

Nuts and Bolts and Watts and Gates - all that mechanical and hardware stuff

cisco Live !

LJ Wobker
Principal Technical Marketing Eng.

Mandatory Administrative Stuff....

Cisco Webex App

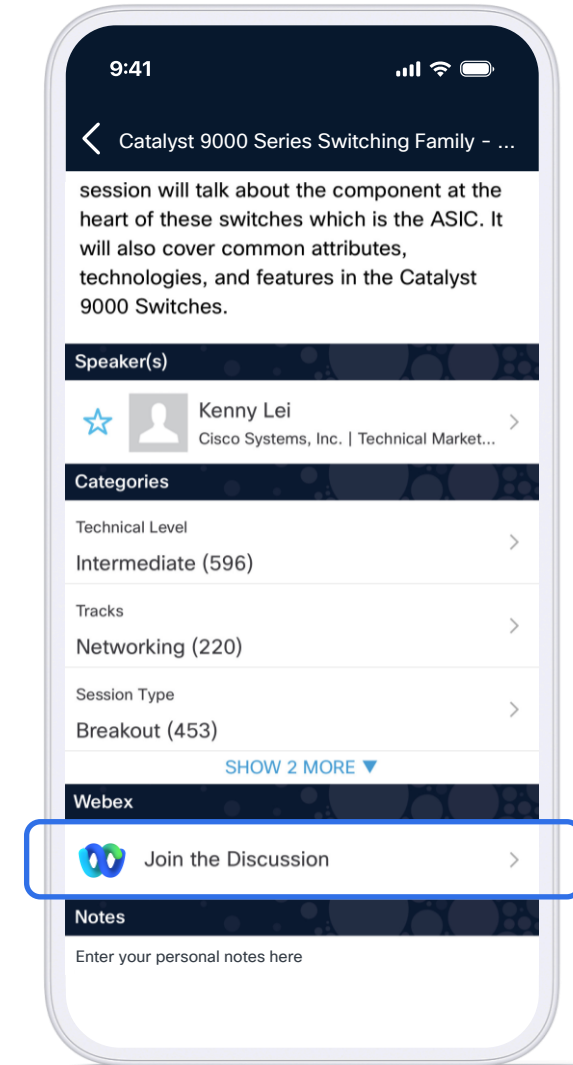
Questions?

Use Cisco Webex App to chat with the speaker after the session

How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until 14 November 2025.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKMSI-2501>

Agenda

- 01 Intro / Overview
- 02 Utility scale
- 03 Building chassis
- 04 Thermals
- 05 Board design
- 06 Signal Integrity
- 07 Optics
- 08 Chips

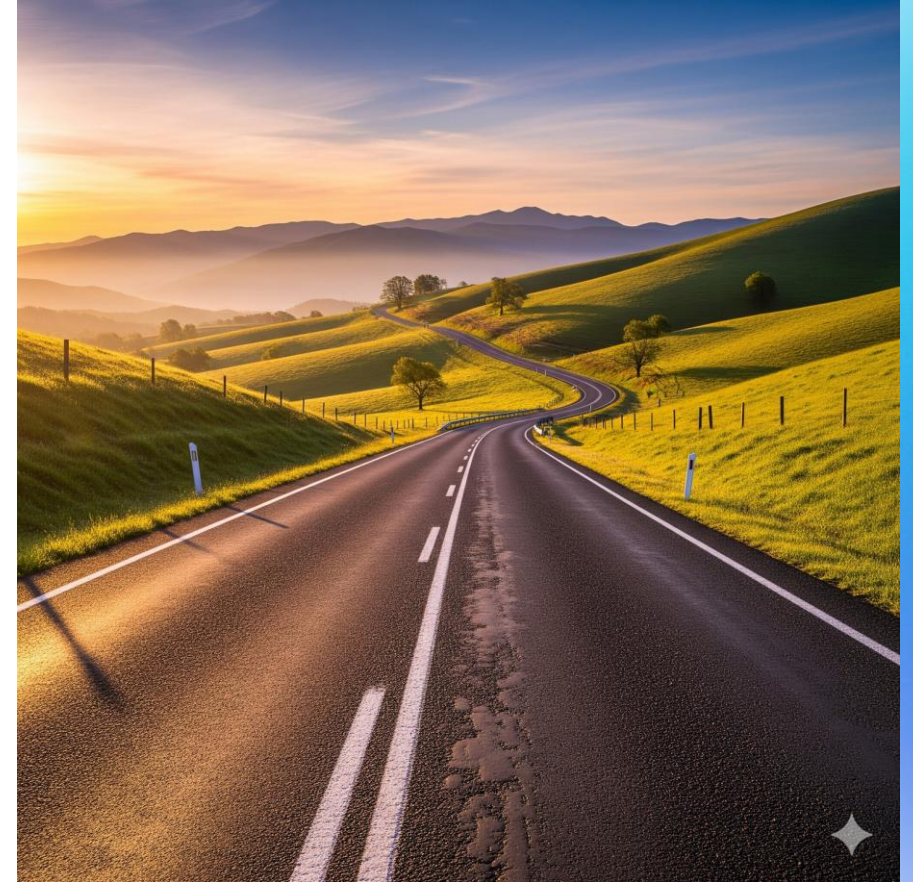
Speaker Intro

- RTP TAC University intern '97-'99
- TAC and ISP Expert Team '99-'01
- GSR Escalation / Deployment '01-'03
- GSR, CRS TME '03-'06
- AS (CX) Consulting Engineer '06-'07
- ASR 9000 / NCS 6000 TME '08-'14
- Barefoot Networks '14-'15
- Silicon One and Cisco 8000 TME '16-'??
- *This presentation contains AI-generated content.*
- *Your speaker is not (yet?) AI-generated.*



What we're going to really talk about...

- The “hardware” – what goes into building these things?
- What are the mechanical and electrical challenges?
- The presentation is slightly circular!
- Slightly more “wide” than “deep”
- COME UP WITH QUESTIONS!
- Such as ...
 - *Why on earth did you guys... do ... !?!?!*



Really Big Themes

- Hardware is ... Hard.
- It takes a long time to change things.
- It costs a lot to change things.
- Tradeoffs, tradeoffs... Tradeoffs!
- All designs are a tug-of-war



Really Big Themes

- Tradeoffs, tradeoffs... Tradeoffs!
- Stuff becomes nonlinear.
 - Sometimes mathematically, sometimes non-technically
- “Exchange rates” shift over time
- There are always reasons.
 - Sometimes, there are even good reasons.

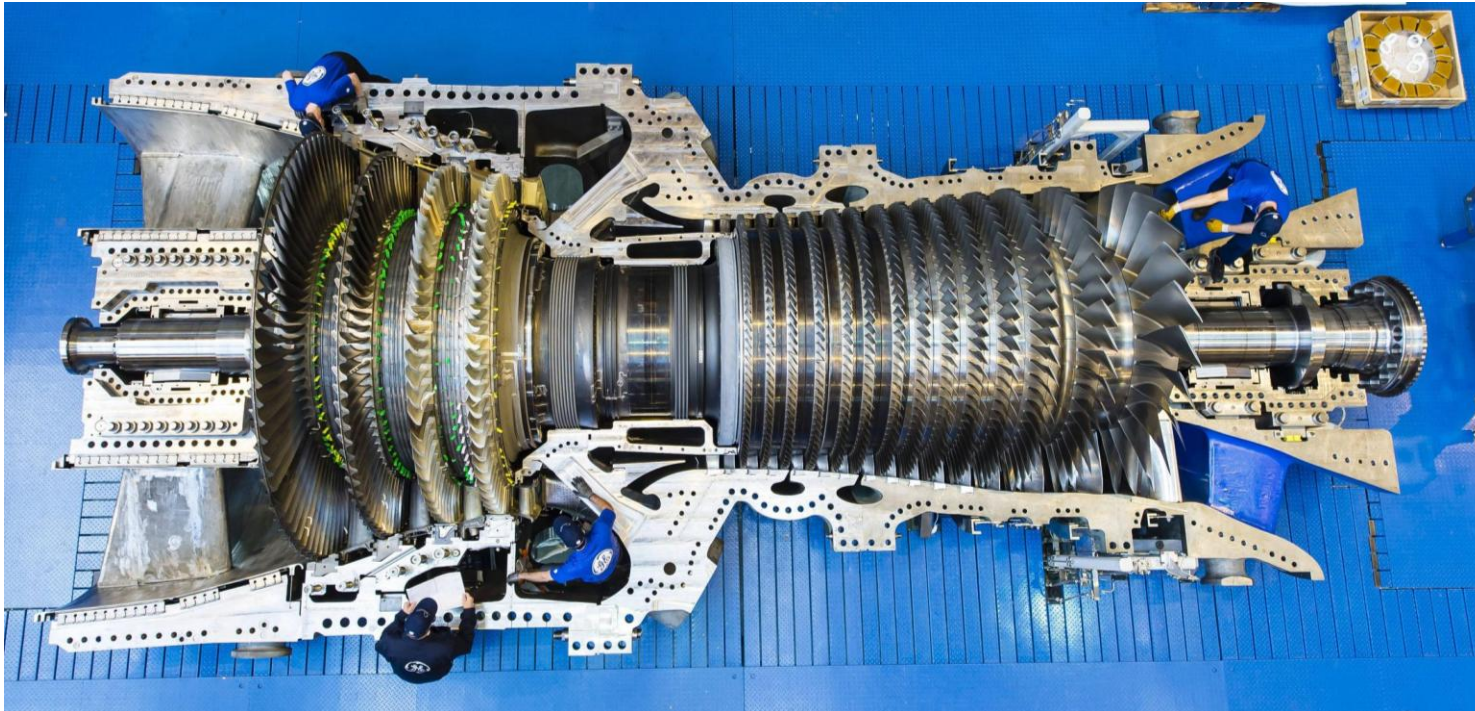


Buildings and Utilities

ONE-POINT-TWENTY-ONE GIGAWATTS ?!?!?!?

What the hell is a gigawatt?!?!?

- That's 1210 MW
- Or about 1/5 of Victoria's average consumption
- Very rough: 1k homes per MW



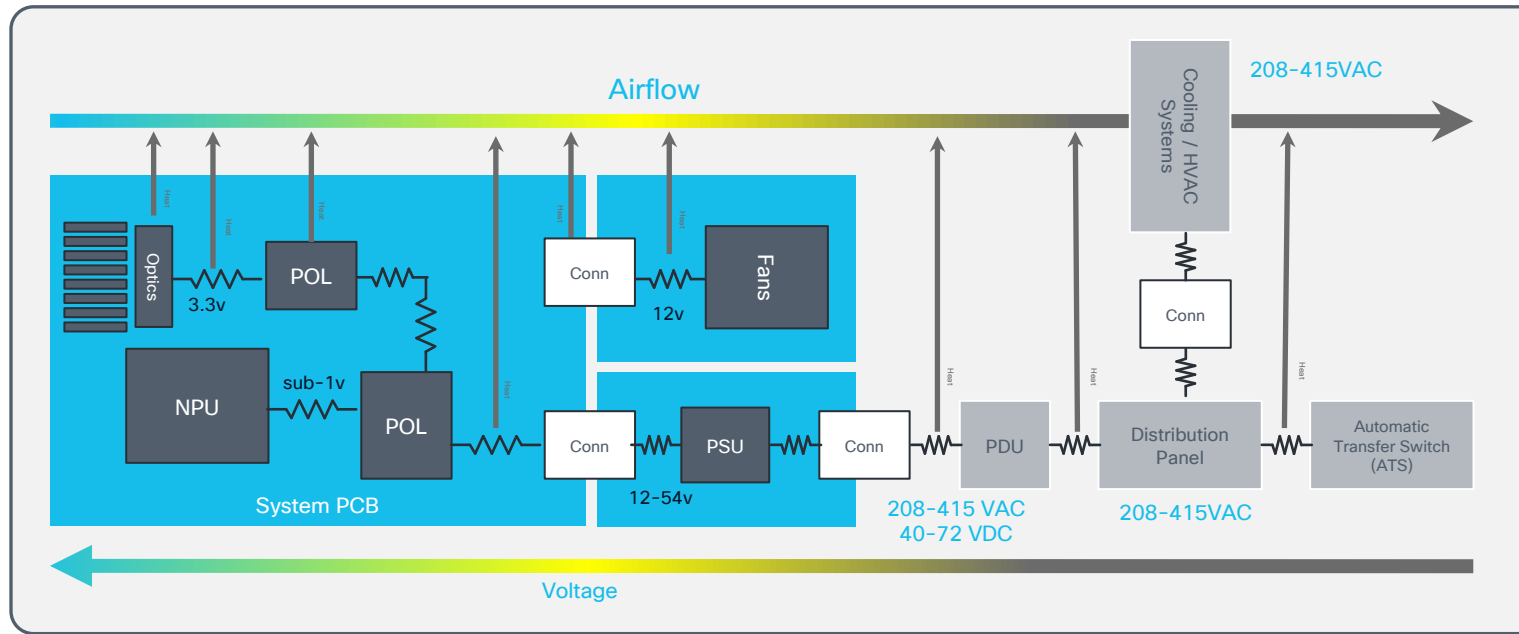
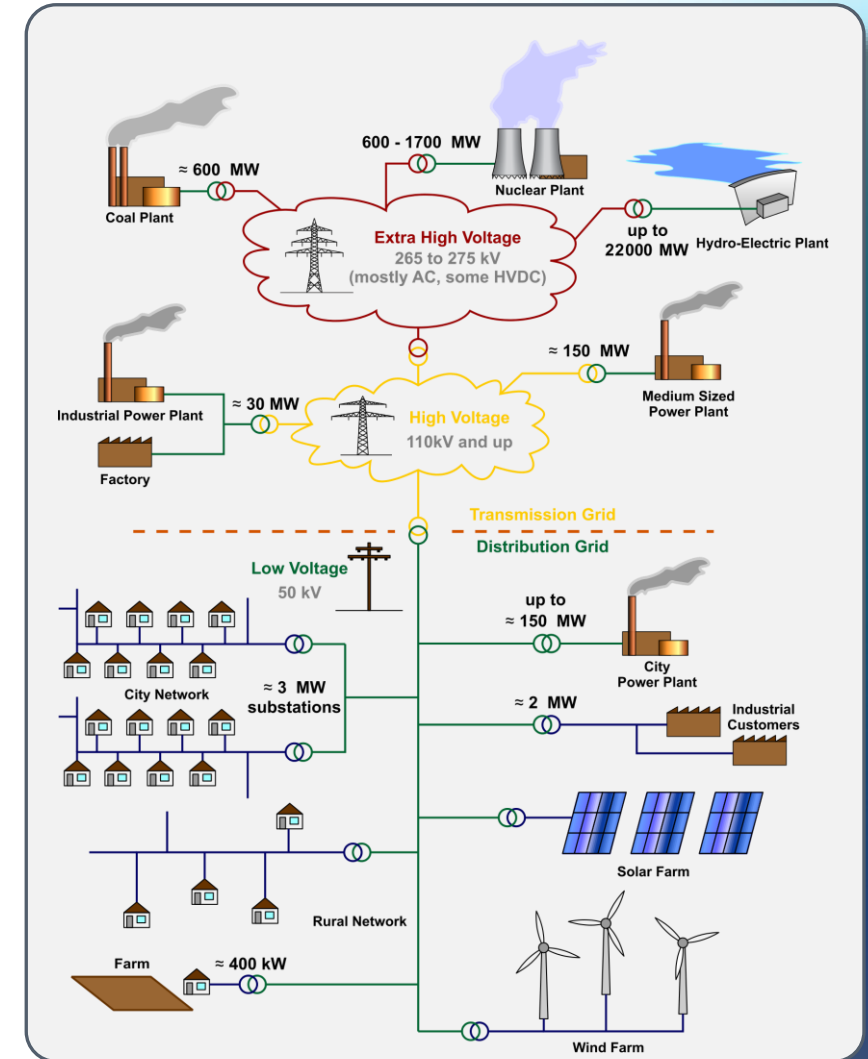
~ 600MW gas turbine

~ 3.5m diameter

~ 400 tonnes

Getting the power in

Generation, Transmission, Distribution (utility)
PSUs, POLs, Busbars, VRMs, oh my!



Free career advice: Do you know your facilities contacts yet???

