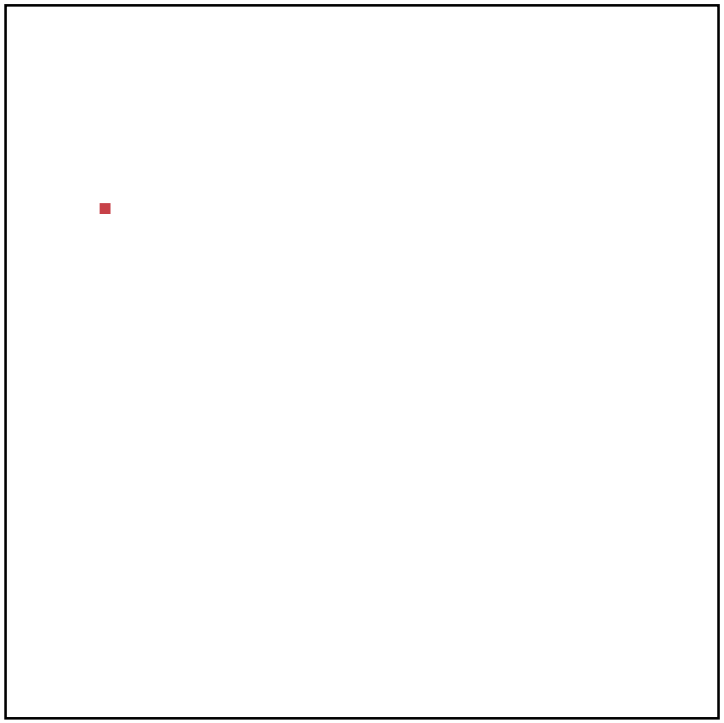
- XCP - after Dina's talk at SigComm 2002
- Checked w/ some h/w, s/w folks on "what would it take" a couple years out…
- Varied responses

  S/W folks: "real inline multiplies and divides tells me that the authors have not thought hard enough yet"

  H/W folks: brute force cell-based 8x16 multiplier: 128 AND for partial products, 175 full-adders; in .13um technology, that's 7 um^2 per AND, 36 um^2 per adder, 0.007 mm^2 total. And given typical propagation times, cold run ~67Gpps. But make sure this is *the* winner…

XCP(2)

- So…. How big is 0.007 mm^2 (.083mm x .083mm)?
- Typical large die might be 15x15..20x20 = 225..400 mm^2
- So that's 0.0017% of available area
- (OK, the red dot is really much too big…)



XCP (3)

- A while later, TechCenter folks connected w/ Aaron Falk & co. doing XCP implementation
- When I checked, neither “side” was happy with the other
- Researchers: Cisco folks not responsive, iterative (paraphrase)
- Cisco folks: ground-rules keep changing- not what they signed up for
- Lesson: mis-matched expectations, cultures

What Others Say

What other say… on Internet Lessons Learned

- From Doug Comer talk on “Lessons Learned from the Internet Project”
- Cerf/Kahn paper proposing packet-gateways, versus (phone-system style) application gateways which require a new box in the middle for every application
- For maximum flexibility, make fewest possible assumptions about underlying nets
- Location of intelligence - intelligent end-systems, simple core is best for innovation, but more vulnerable

