Characterization and Benchmarking Methodology for Power in Networking Devices

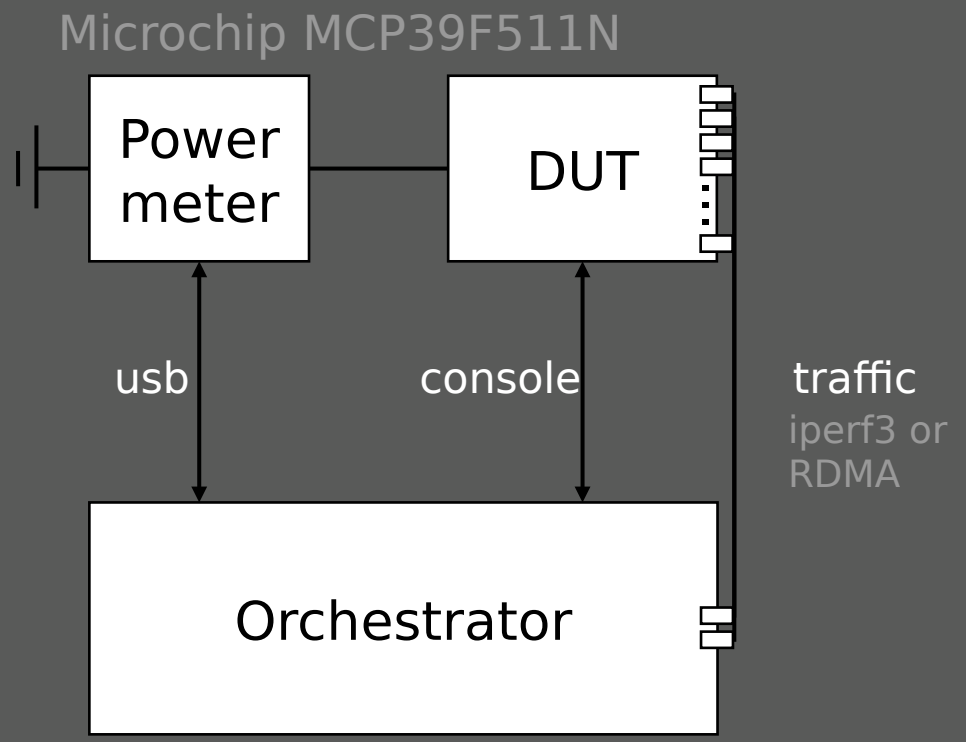
Romain Jacob
ETH Zürich

IETF 121 - BMWG
Nov. 7, 2024

Who am I?
Why am I giving this talk?

(This is not me.)

- I researched ways to reduce energy use in wired networks for ~2 years.
- There are several ideas around, but hardly ever properly evaluated.
- I worked a lot on deriving power models for networking hardware.




```yaml
metadata:
  device_id: aristaDCS-7280CR3K-32D4
  connected_PSUs: 2
QSFP:
  PCC:
    100G:
      P_BASE: 266.7495
      P_PORT: 0.6896069518716569
      P_TRX: 0.554247576871658
      P_TRX_IN: 0.22439062499999984
      P_TRX_UP: 0.3298569518716582
      E_BIT: 1.2146658059270798e-11
      E_PKT: 9.336373009493466e-09
      P_OFFSET: -0.3940915741265485
QSFP-DD:
  FR4:
    400G:
      P_BASE: 266.7495
      P_PORT: 1.8164000000000124
      P_TRX: 5.358466379310332
      P_TRX_IN: 5.427624999999992
      P_TRX_UP: -0.06915862068965994
      E_BIT: .nan
      E_PKT: .nan
      P_OFFSET: .nan
```

The Good

The Bad

The Ugly

1 The Good

Defining “benchmark problems” isn’t hard.

2 The Bad

Making the benchmark practical is hard.

3 The Ugly

Making the benchmark trustworthy is worse.

1 The Good

Defining “benchmark problems” isn’t hard.