Chassis

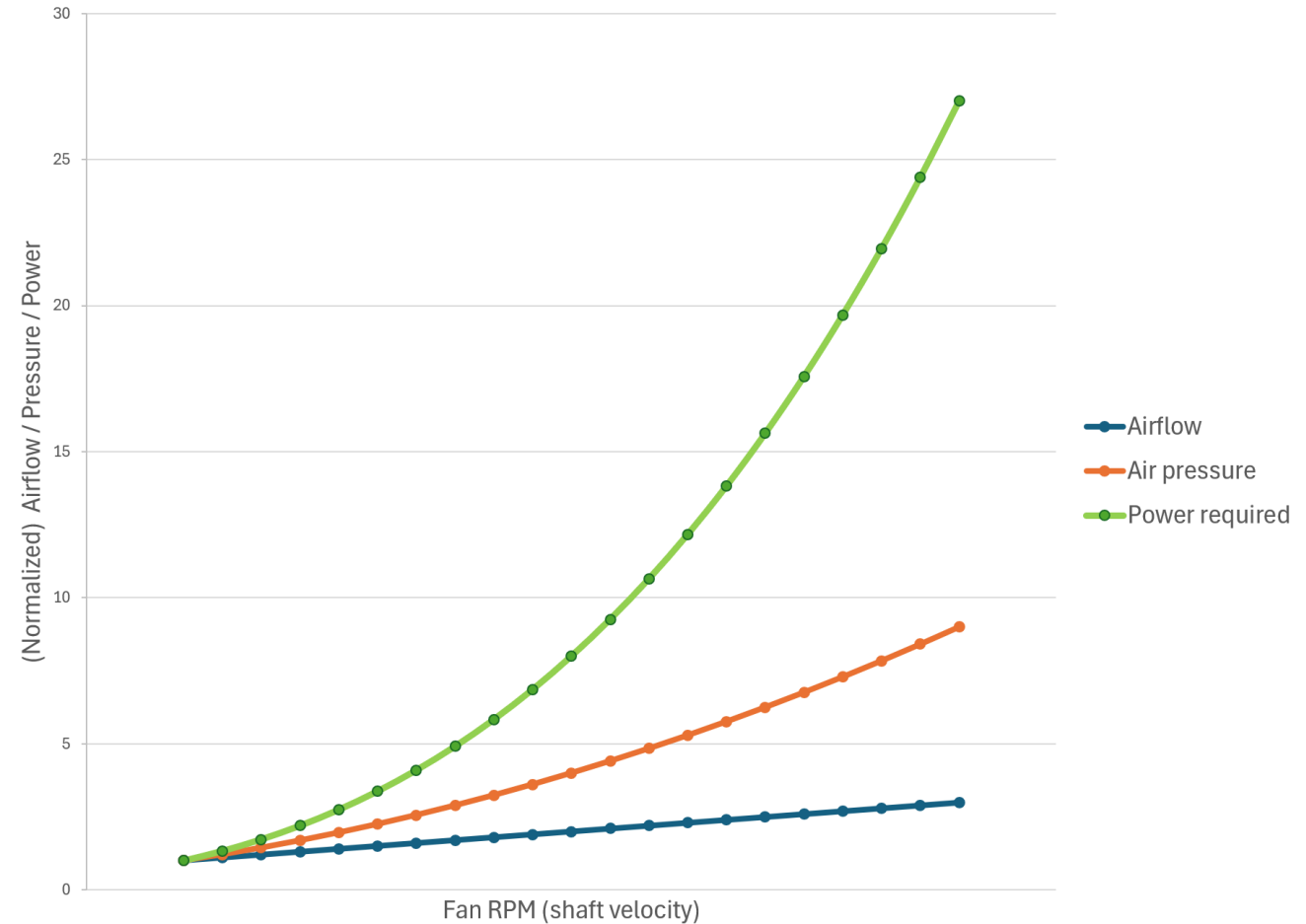
System design: transformation / stepdown losses

- Wide range of voltages, lots of steps...
- High power : BIG bus bars...
- Hundreds of volts but **thousands** of amps!
- Each transform costs \$\$\$ and W



Getting the heat away from the gear...

- A watt in is (pretty much) a watt out.
- Fans = easy to understand.
- But! they have painful laws.
- Liquid is better, except when it isn't.



Chassis: large modular systems

- Multiple linecards
- Everything field-replaceable
- Large high-power busbars
- Very large individual boards
- Shipping / handling
- Connector alignment / insertion force
- Very high total power input / heat output



An actual conversation with one of our mechanical engineering leads

8/29/2025, 4:52 AM

(j/k) what "small stuff"? you calling me/us small?

8/29/2025, 4:54 AM

adding our ME+thermal leader

8/29/2025, 4:54 AM

also adding for visibility and comments if any



You 8/29/2025, 4:54 AM

This procedure outlines the steps for installing the **pull-bars** on the 8818 chassis. The **pull-bars** serve as an aid for moving the chassis to the lifting device or closer to the rack where the chassis is installed.

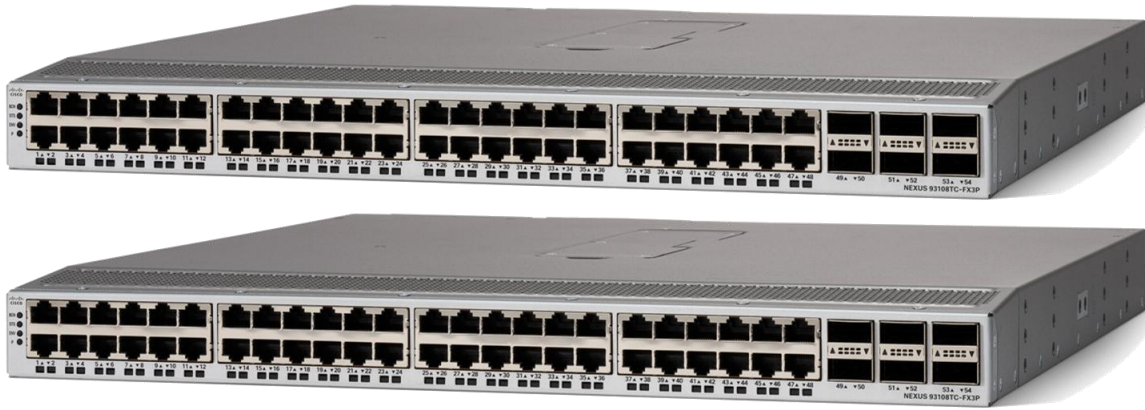
 Do not use these **pull-bars** to lift the chassis. Follow the orientation specified on the **pull-bars** to identify the correct right and left orientation.

lol - if the install guide doesn't start with something like "install the special pull bars so you have a chance of moving the chassis" ... then I consider it a "small" system ;-P



1RU / 2RU “Pizza boxes”

- Port density
- Volume constraints
- Microelectronics are smaller than fingers!
- Small diameter fans
- No “real” standard for depth

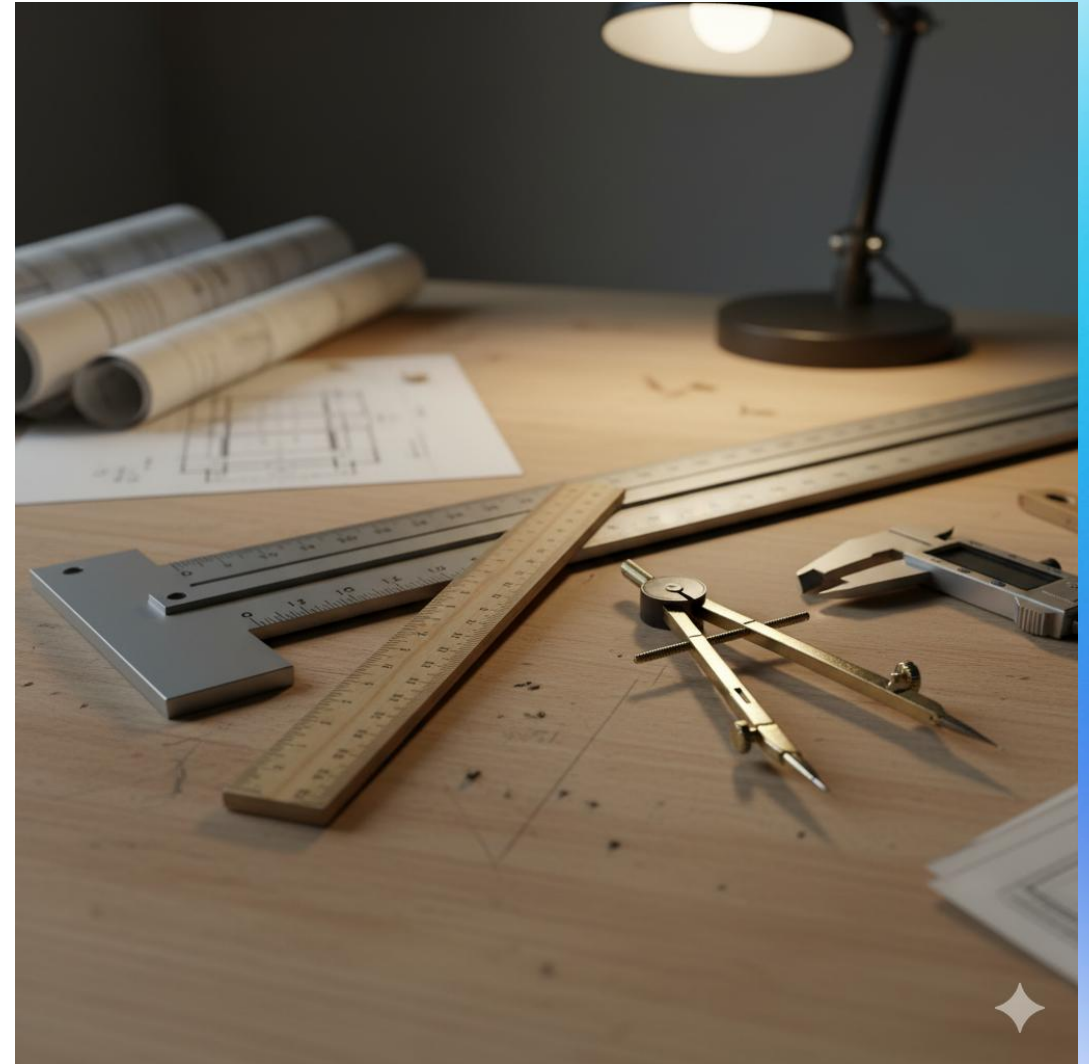


Physical dimensions

Rack sizes are standard, right?

- Width: yes (mostly)
- Height: kind of?
- Depth: Oh, most certainly not.

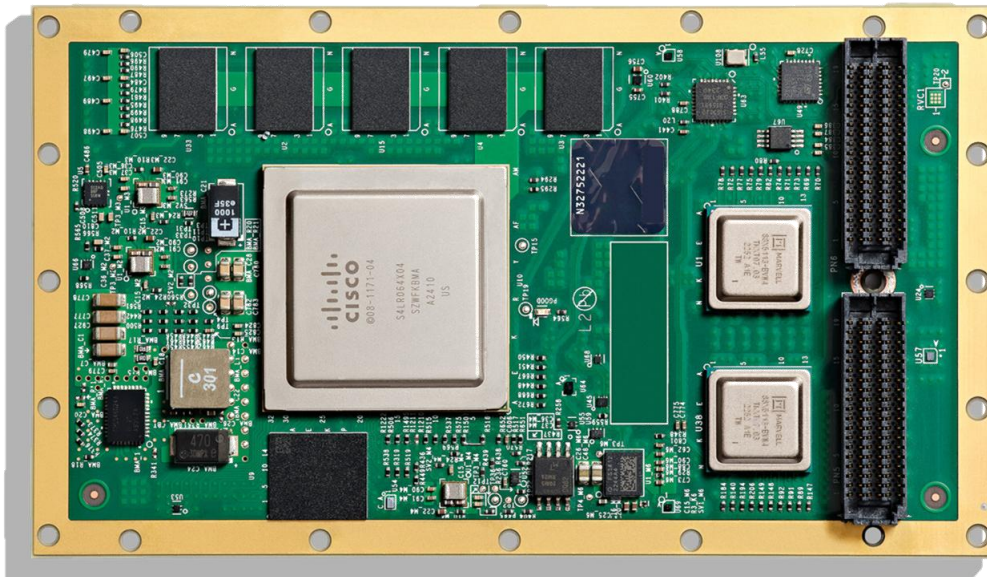
There are some step functions in sizes, too.



Industrial / extreme environmental switches

- Industrial / extreme environment
- Connectors
- Salt water

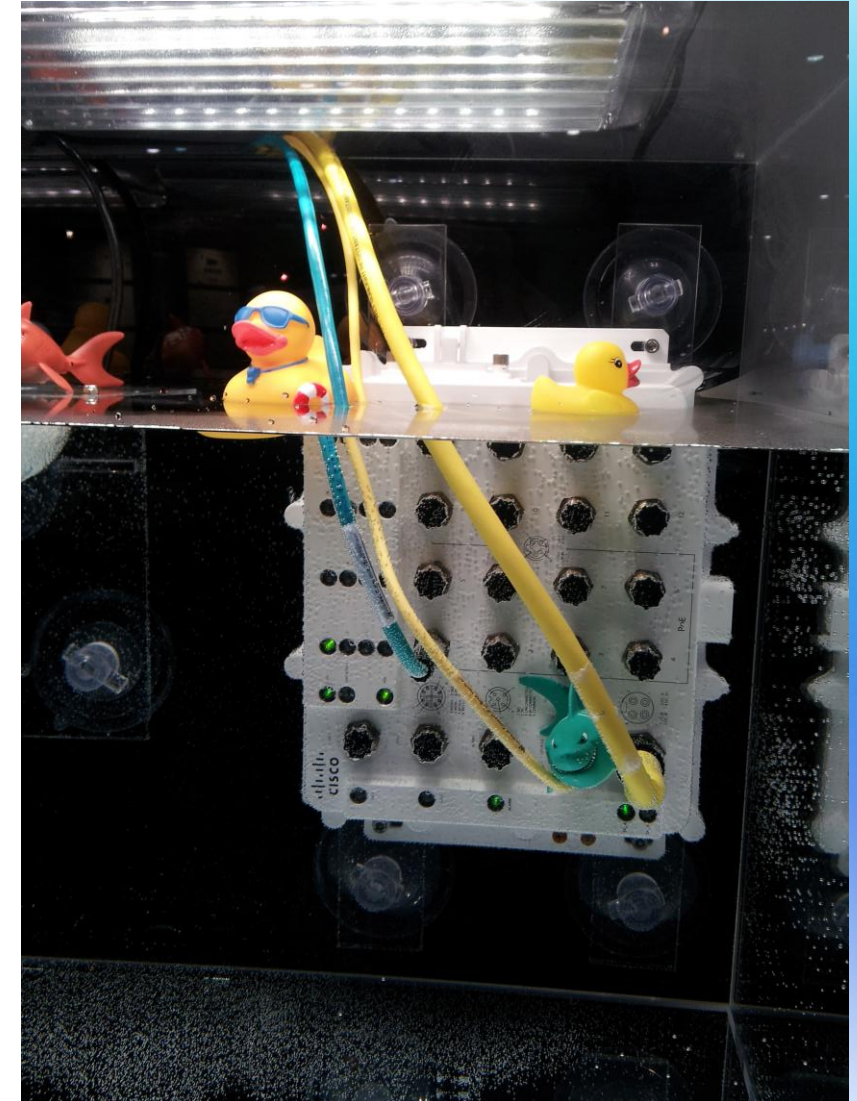
ESS-9300-SX embedded switch



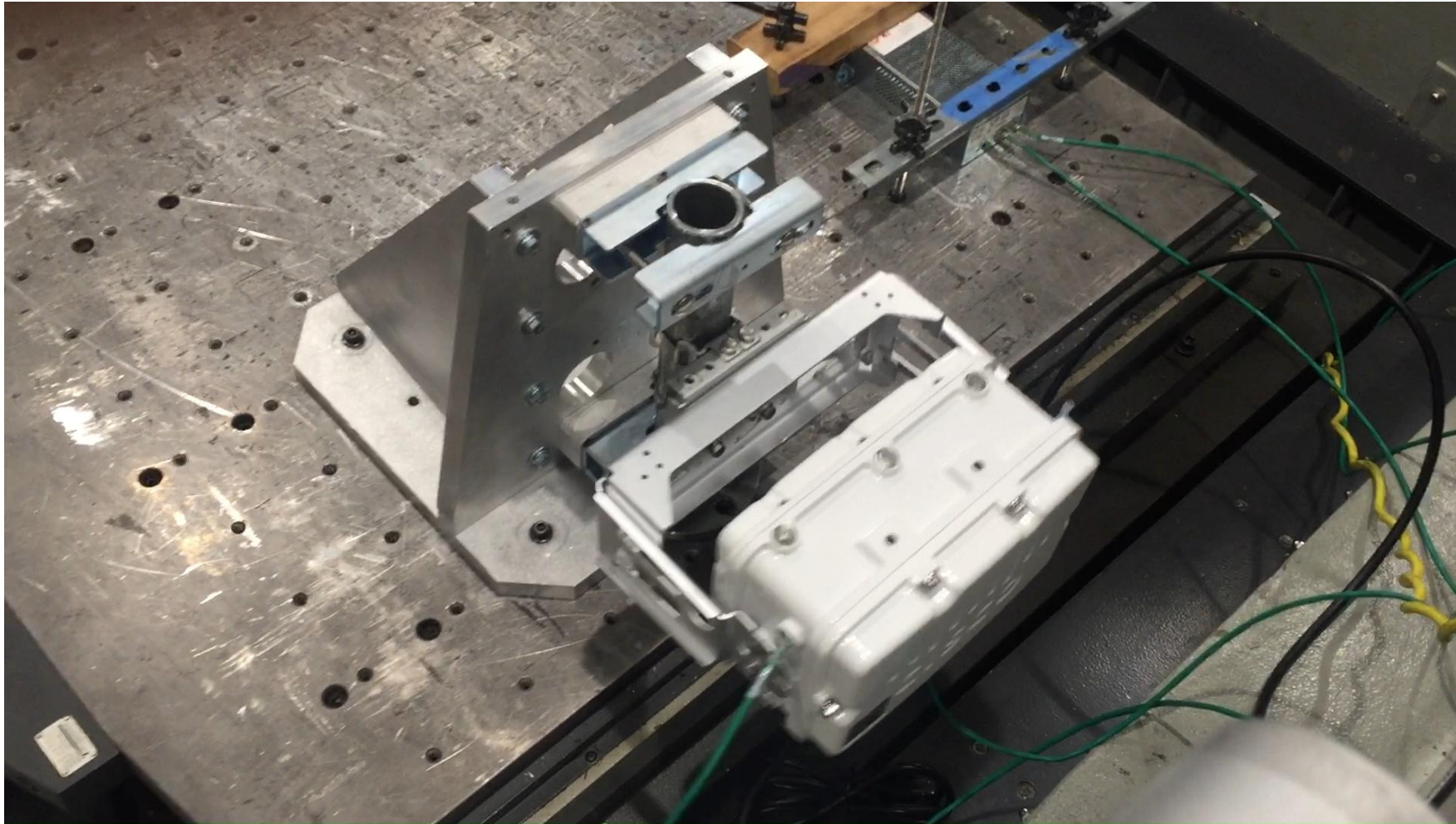
IE3400H Heavy-Duty switch

IE2000 Submersion demo

- Pressures vary w/ depth, time.
- No rubber duckies were drowned during this testing.