What others say… on inevitability

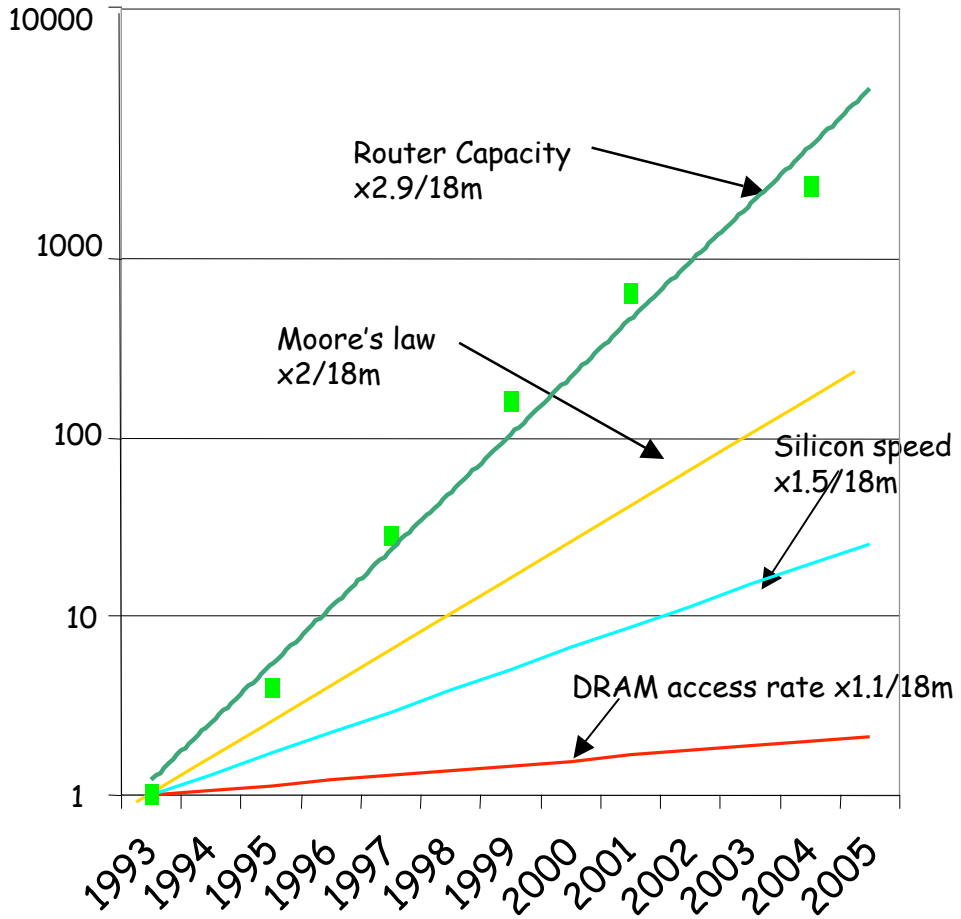
- **Well-respected Stanford Prof.**
- **Future Congestion Control will involve routers - it's inevitable**

What others say… on what's important

- Senior h/w designer at router-vendor
- Power considerations are increasingly important
- System requirements scaling faster than both silicon and cooling technology
- Approaching limits in total system power, individual device power, cooling infrastructure
- Buffer memory <10% of power; ASICs ~50%
- So what?
  Provide perspective on what's on designers' minds.
  Not "save 20% memory with new address-lookup variant."

