Power Consumption in High-End Routing Systems

Lawrence J. Wobker
System Architecture / Technical Marketing, Cisco
lwobker@cisco.com

Abstract

Power consumption and efficiency continues to be a major challenge in the networking industry. Current-generation products are increasingly constrained by the availability of power in both the developed and developing markets. While the unit power consumption (i.e. functionality per watt) continues to decrease, the rate of decrease is much lower than the explosion of bandwidth. This presentation discusses a number of topics related to power consumption in high-end routing/switching systems, with the primary goals being:

- educate network operators in the “physics” of system power consumption
- identify which performance metrics and requirements do and do not drive increased power consumption
- identify the mechanical/cooling impacts of system design
- discuss system hardware design challenges around power and cooling
- discuss the impact/costs of mechanical standards compliance (i.e. NEBS)

Power consumption in high-end systems

(and hopefully what we can do about it)

- why does power consumption matter?
- the “peak load” stat-mux problem
- the physics of power consumption
- non-linearities
- router power consumption
- NEBS vs. the environment

Why does power consumption matter?

(or at least some reasons)

- economics
- it limits network growth
- environmental impact
  - altruistic
  - regulatory
  - political

Router/switch power consumption

(Why are these systems so high power???)

- density sells routers, density drives power
- silicon capability grows FASTER
  - than power usage decreases
- someone, somewhere can power them.
  - just maybe not you…
- Summary:
  - unit power **will** continue to decrease
  - total power **almost certainly will** increase

The "peak load" stat-mux problem

(and what we probably can't do about it)

- power is provisioned for worst case
- facility cost model: cost = Ax + By + Cz
  - *A: power consumed (kWh) * rate x*
  - *B: provisioned power (breaker-amp) * rate y*
  - *C: cooling/thermal (rack? amp? kWh?) * rate z*
- B is often >> A
- C is often a direct function of B.
- Summary: paying for unused power

The physics of power consumption

(things we definitely cannot do anything about)

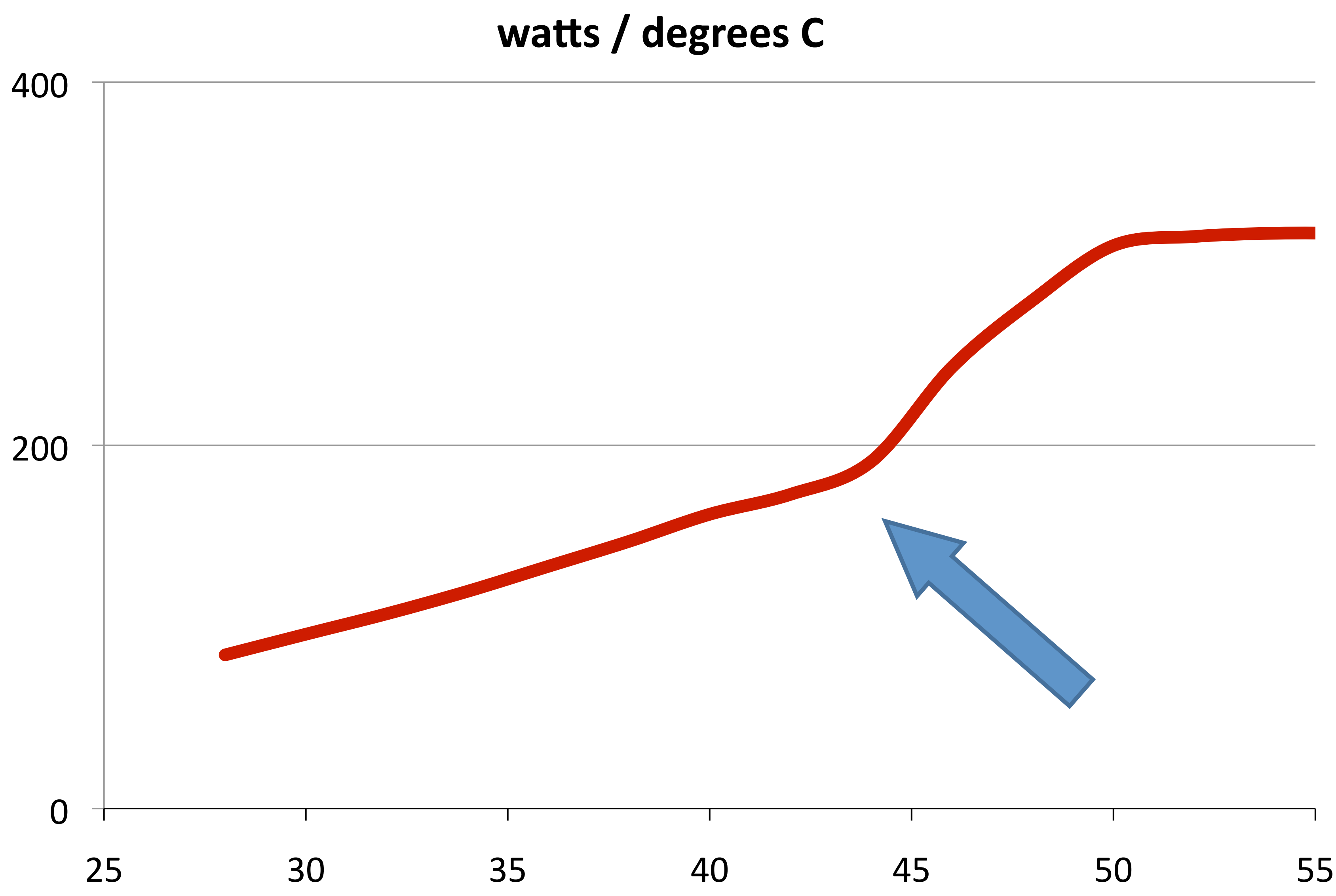
- water cooling is ***probably*** not realistic
- mechanical cooling == fan trays
- larger fans are more efficient
- ***R = cT:*** *R = resistance c = material coeff T = temp*
  - translation: hot routers use more power