1552 Vibration Testing



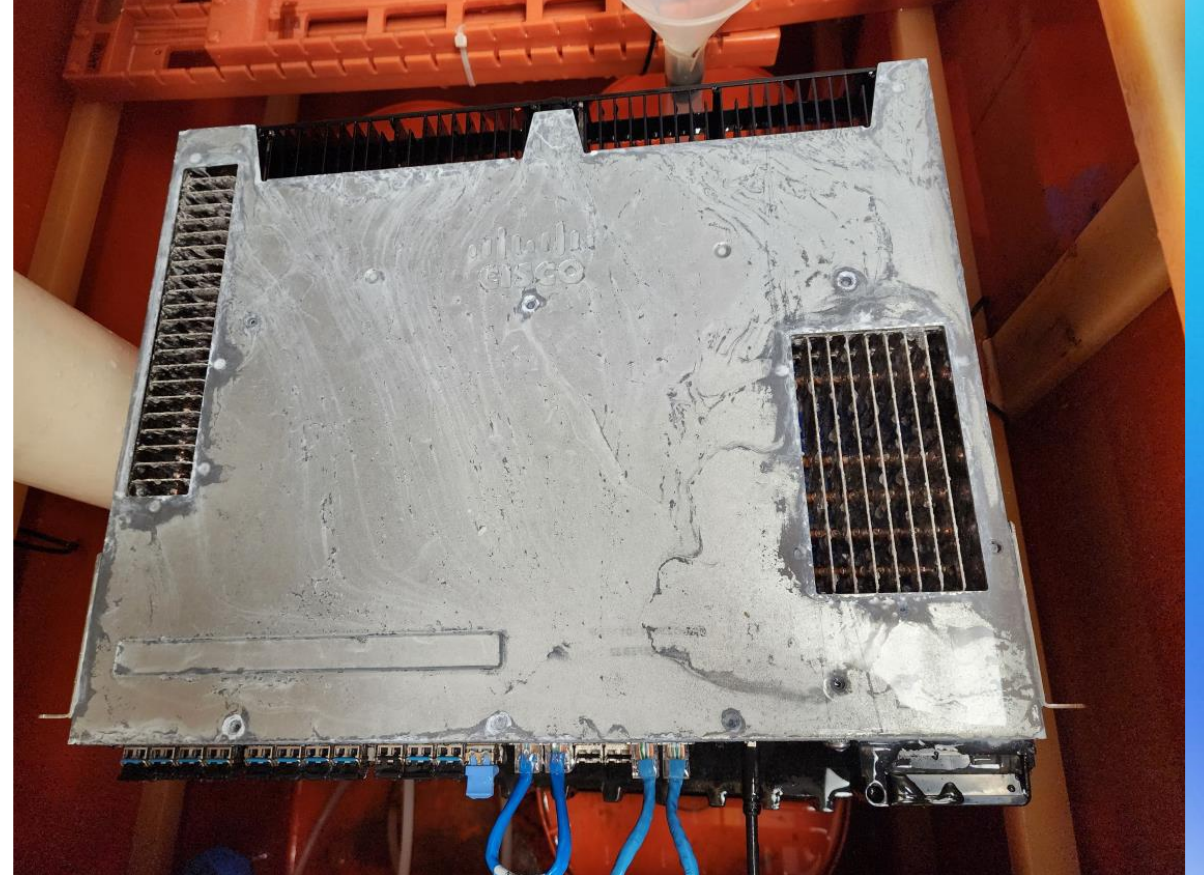
1552 Dust Test – Part of IP67 Certification



1552 Hosedown Testing



IE9K Salt Fog



Thermals

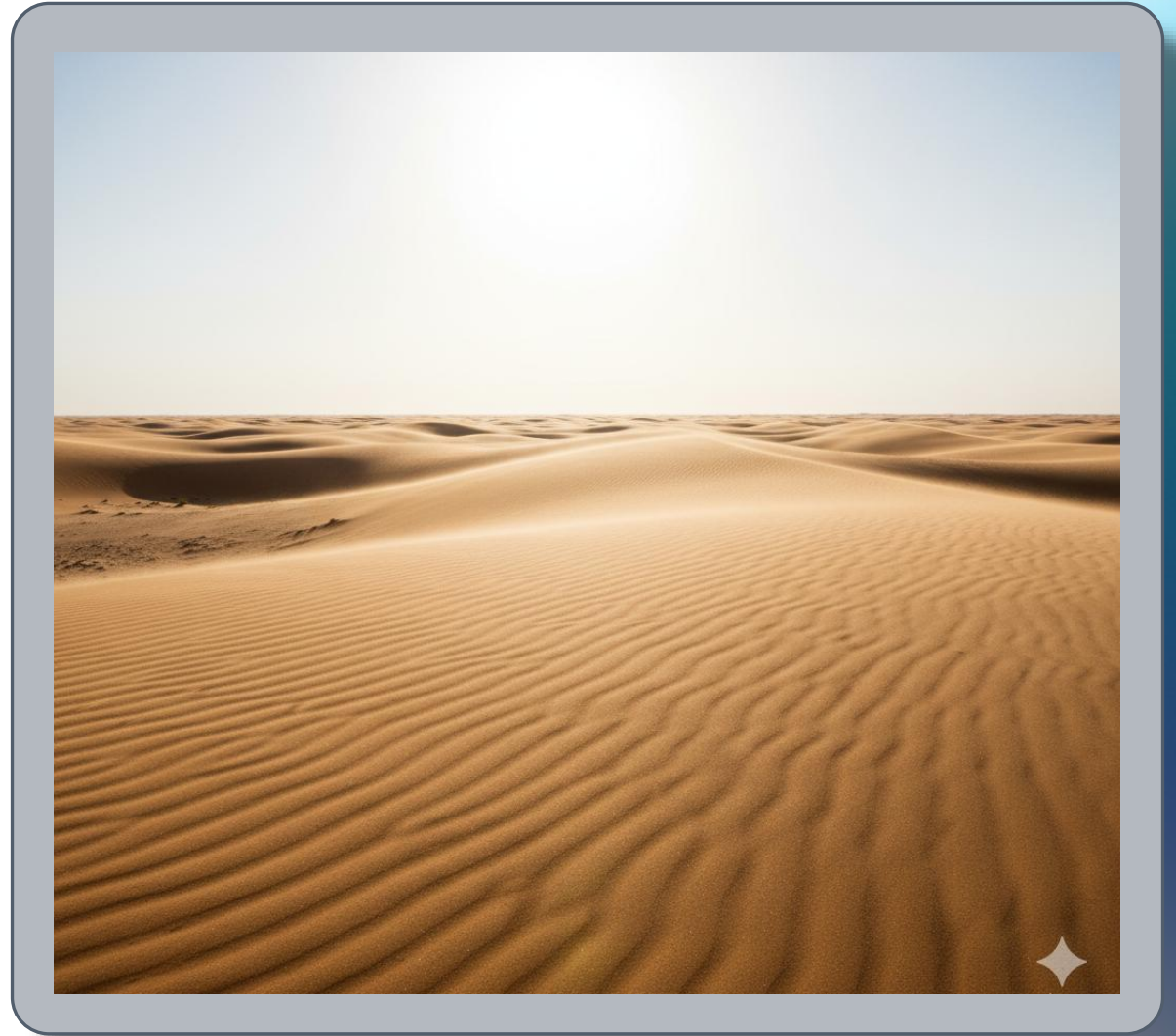
Some Physics (sorry...)