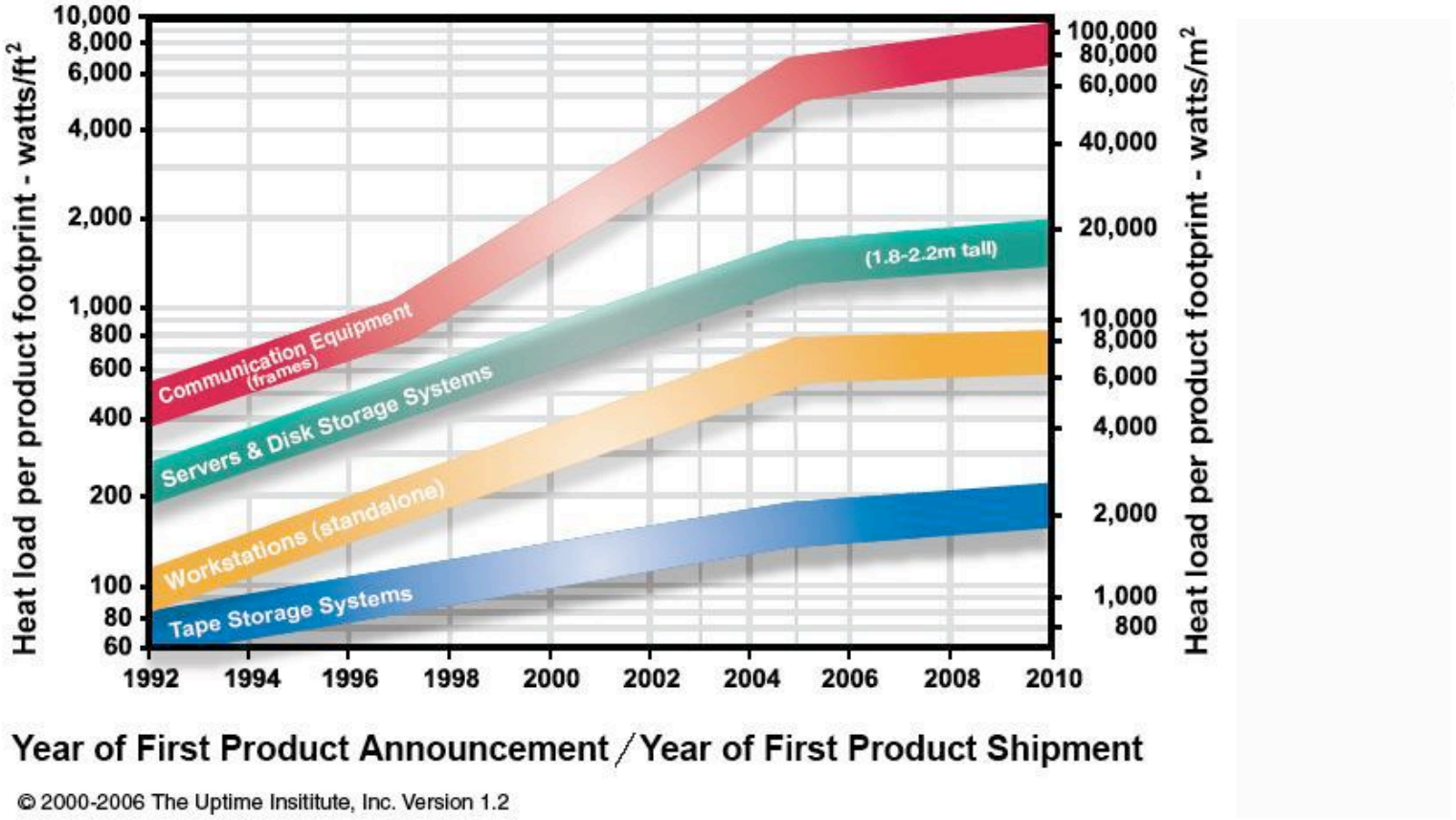
What's important(2)… growth/power

- **Router capacity growing faster than component growth, and faster than Moore's law**



What's important(3)… power density

Figure 1: 2000-2010 Product Heat Density Trends Chart



Data from Thermal Management Consortium (includes Cisco)

What's important(4)… power density/facilities

What others say… on technology trends

- Ray Kurzweil - on exponential trends everywhere
- See www.kurzweilAI.net/pps/sc06

  And “The Singularity Is Near: When Humans Transcend Biology”