Non-linear equations

(and not that really hard math course you didn't take)

- cooling capacity = airflow * heatsink
  - $\textit{\textbf{cooling cap}} = (\, fan_{cfm} * \Delta T \,) * (\, mm^2 * tC \,)$
- fan volume (CFM) nonlinear vs. power used
- cooling capacity (dT) nonlinear vs. temp
- heatsink size and materials are fixed
- power supply efficiency nonlinear vs. load

fan power required vs. ambient temp



Router/switch power consumption

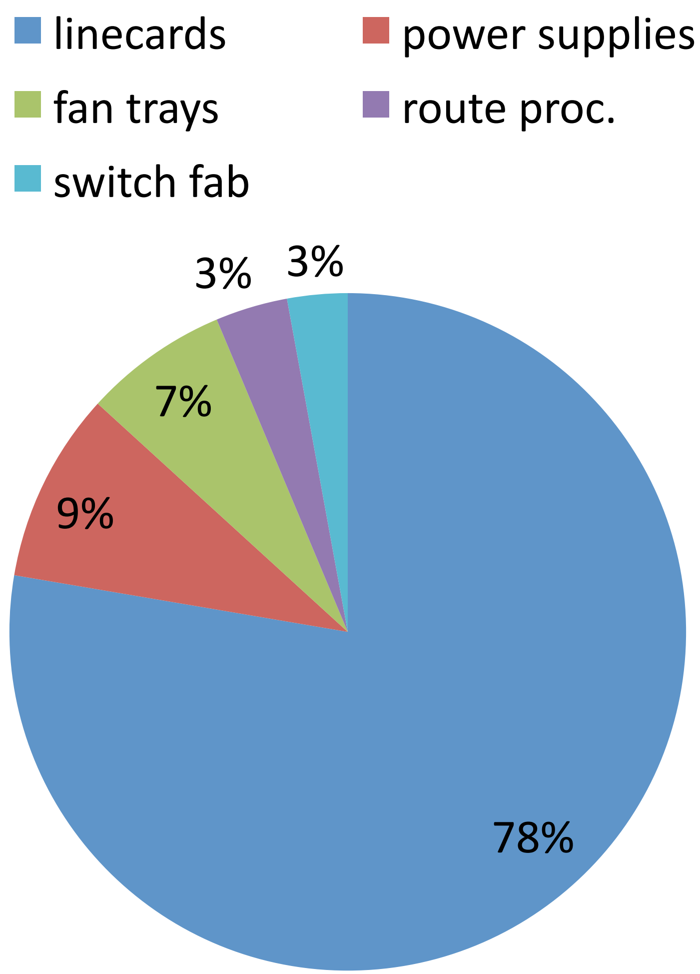
(Where do we use a lot of power?)