- Heat transfer equation
- Fan laws
- Thermal conductivity
- Ohm's Law



Forced-air cooling

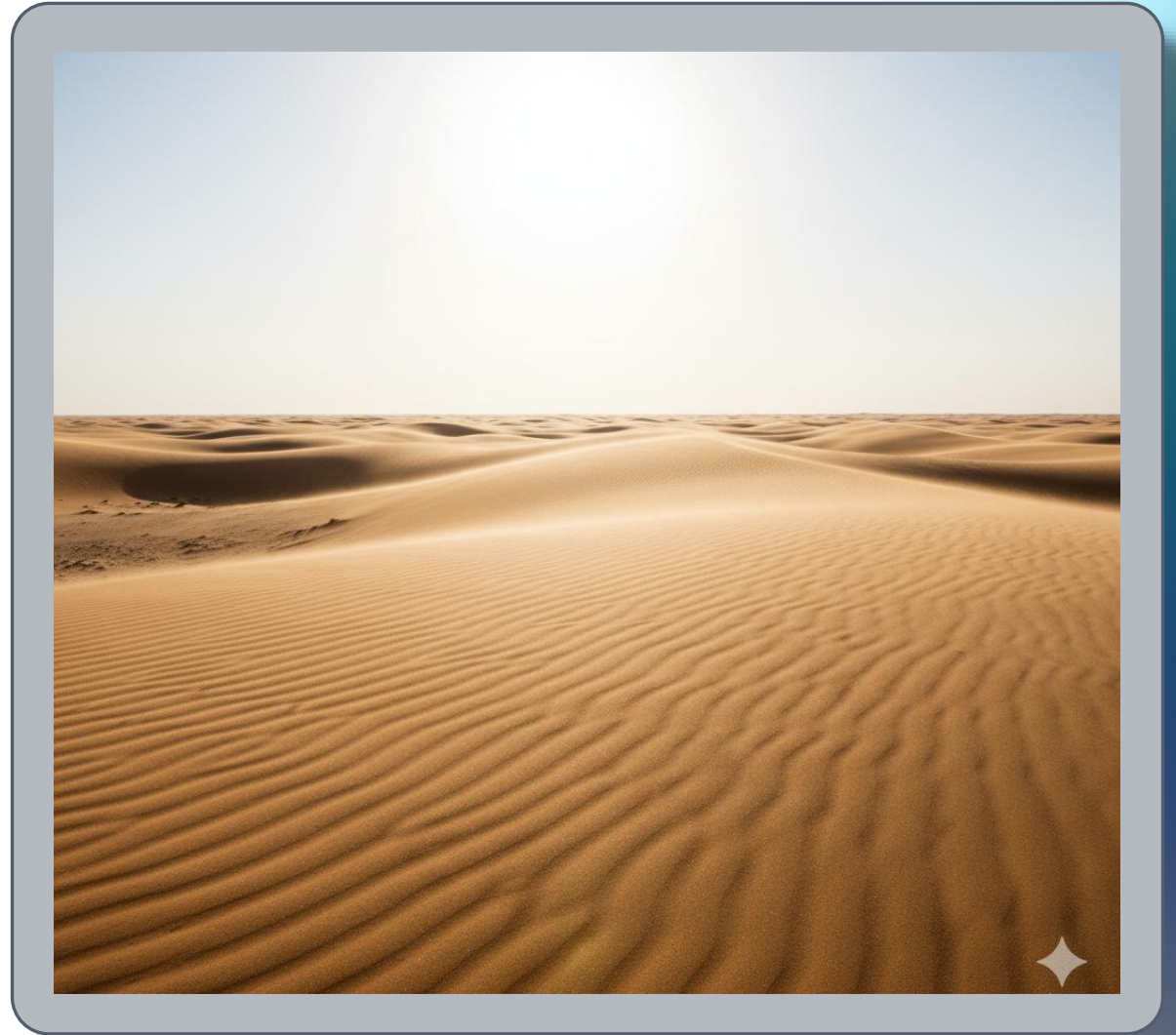
Key inputs



Forced-air cooling

Key inputs

inlet temperature

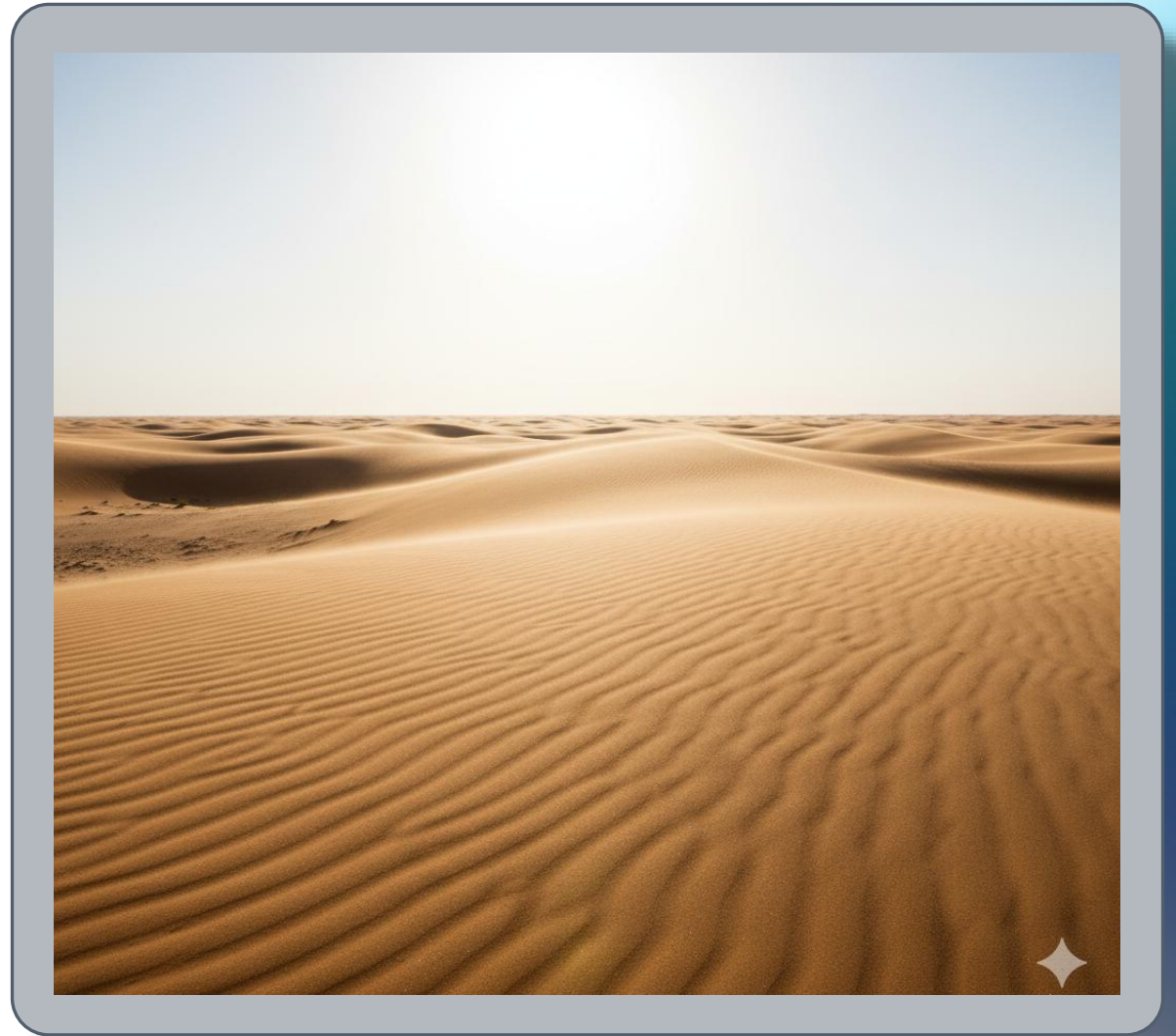


Forced-air cooling

Key inputs

inlet temperature

fluid (air) flow rate



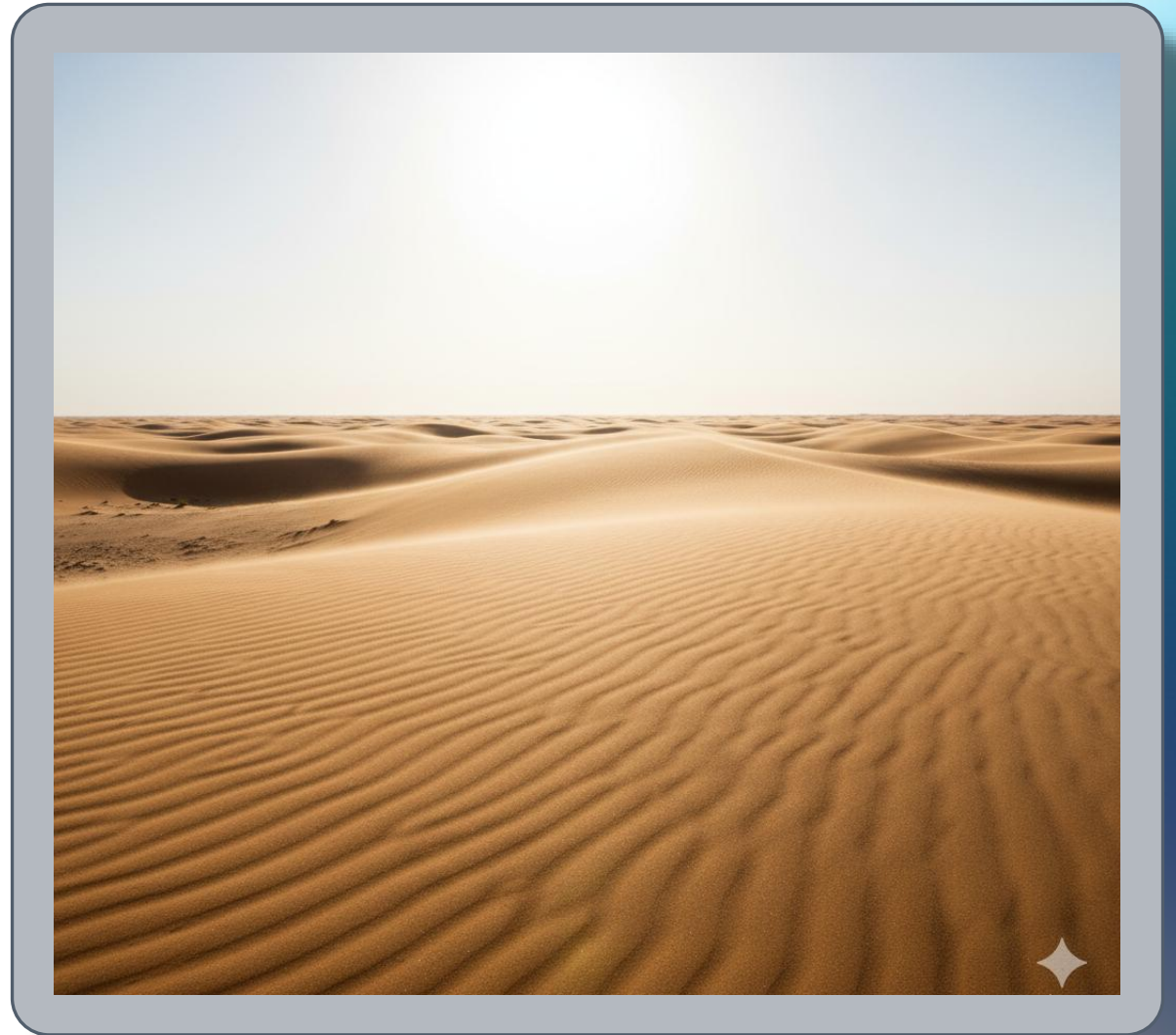
Forced-air cooling

Key inputs

inlet temperature

fluid (air) flow rate

surface area



Forced-air cooling

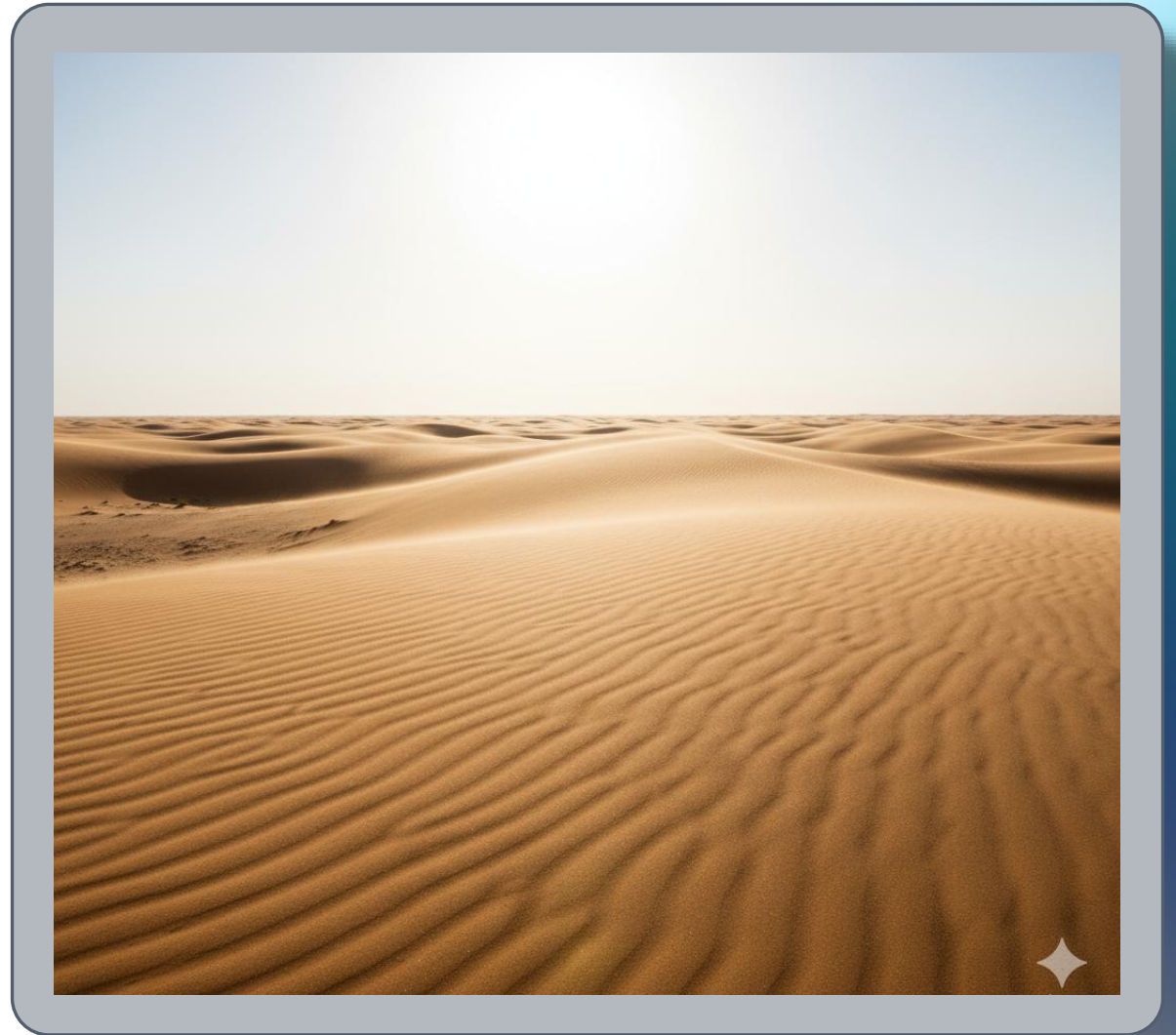
Key inputs

inlet temperature

fluid (air) flow rate

surface area

thermal conductivity



Liquid Cooling

Can anyone tell me what the key inputs are?



Liquid Cooling

Key inputs

inlet temperature

fluid (air) flow rate

surface area

thermal conductivity



Liquid Cooling

Key inputs

inlet temperature

(sort of?)

fluid (air) flow rate

(less so...)

surface area

(still yes...)

thermal conductivity

(still very much yes...)



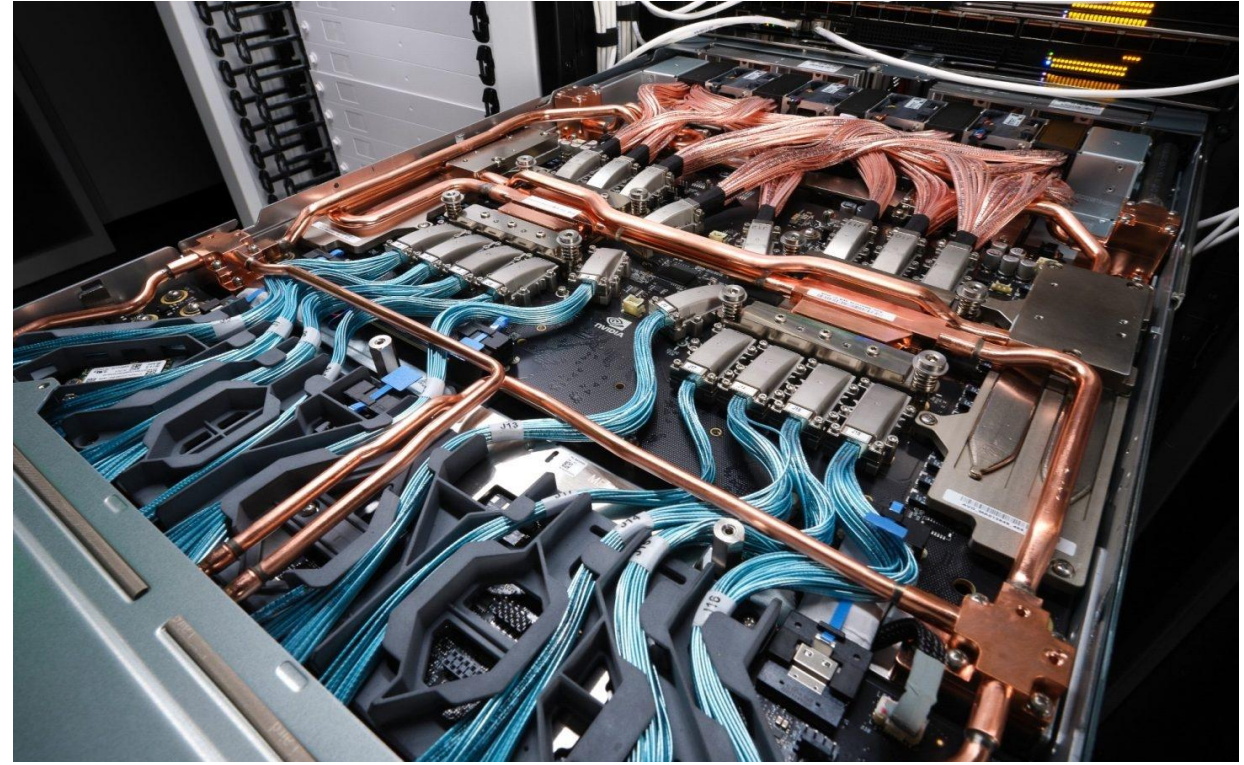
Liquid cooling word cloud

- CDU
- Sidecar
- End-of-row
- In-rack
- {Water, Power, Carbon} usage efficiency
- Fittings, couplers, manifolds, feeds, returns

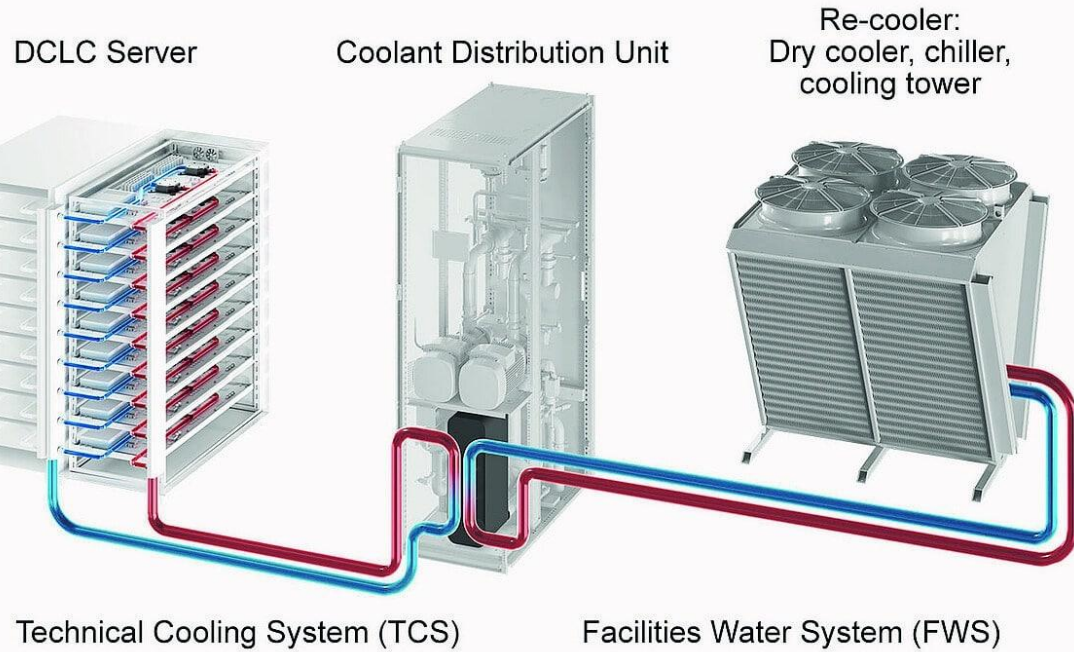


Immersion vs. Direct

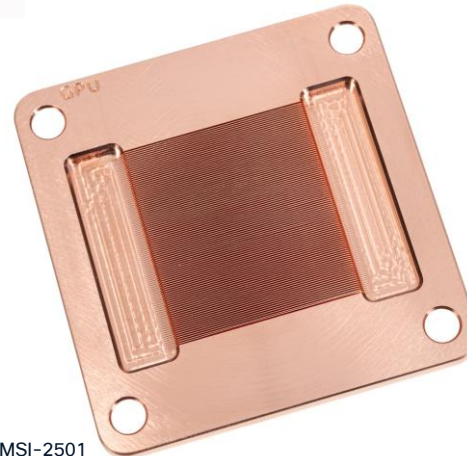
- Immersion was popular for crypto, some x86
- OCP has converged on single-phase Direct-to-Chip for next gen of GPUs



Key Bits



- Leak detection
- Flow sensors
- BMC control
- Secondary loop fluid (chemical engineers?)
- QD vs. Direct Mate



BRKMSI-2501



Show and Tell

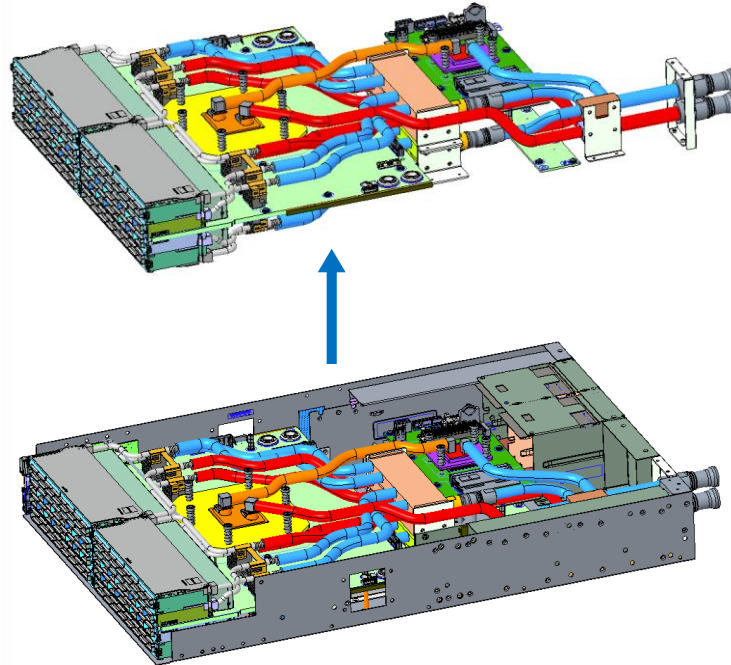
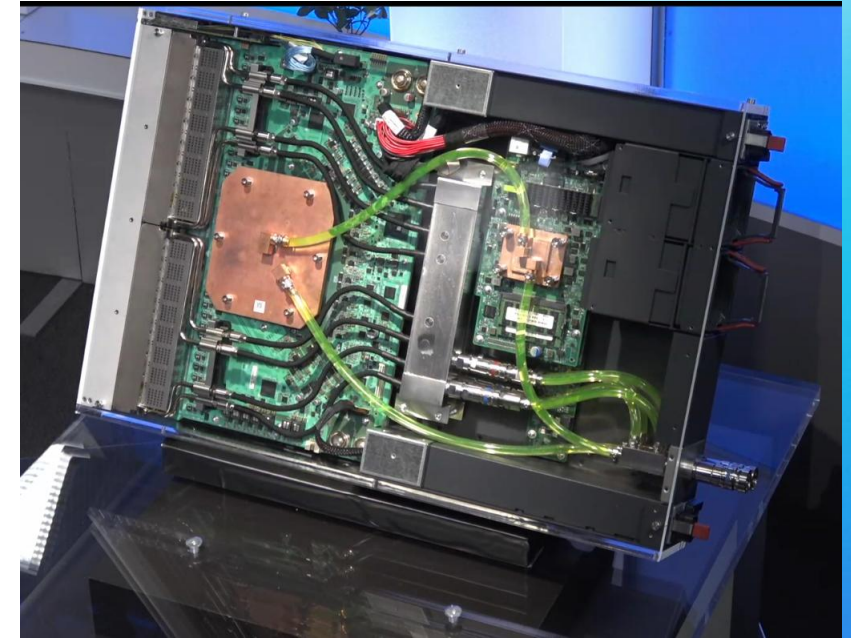
CiscoLive 51.2T Demo June 2025

OFC 51.2T Demo March 2025



Key Features

- Liquid cooled components: G200 ASIC, CPU, 64xOSFP 800G optics
- Up to 90% of system heat is removed via liquid cooling
- Future proof design to support up to 40W of optics and over 1KW of NPU
- Compact 2RU platform



Key Features

- Full switch passing traffic
- Liquid cooled
 - NPU
 - Optics
 - CPU
- Industry standard UQD connectors

Board design & Signal Integrity – Part I

Board design

NOT THIS.