- Themes

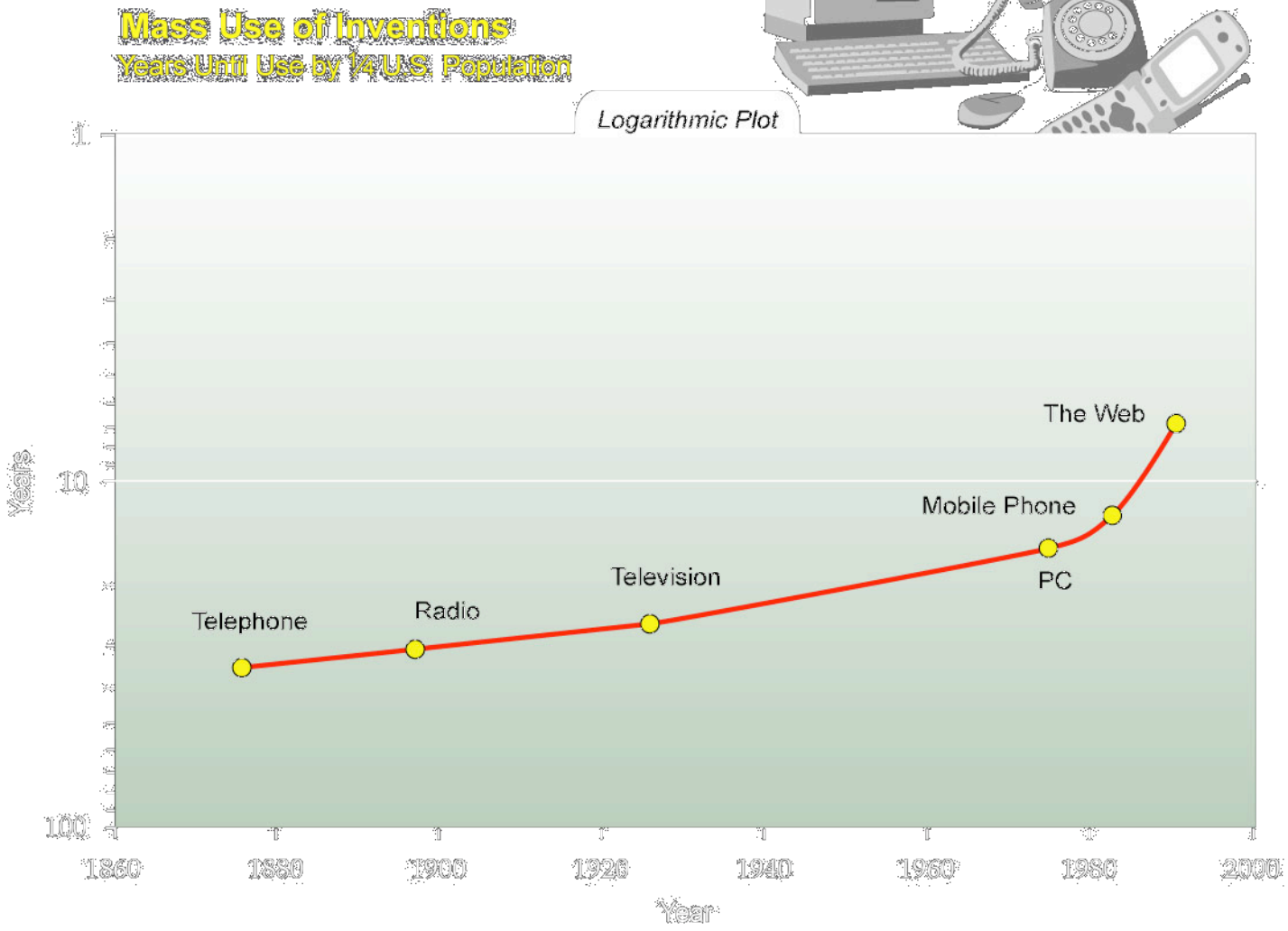
  Exponential trends in nature, technology - for 10’s, 100’s, millions of years