The Bad

Making the benchmark practical is hard.

The Ugly

Making the benchmark trustworthy is worse.

The Good

Defining “benchmark problems” isn’t hard.

2 The Bad

Making the benchmark practical is hard.

The Ugly

Making the benchmark trustworthy is worse.

The power model we are trying to derive is fairly simple.

Device power = Base power

+ Static power per port

+ Energy per packet * packet rate

+ Energy per bit * bit rate

f(device config)

To identify those parameters, we need different experiments.

	Config.	Transcievers	Traffic
▪ Base	None	Unplugged	None
▪ Idle	None	Plugged	None
▪ Static	Admin-up	Plugged	None
▪ Trx	Link-up	Plugged	None
▪ Snake	VLANs	Plugged	Load testing

* transceiver types
* port types
* port speed

* packet sizes
* bandwidth

* multiple runs for replicability

To identify those parameters, we need different experiments. A lot of them.

	Config.	Transcievers	Traffic
▪ Base	None	Unplugged	None
▪ Idle	None	Plugged	None
▪ Static	Admin-up	Plugged	None
▪ Trx	Link-up	Plugged	None
▪ Snake	VLANs	Plugged	Load testing * packet sizes * bandwidth

* transceiver types
* port types
* port speed

* multiple runs for replicability

This can be automated (we've done it) but finding the proper config remains a pain every time.

- Every router model and OS has its quirks.

  Different command types
  Commiting or not
  Different ordering required

- Knowing that "It has to be possible"
  does not help finding the correct commands...

There is an important hidden assumption here:

Ports of the same type are assumed to be equal.

- This is not always true.
- It is hard to detect this before measuring.

  Datasheets won’t necessarily tell you.

  You do not know how the HW looks like inside the box.
- It is hard to detect this after measuring.

  Power effects are small and noisy.

  Randomization of used ports makes it hard to detect.

Okay, we've got it done!

Can we trust it?

Benchmarking is about **ranking.**

- “Is A better than B?”
- “Is new-A better than old-A?”

Trust is a “MUST” in the RFC sense.

The Good

Defining “benchmark problems” isn’t hard.

The Bad

Making the benchmark practical is hard.