At least, not for us. ;-)



Board design

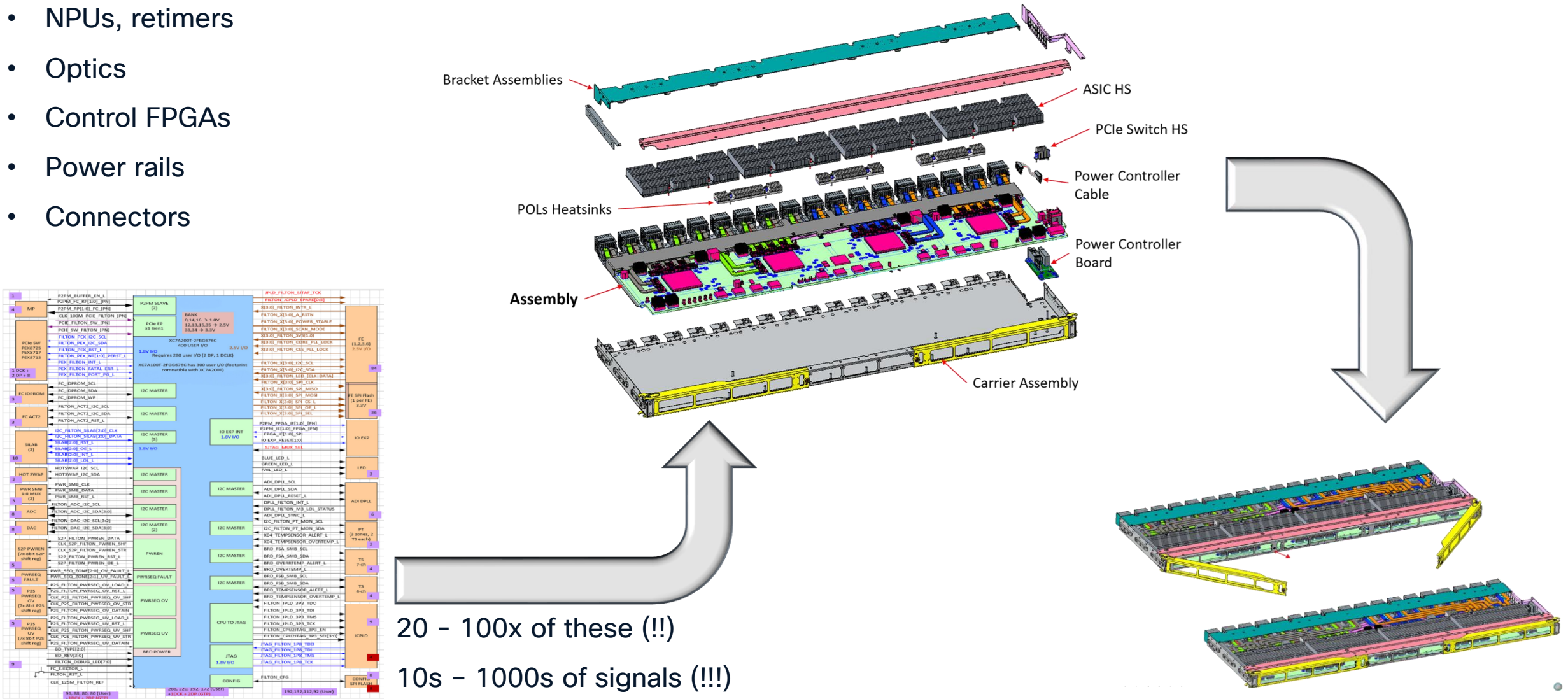
Also, NOT THIS.

But it's a little closer to the truth...



Board design : complexity and scale

- NPU, retimers
- Optics
- Control FPGAs
- Power rails
- Connectors



Why is this so hard? -- “Art” vs. “science”



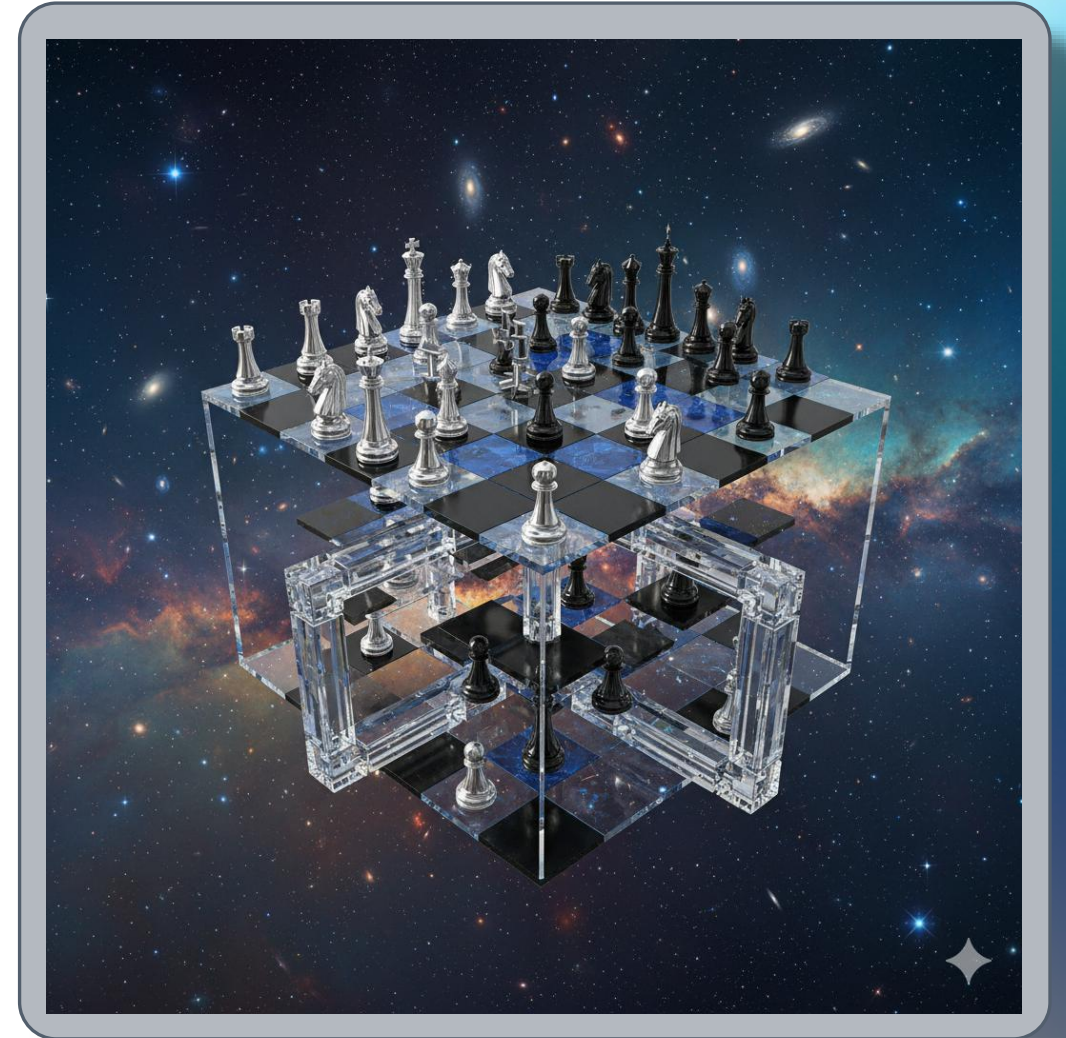
Why is this “hard” ?

- “Art” v. “Science”
- LOTS of competing priorities.
- Physics and space conspire against you.
- Each decision impacts others



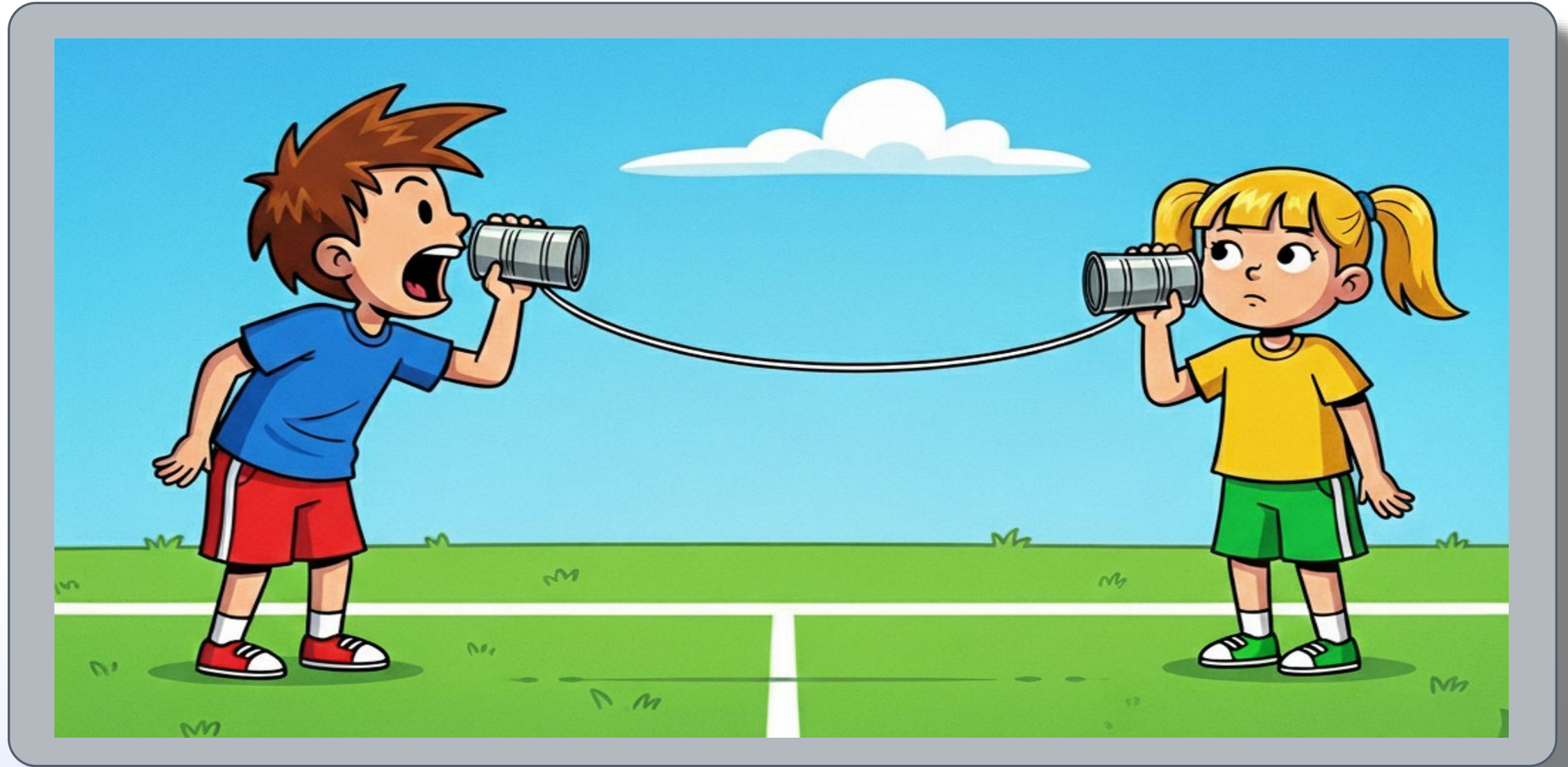
3-D chess: Making everything “fit”

- There are lots and lots of pieces
- They all want many of the same things
- Very high inter-dependence



Board design & Signal Integrity – Part II

What is signal integrity, anyway?



What is signal integrity, anyway?

NOT THIS.

At least, not for us. ;-)

Wave equation

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

Not to be confused with [Wave function](#).

The **wave equation** is a second-order linear [partial differential equation](#) for the description of [waves](#) or [standing wave](#) fields such as [mechanical waves](#) (e.g. [water waves](#), [sound waves](#) and [seismic waves](#)) or [electromagnetic waves](#) (including [light waves](#)). It arises in fields like [acoustics](#), [electromagnetism](#), and [fluid dynamics](#).

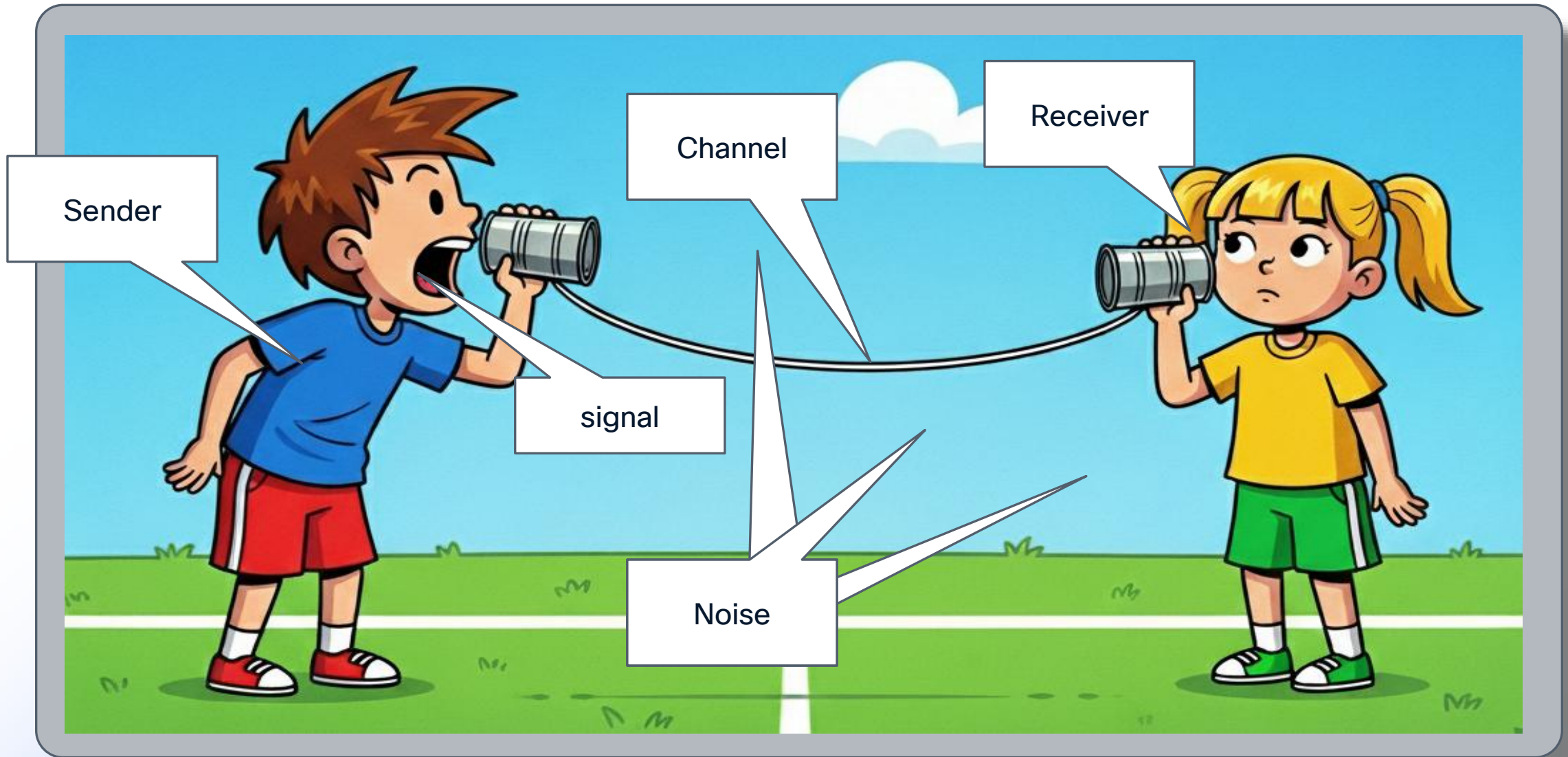
This article focuses on waves in [classical physics](#). Quantum physics uses an operator-based [wave equation](#) often as a [relativistic wave equation](#).