  S-curves, sure - but S-curves often stack

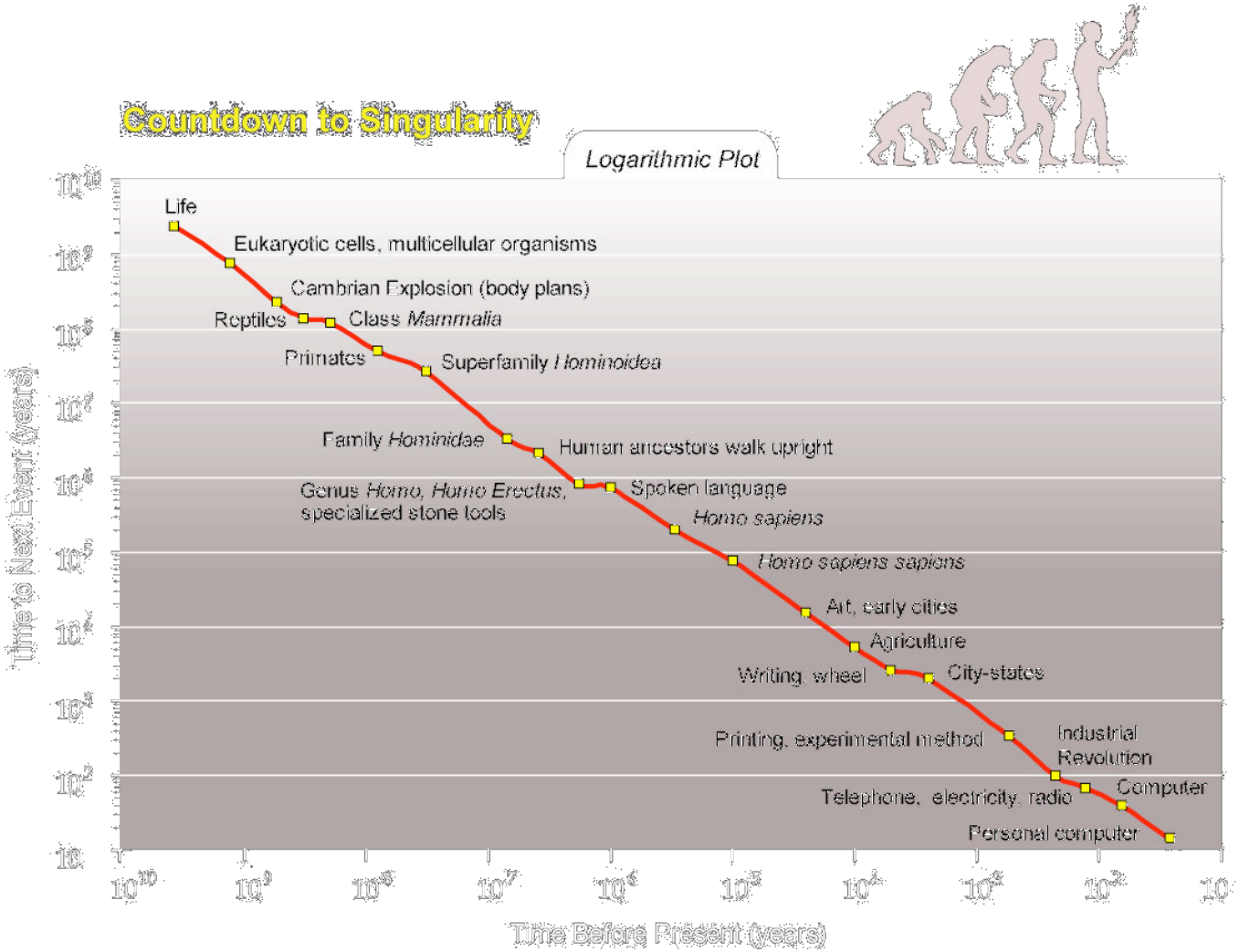
  Horizon: 3D circuit technology, carbon nanotubes

  (related: Intel announce on Hafnium vs. SiO2; ACM TechNews 070129; similar from IBM)