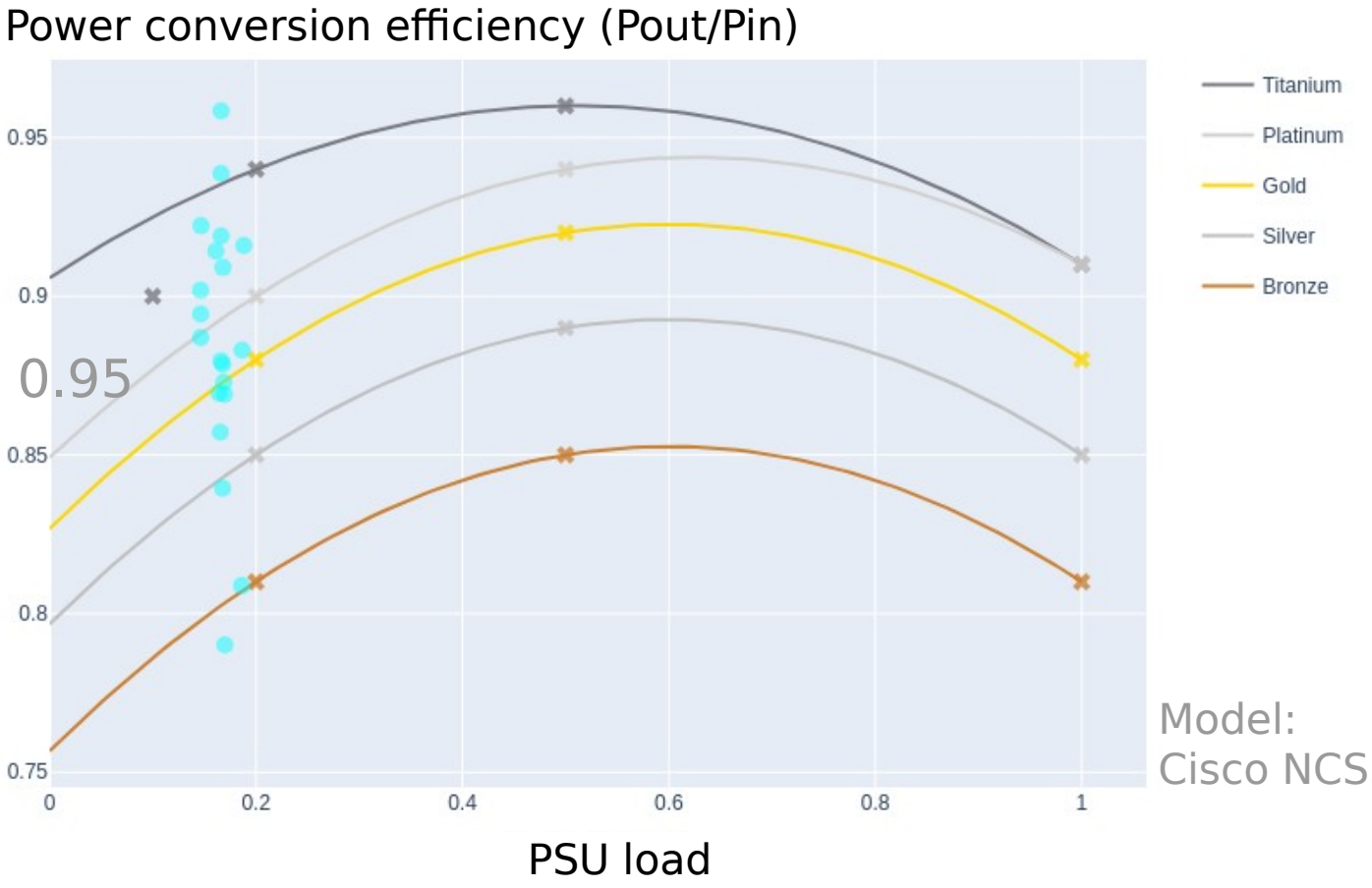
3 The Ugly

Making the benchmark trustworthy is worse.

Things change in unpredictable and sometimes surprising ways.

The power conversion efficiency of the same router model can vary from “very good” to “very bad.”

- Temperature?
- Aging?
- Manufacturing fluctuations?