What is signal integrity, anyway?

WHEW. This is more like it.

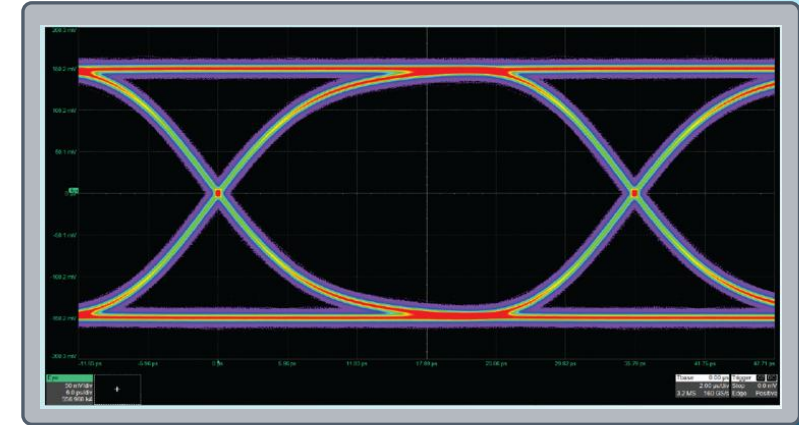
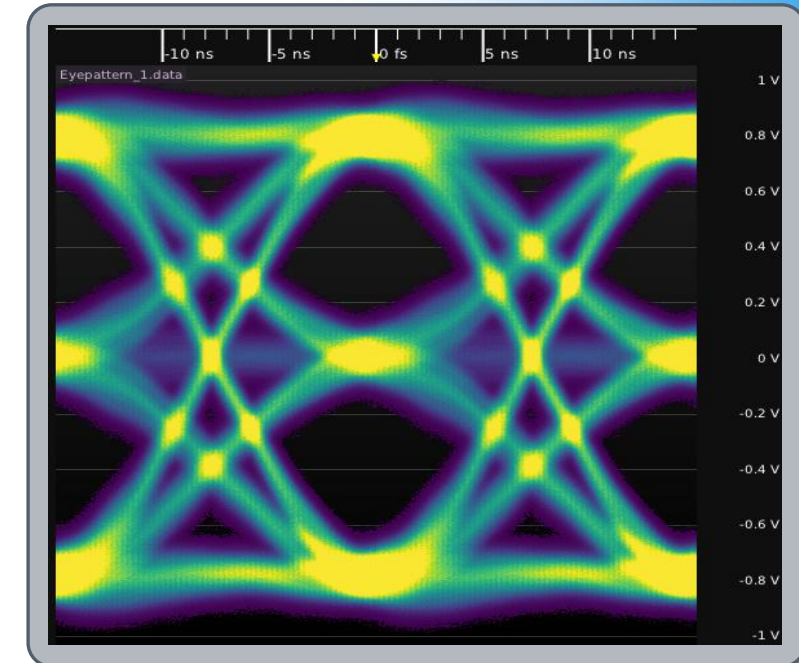


What is signal integrity, anyway?



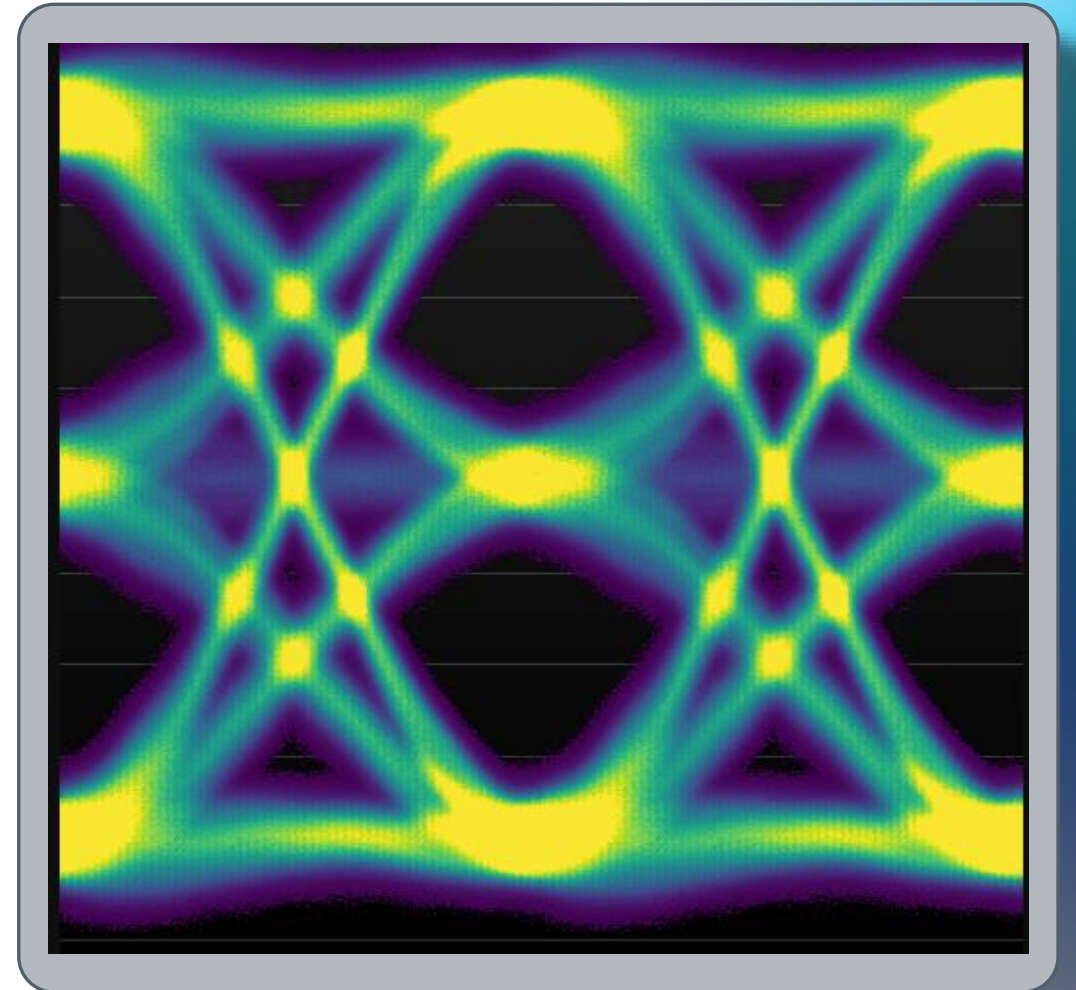
How do we measure / quantify SI?

- Eye diagrams.
- How many symbol levels?
- How many bits is this?
- “Open” v. “closed” eyes
- Time, power, noise shifts...

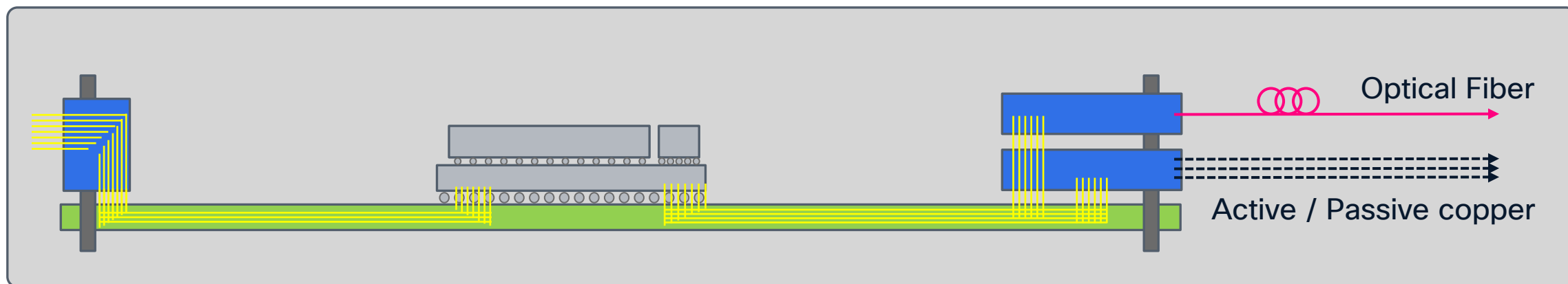
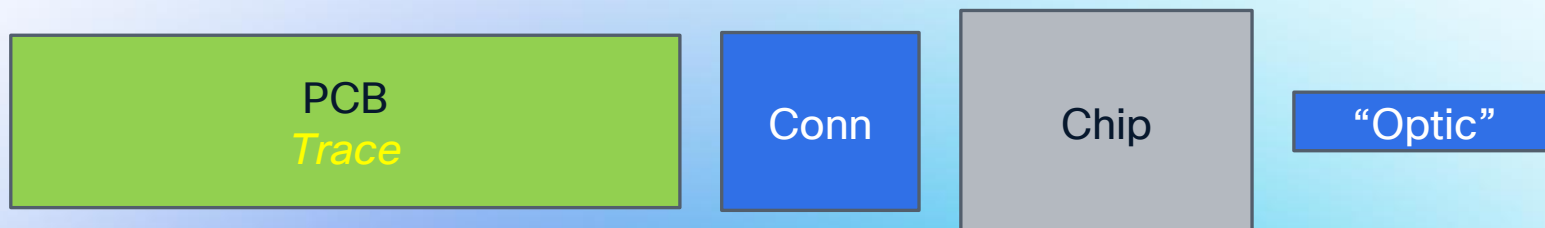


Improving signal integrity...

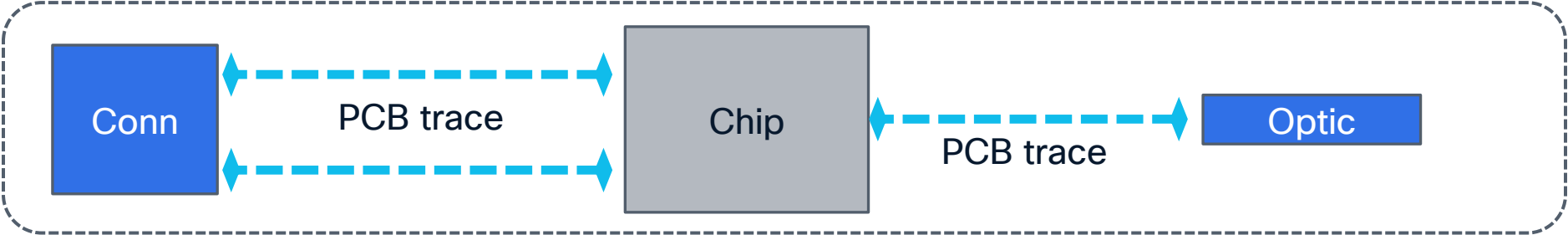
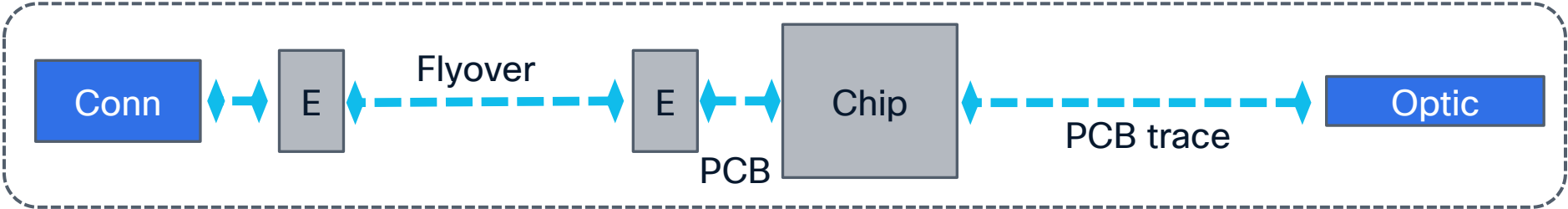
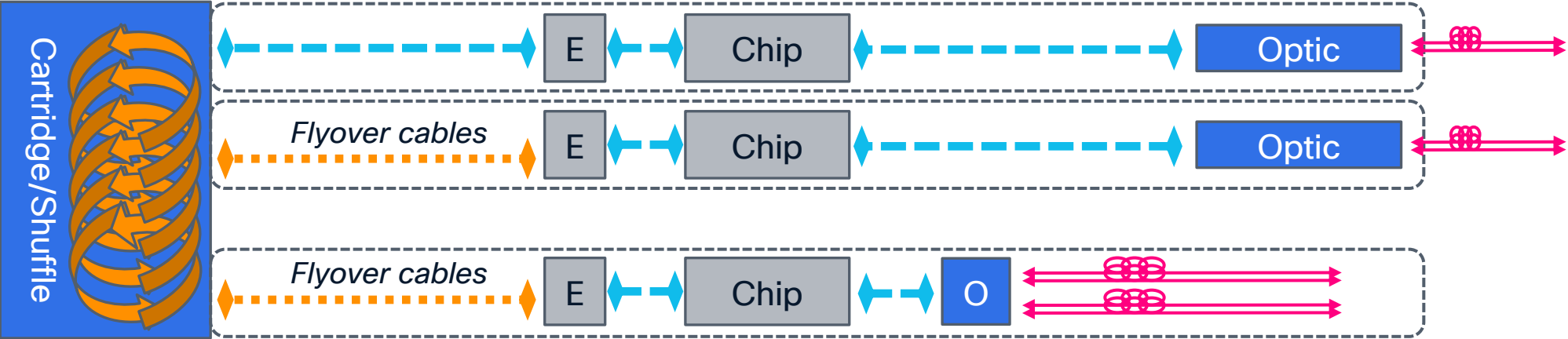
1. Shout louder.
2. Talk slower.
3. Get closer.
4. Get better string.
5. Remove background noise.



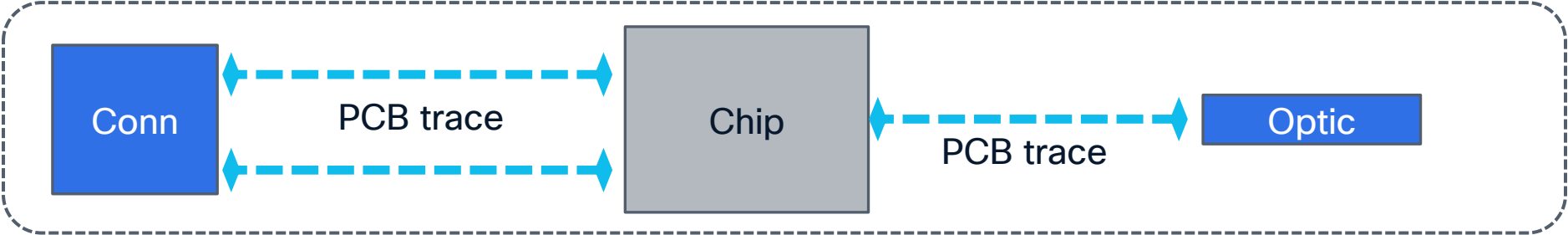
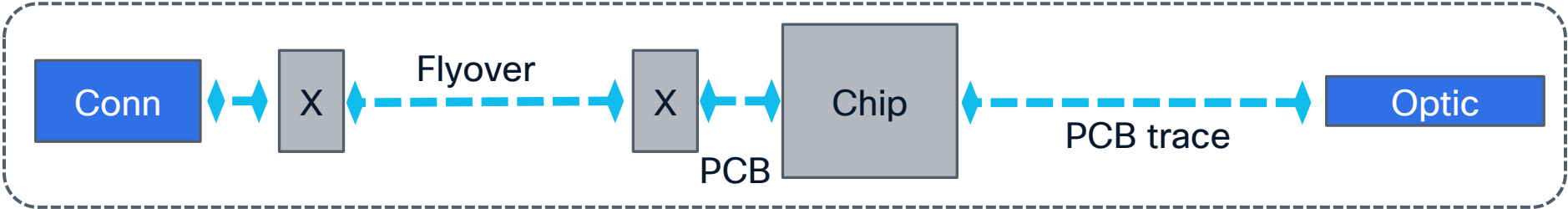
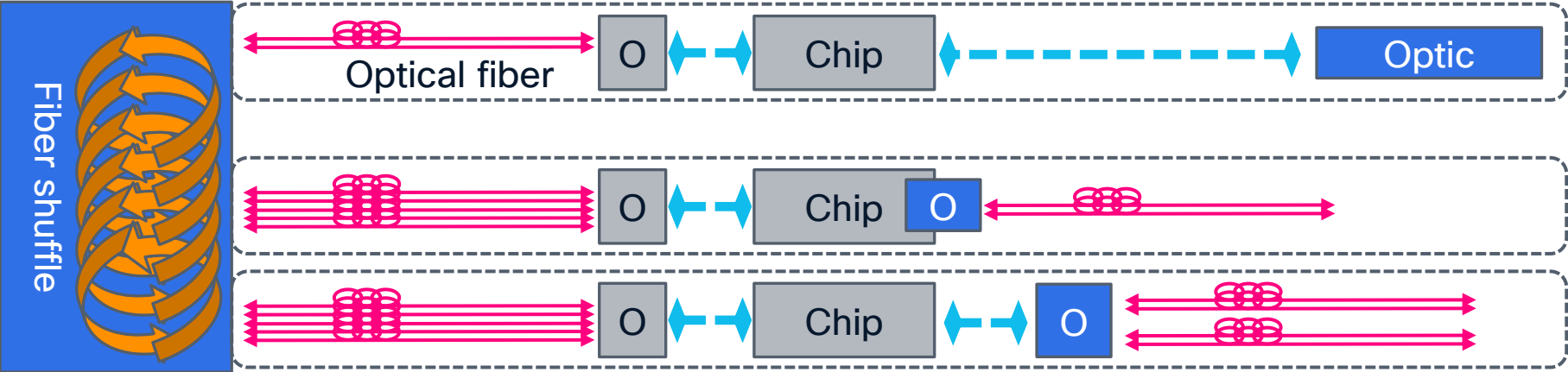
A Very Excellent Generic Board Side View Schematic