- control plane, fabric, and mechanical components are amortized
- cooling 1 slot ~= cooling ALL slots
- fans: lots of variance due to temperature
- “full” systems are more efficient
- linecards are dominant consumers
- power % shifts to fans at higher temps

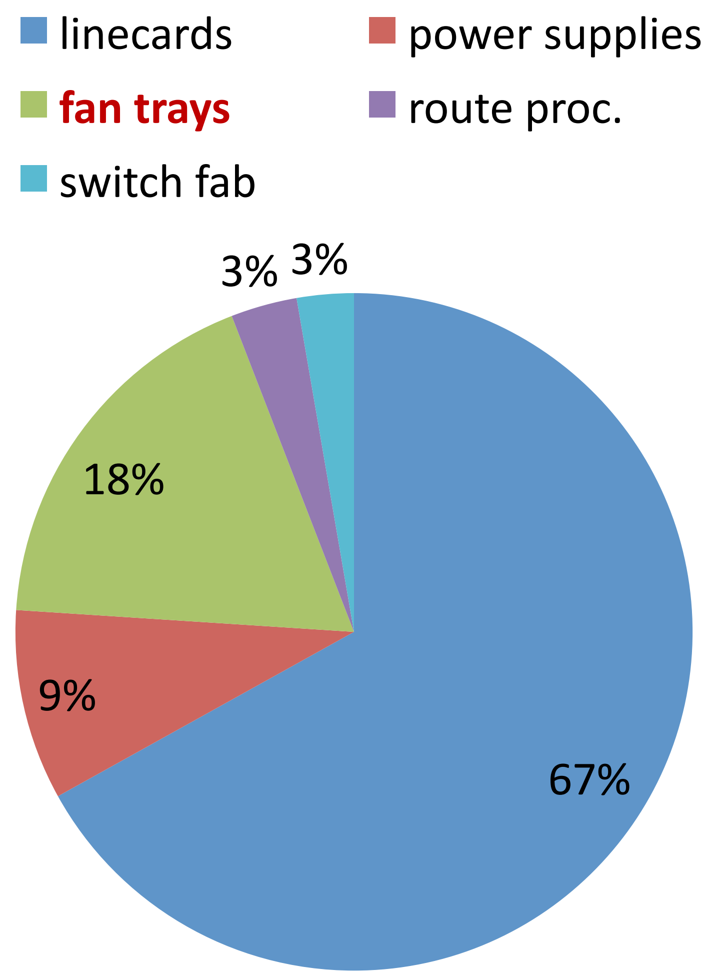
System: consumption by component

(large system at 25 vs. 50 degrees C)