Kurzweil(2)…time to adoption

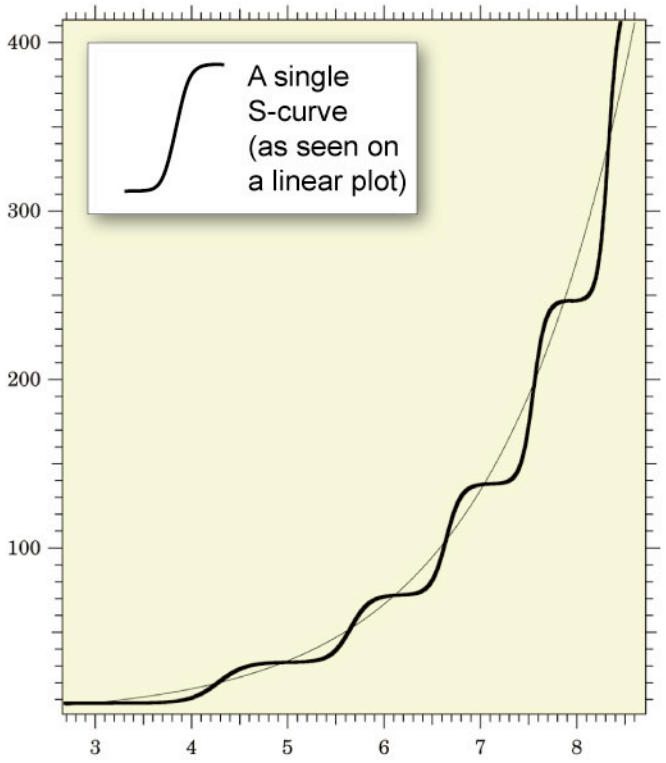


Kurzweil(3)… countdown to singularity

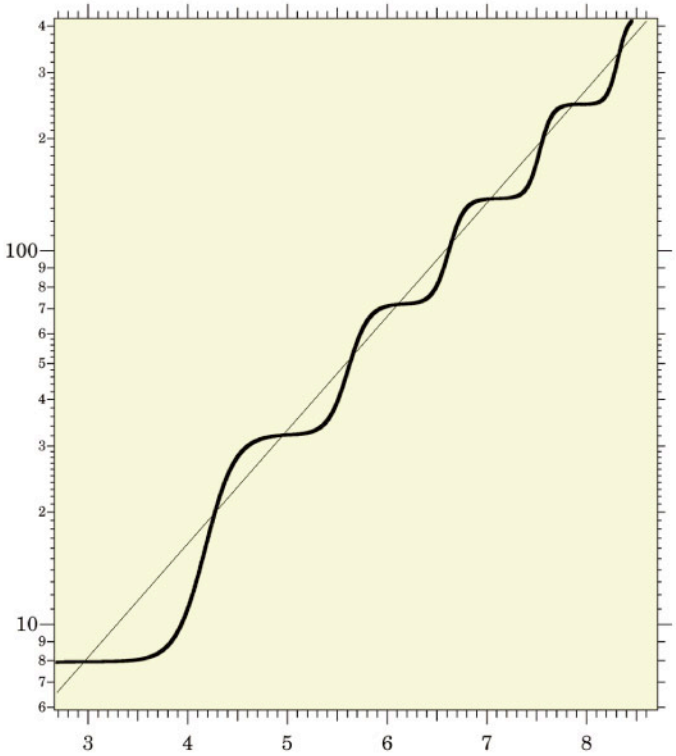