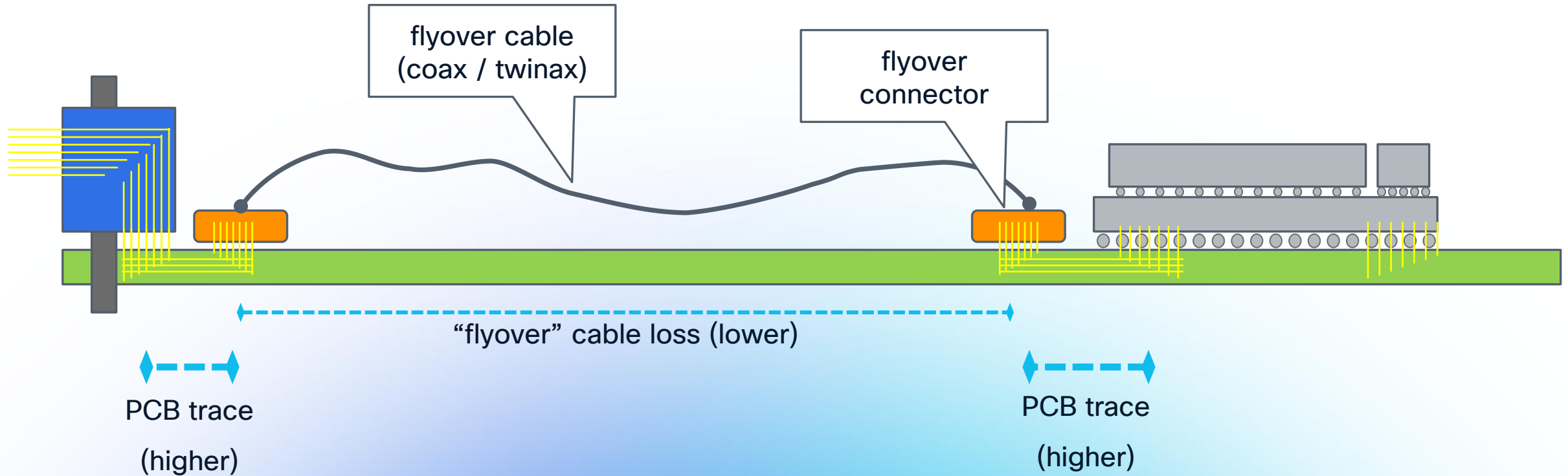
A bunch of ways to solve the problem



A bunch of ways to solve the problem

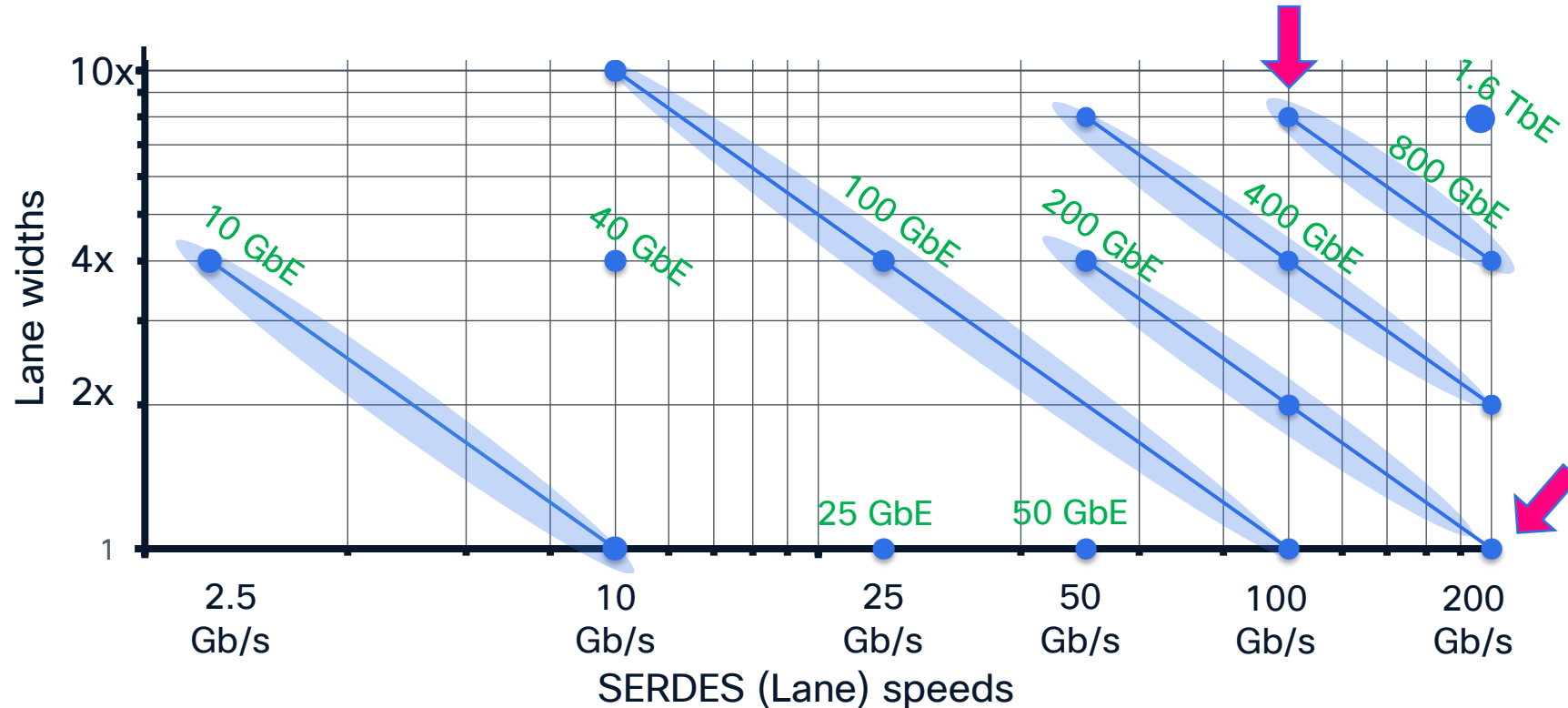


A bunch of ways to solve the problem



Optics

Evolution and reuse of underlying optics / serdes tech

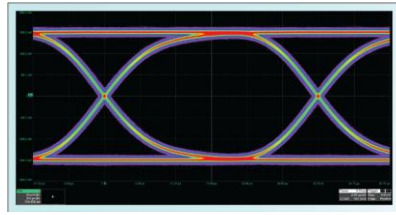


- (almost) nothing is new or only used once
- 8x100 “today”
- Nx200 “tomorrow”
- Requires (yet more) integration
- (yet) higher mechanical and thermal densities

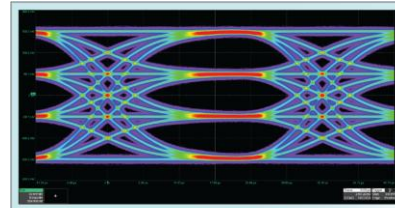
Fundamental technologies for high-speed optics magic

Advanced Modulation

Client optics (IMDD)

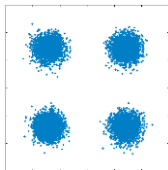


NRZ
1 bit/sec/symbol

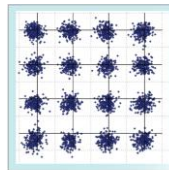


PAM4
2 bit/sec/symbol

Coherent optics



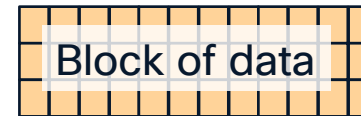
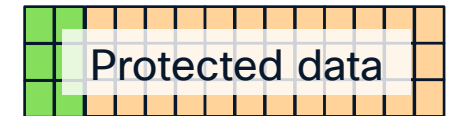
Coherent QPSK
2 bit/sec/symbol



Coherent 16-QAM
4 bit/sec/symbol

Using more complex modulation allows us to increase the data rate (Gb/s) without increasing the signaling speed (Gbaud)

Forward Error Correction (FEC)



Allows correction of errors at receiver

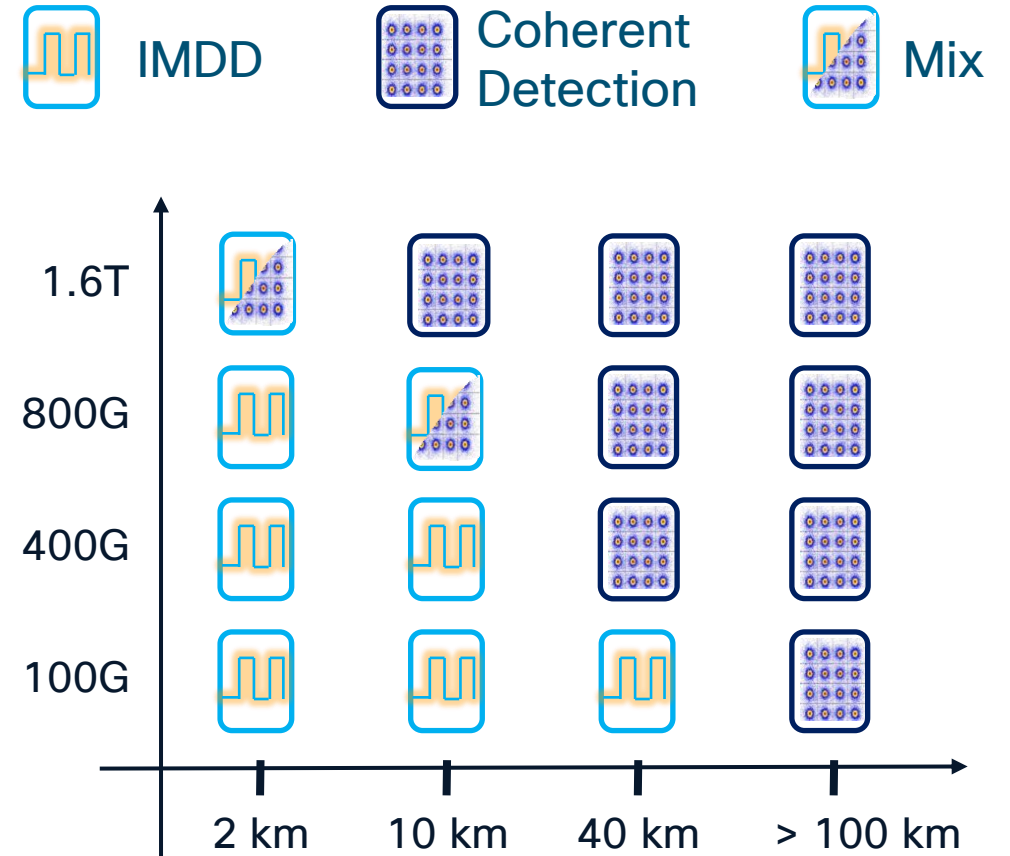
Enables use of relaxed specs (saves \$) to get same performance or enables much higher performance.

Usually embedded in Ethernet switch or module DSP ASIC

Some overhead (faster data rate needed)

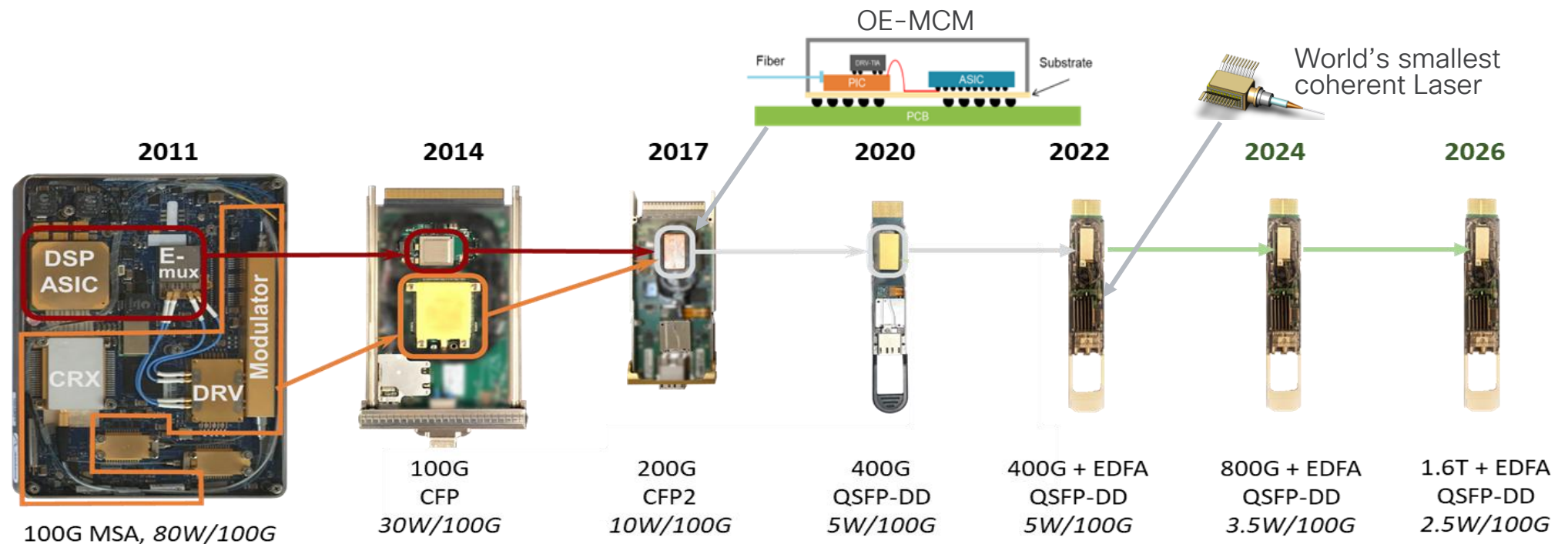
Coherent detection allows greater reach × speed

- Both IMDD and Coherent technology will continue to be used
 - 800G, 1.6T and beyond
 - Coherent pushing towards shorter reaches. Not only in DWDM but also Grey applications
 - Focus shifting from performance enhancements (\$\$\$) to interoperable interfaces and pluggables
- Coherent multivendor Interop more and more prevalent
 - 100G, 400ZR/ZR+, 800ZR/ZR+