0.95

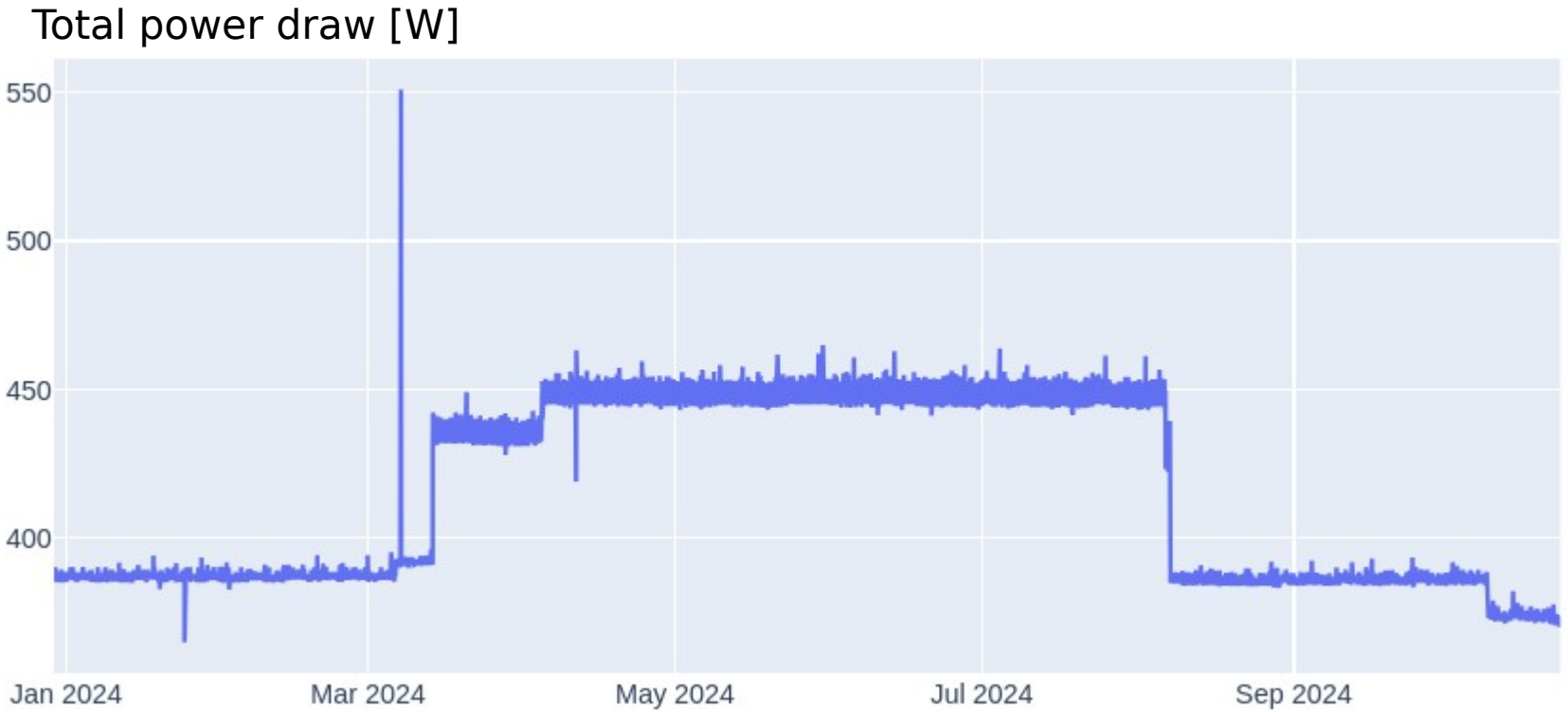
0.78

Model:
Cisco NCS55A1-24H

Things change in unpredictable and sometimes surprising ways.

Large power changed resulted from an OS update!

- Different fan policy
- Eventually reverted (it seems)