Kurzweil(4)… stacking S-curves

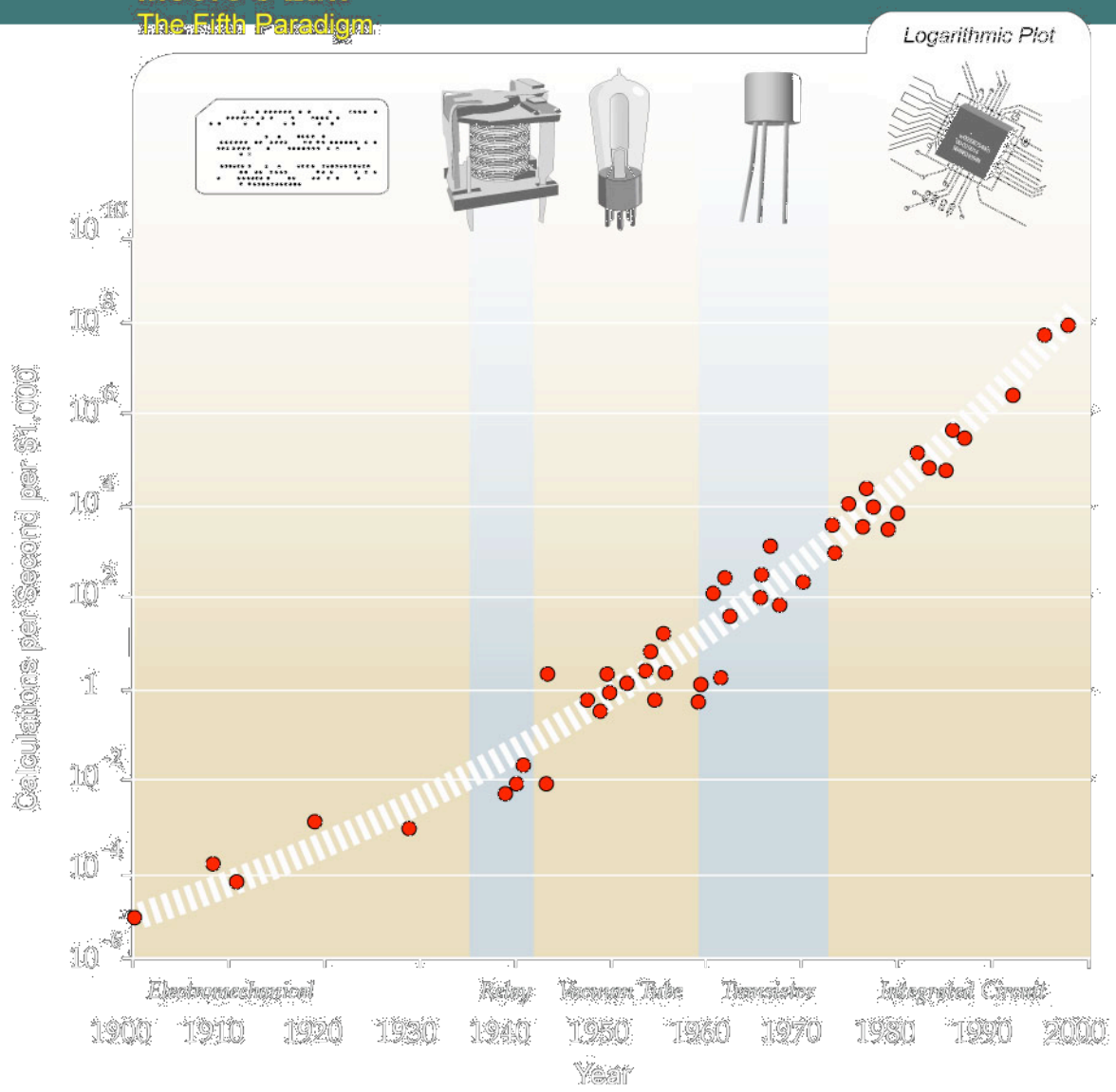
An ongoing exponential sequence made up of a cascade of S-curves (linear plot)