25C power: 10 KW

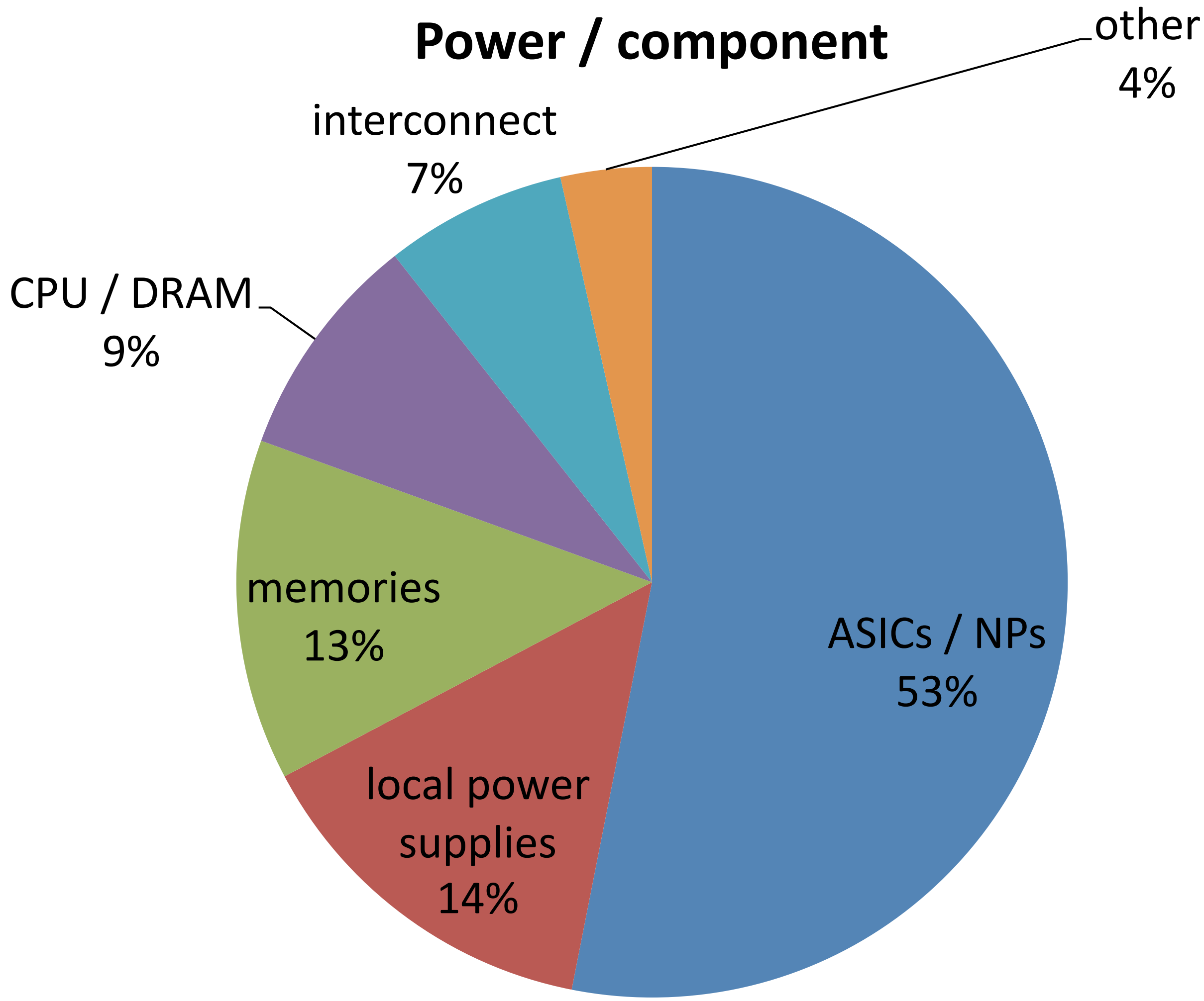


50C power : 14 KW



Linecard: consumption by component

(what things matter and what things do not)



Router/switch power consumption

(linecard power usage)

- silicon power is dominated by chip complexity
- chip complexity = gate count (# of transistors)
- gate count is mainly a function of
- peak/top end performance requirements
    - packets / second + bits / second
- total feature complexity
    - # of lookups, classifiers, stats, inspections

NANOG Fables : the Porsche and the Prius

(why top –end incremental performance is expensive)

power output: 98hp
mass: 3042 lbs
0-100km: 9.7s
$25k @ 50mpg

output: 500hp
mass: 3312 lbs
0-100km: 3.5s
$125k @ 17mpg

Moral of the story: very different energy usage & cost when performing "nearly" identical tasks.

NEBS vs. the environment

(what things matter and what things do not)

- NEBS requires 55C for “small” systems
- 50C for “large” systems
- high temperature operation for short periods
- “short” helps reliability but not cooling

Things to ponder

(how do we manage this going forward?)

- Porsches or Priuses?
- performance vs. power (IMIX vs. 64B?)
- acceptable max ambient temps (40? 55?)
- lower cooling costs → higher power costs?
- marketing myths vs. realities

A cry for help…

(anyone feel like collecting data?)

- lab != deployment
- vendors care about selling
- vendor testing is primarily around boundary conditions and compliance
- improving instrumentation
- we need data to know what to build…

Thanks.

Lawrence J. Wobker
System Architecture / Technical Marketing, Cisco
lwobker@cisco.com