Model:
Cisco 8201-32FH

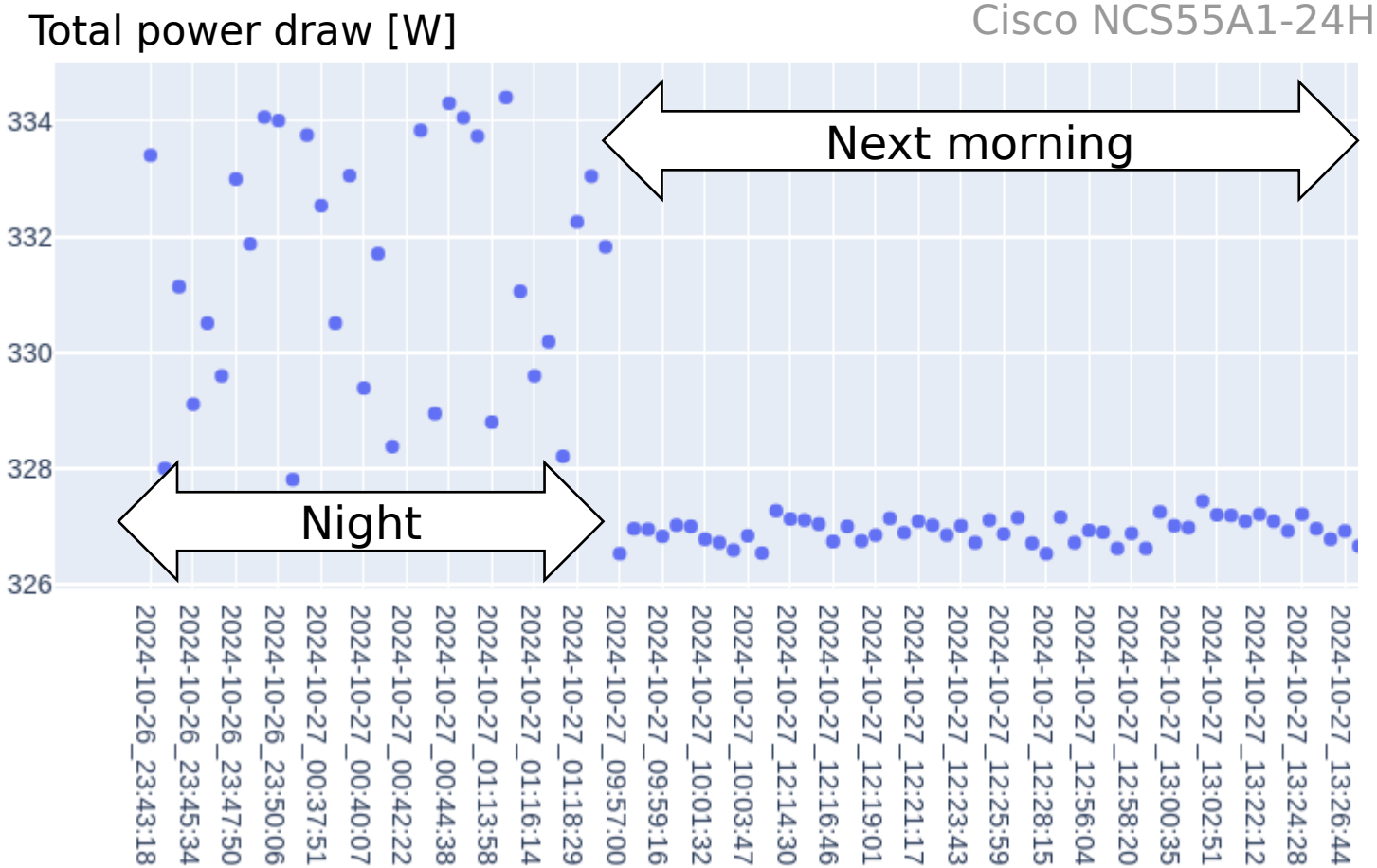
Things change in unpredictable and sometimes surprising ways.

- Same machine
- Same experiment
- Same room

- Different... time?

Still no clue
what happened.

Model:
Cisco NCS55A1-24H



Benchmarking is about **ranking.**

- “Is A better than B?”
- “Is new-A better than old-A?”

Trust is a “MUST” in the RFC sense.

... but it seems we ain’t quite there.

So what?

The Good

Defining “benchmark problems” isn’t hard.

The Bad

Making the benchmark practical is hard.

The Ugly

Making the benchmark trustworthy is worse.

We MUST make it easy(ier) to “benchmark your own device.”

- Our power profiling toolchain will be open sourced.
  Currently focusing on writing up the paper...

- IETF contributors could undoubtably help
  build a much better and more robust toolchain.

 Help welcome!

We SHOULD increase trust by using large numbers.

- I think we should not specify a router power benchmark so strictly as to guarantee replicable results.

  Maybe this can be done, but
  how useful would the resulting benchmark be?

- Instead, I propose to make it easy to run somewhat imprecise benchmarks and aggregate the results publicly.

  Akin to “meta-analysis” performed in other fields
  with high uncertainty in single experiments.
  E.g., psychology, medicine, environmental sciences, ...

We created a public database for power data: **NetPowerDB**

The database contains

- Datasheet information
- PSU readings
- External measurements
- Power models

Welcome to the
networkpowerzoo.ethz.ch

Prototype stage

Characterization and Benchmarking Methodology for Power in Networking Devices

Romain Jacob
jacobr@ethz.ch

networkpowerzoo.ethz.ch