The same exponential sequence of S-curves on a logarithmic plot

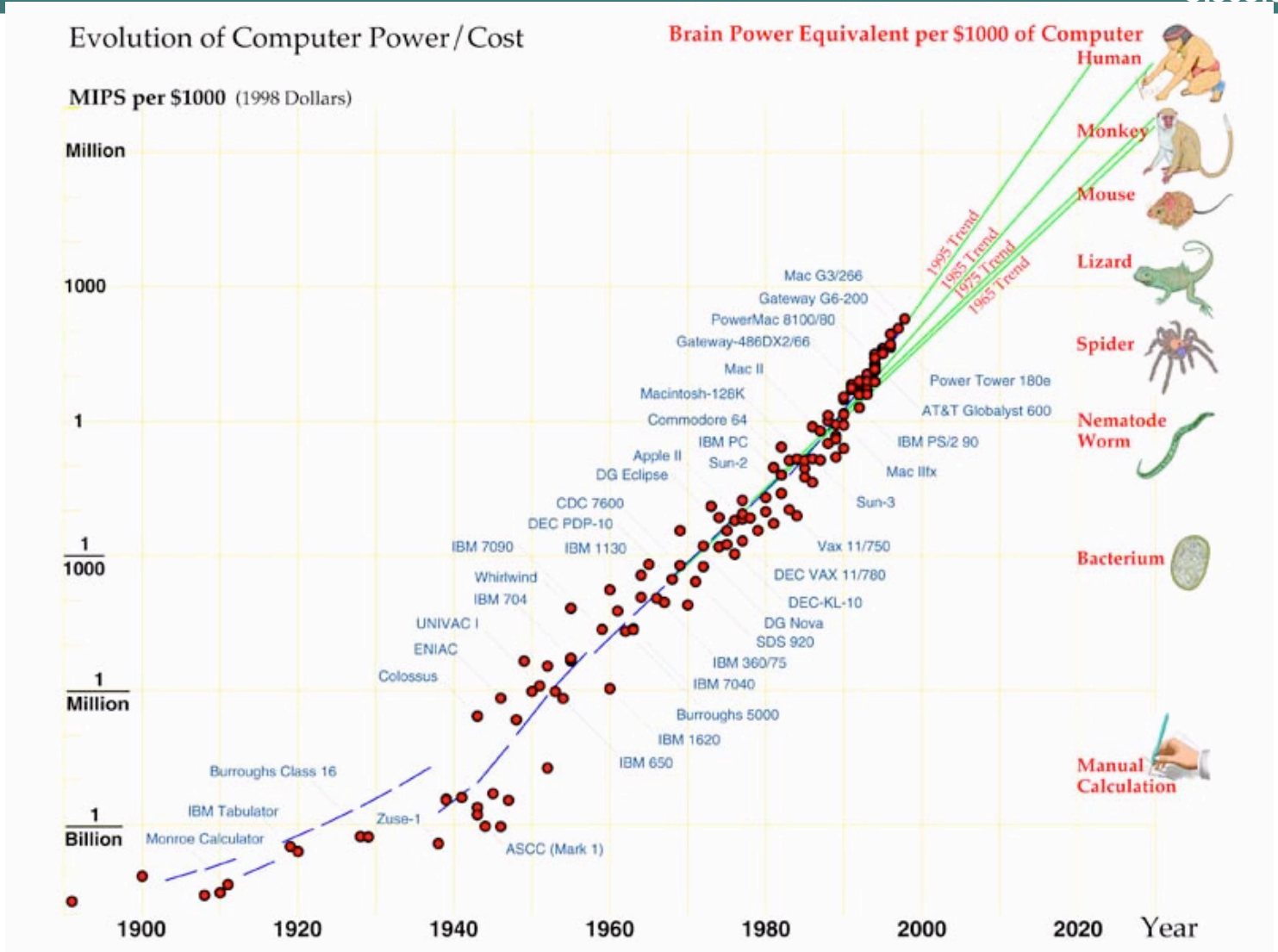


Kurzweil(5)… computing - 100yr view



Kurzweil(6)… compute-power trends(brain)

Cisco.com



What others say… IETF

- Paraphrase from discussion/comment on IETF list in past few months…
- Deployment characteristics are dominating, given more consideration than technology goodness

What others say- So what?

- Well-respected views on Lessons Learned, vs. gates-are-available, inevitable
- Real-router Power consumption, vs. Kurzweil observation on long-term technology trends
- PFLDnet - TCP vs. UDP vs. “other” prevalence/trends?
- How does this affect you?

Three sides of the coin

Side 1: It's "easy" to get ideas into routers

- If there's a unified voice
- And/or significant customer demand
- And timing expectations match
- Examples: BGP, QoS/queuing, RED
- Or in controllable domains with good incremental deployment properties (e.g. XCP ideas in MPLS/TE)

Side 2: It's "impossible"

- If the community doesn't have a single voice
- And/or there's no clear "winner"
- And/or customer demand doesn't materialize
- And/or gains are marginal or not aligned with perceived problems

Side 3 (the edge?): It's tricky

- Sometimes it's timing
- Luck
- Who you know
- Domain you can control (startup, niche population)
- Is this an example of "Crossing the Chasm?"

- Old quote: "If it were about coolest technology, NeXT would be thriving"

Implications for PFLDnet

Common themes heard at PFLDnet

- 25% “There are 50 metrics, we tested on 3, and are the best on one: x”
- 75% “Here is why ‘x’ is the metric we should all care about”
- “You have incorrectly implemented an old version of my algorithm, so your comparison is invalid.”

Single voice, common methods?

- Can we agree on a common set of metrics/methodology (or is that antithetical to PhD production?)
- TMRG vs. ICCRG - I still don't fully get it, even after Sally explained it. ;-)
- TMRG (Transport Modeling) draft cool - yet I thought that result might have come from ICCRG folks…

  draft-irtf-tmrg-metrics-06.txt

  (e.g., throughput, loss, delay, transients, stability, convergence, fairness)

Possible tools

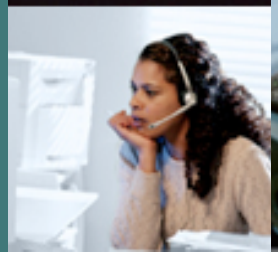
- **Injong's work on testbeds/scripts?**
- **TMRG or other metrics/testing as price-of-entry to PFLDnet?**

  Baseline plots/comparisons, then your unique attributes/argument

On sender-side-only vs. router-assist

- Entire class of “If I can get a little (per packet?) help from the router(s), life would be good.”
- Does this violate simple-core, innovation lessons? Do we care?
- If router-assist is “10” on some metric of goodness, and sender-side-only is “8”, what happens?
- Research topic: fundamental limits on achievable performance/behavior of the two styles?

Wrapping up…



Cisco research plug

- Doug Comer on board as VP for University Research
- Previous programs: University Research Program (URP), and Cisco Applied R&D (CARD) (some folks in room have seen +/- of this)
- Moving towards collaborative research - Requests for Proposals, so Cisco folks can give hints on what they're struggling with
- www.cisco.com/research

Conclusions

- There are conflicting views on the viability/advisability of incorporating new stuff into routers
- So any of easy/impossible/tricky might be correct, depending on perspective/experience
- PFLDnet might help the situation w.r.t. TCP, but it'll take cooperation, dedication, and time.