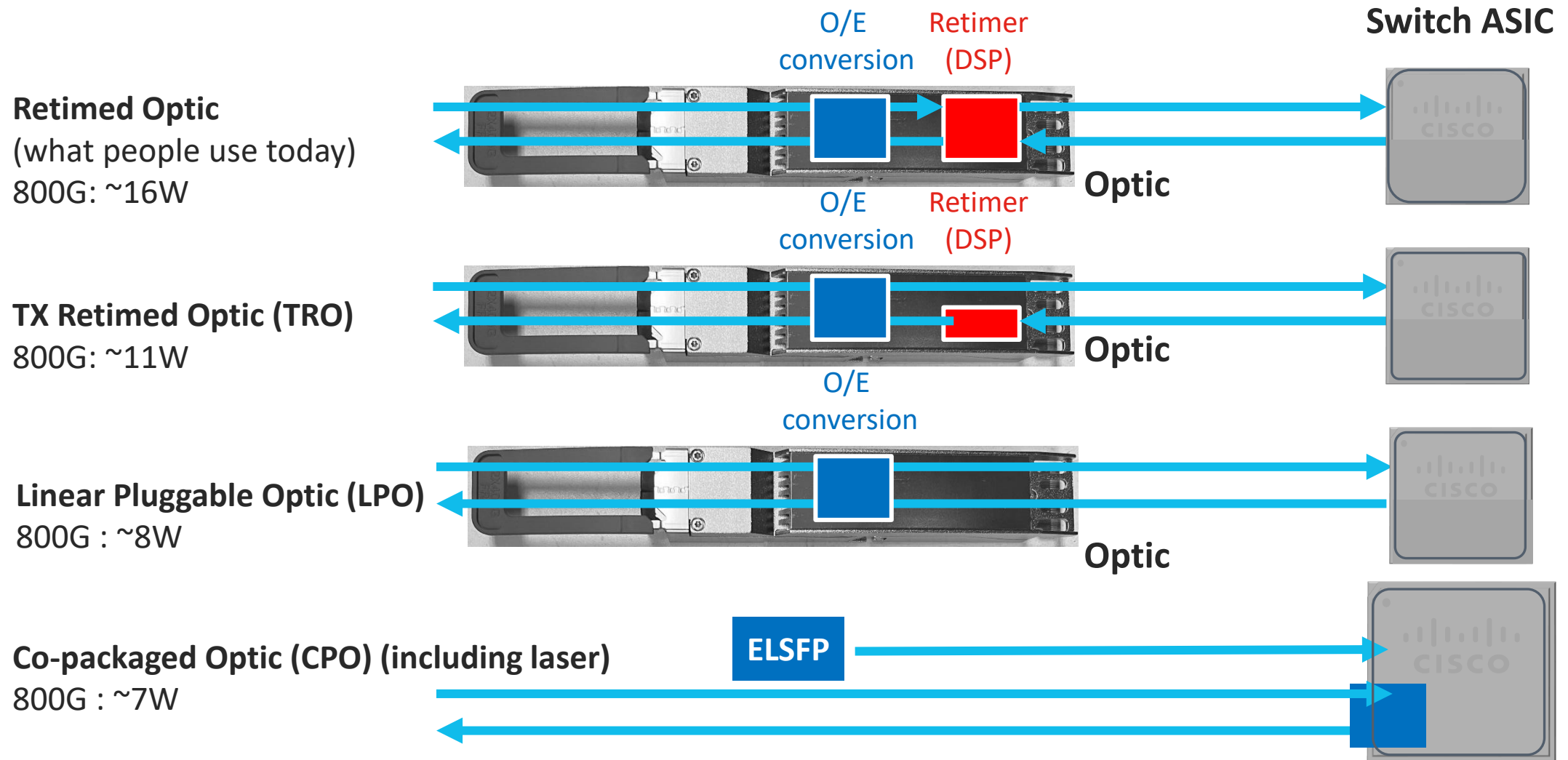
As speed increases, “crossover distance” between IMDD and coherent decreases

Increasing component integration

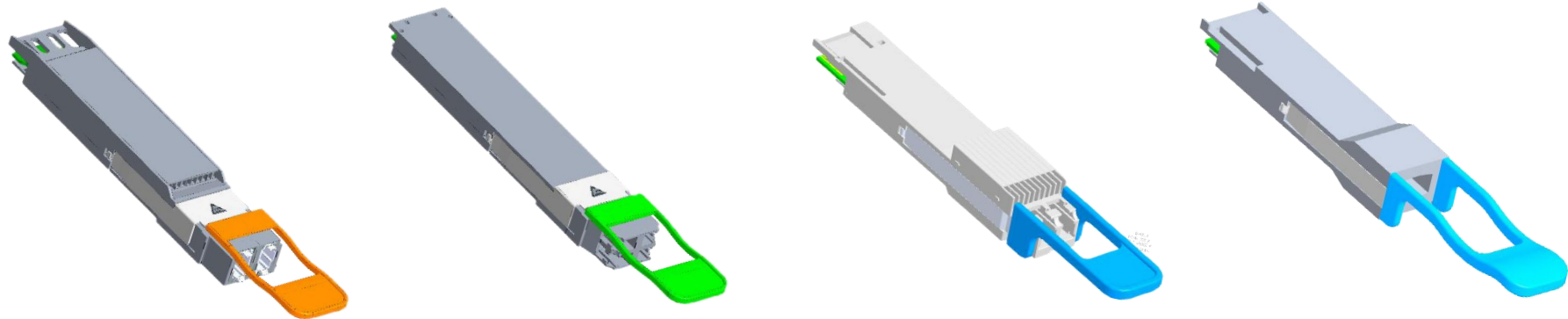


	In-House: DSP + SiPh Optics + OE-MCM + ASIC + Laser						100% Vertical
QAM:	QPSK	QPSK	16QAM	16QAM	PCS	PCS	
CMOS:	40nm	28nm	16nm	7nm	4/3nm	2nm	
Optics:	Discrete, InP, LiNbO3	SiPh	SiPh	SiPh	SiPh	SiPh/TFLN	

Follow the bouncing functionality!



Pluggable Optical Modules Options



OSFP

OSFP-RHS

QSFP-DD

QSFP

Host high-speed pins	8	8	8	4
Module capacity (Gbit)	400, 800, 1600	400, 800, 1600	400, 800, 1600	40, 100, 200, 400
Cooling	Integrated heatsink	Riding heatsink	Riding heatsink	Riding heatsink
High-speed IO Compatibility (Gbps)	50-100-200	50-100-200	50-100-200	1-25-50-100
Port compatibility	OSFP	OSFP-RHS	QSFP-DD & QSFP	

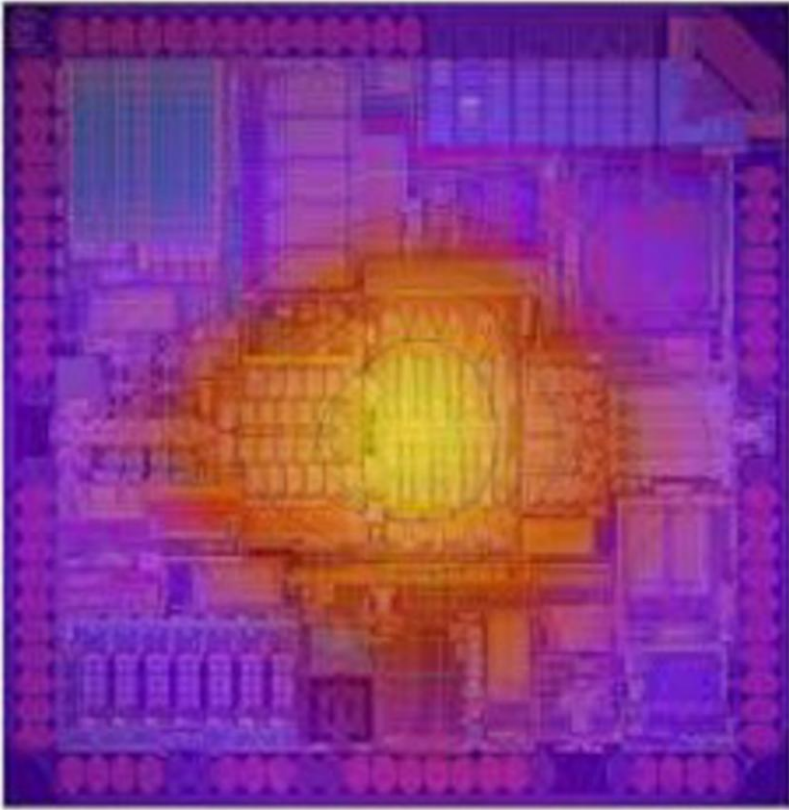
Hot Chips!

How hot is too hot?

- Incandescent bulb (filament) : ~2500C
- Incandescent bulb (glass) : ~250C
- Modern ASIC/NPU : ~100C
- Modern optics / lasers : ~75C
- People fingers : ~55C

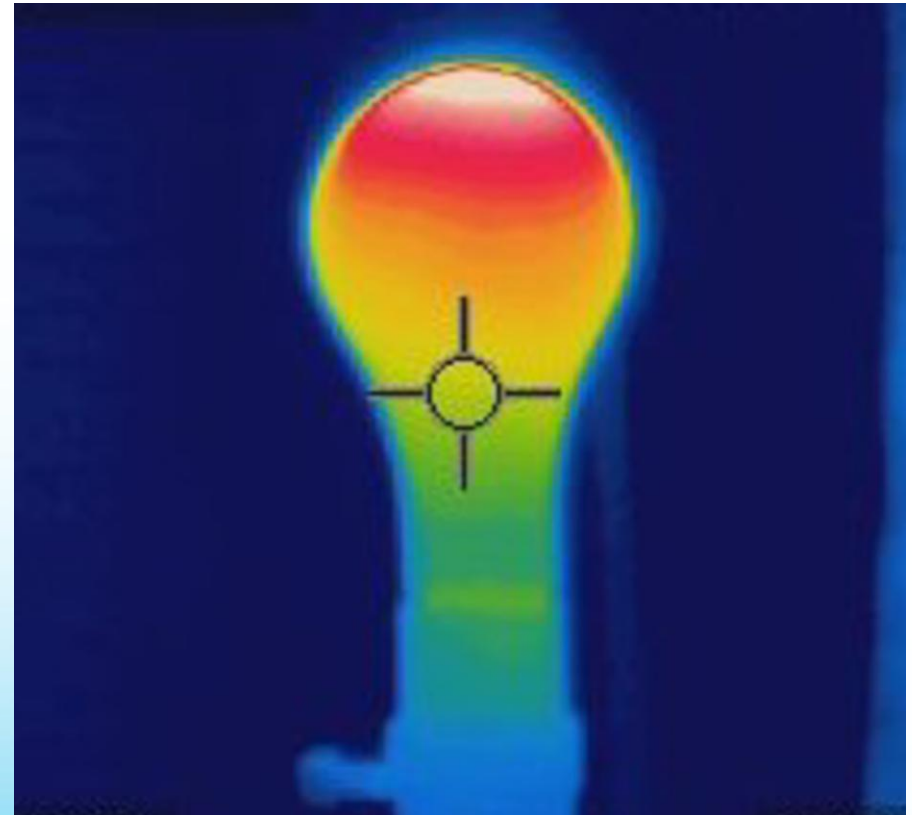


(Is it) Hot or Not?



ca. 2025 :
50 Tbps networking ASIC

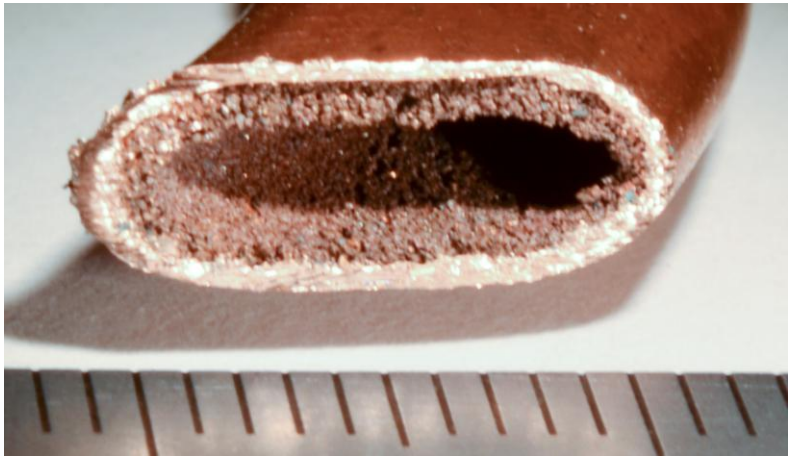
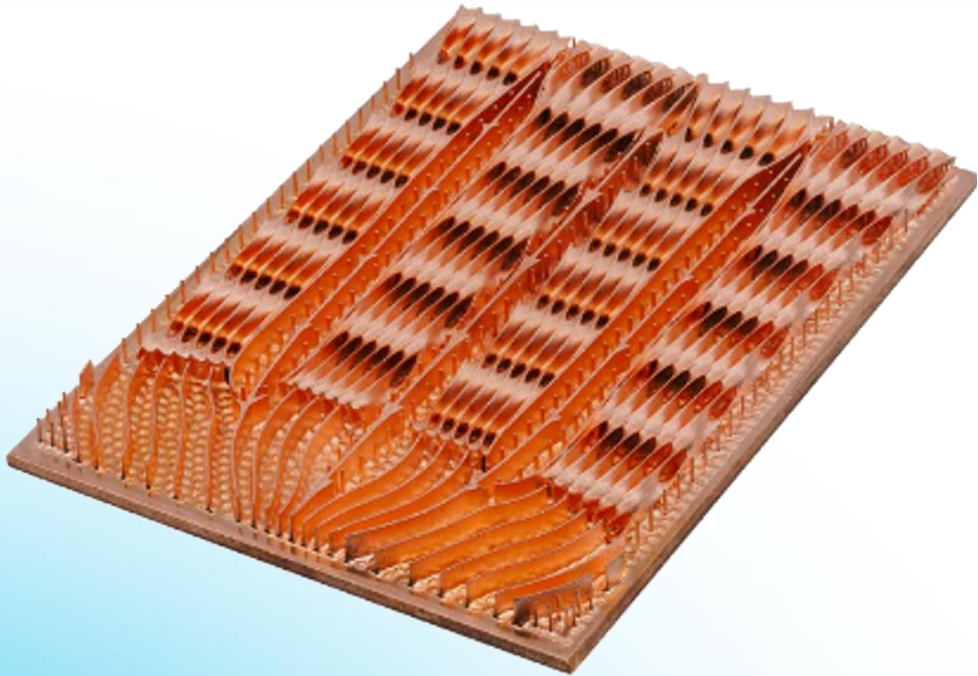
VS.



ca. 1925 :
100W incandescent bulb

Heat sinks, heat pipes

Light bulb			Chip		
area	0.012271846	sq m	area	0.0025	sq m
radius	0.03125	meters	radius		meters
power	100	watts	power	1250	watts
flux	8,148.73	W / sq m	flux	500,000.00	W / sq m
Sun			Induction Hob		
area	6.08735E+18	sq m	area	0.237582944	sq m
radius	696000000	meters	radius	0.275	meters
power	3.80E+26	watts	power	4000	watts
flux	62,424,526.00	W / sq m	flux	16,836.23	W / sq m



Cold plate w/ manifold and connectors



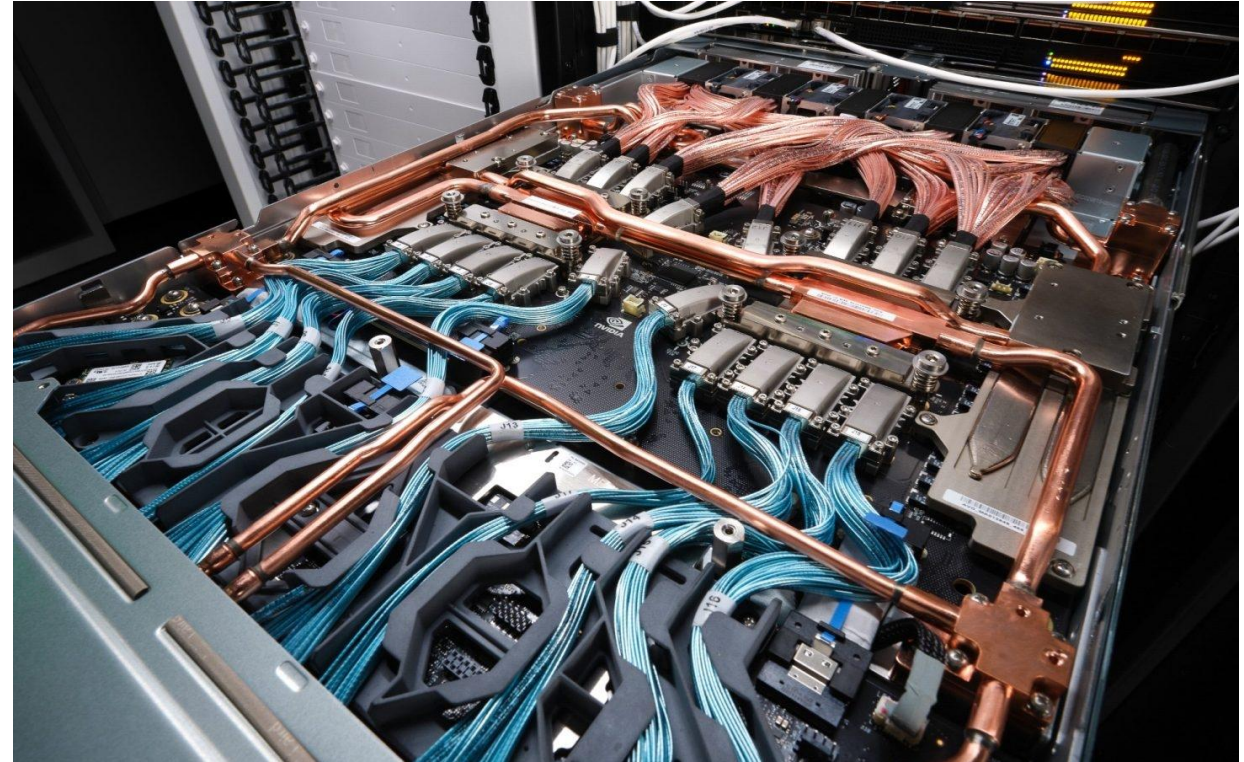
Liquid cooling word cloud

- CDU
- Sidecar
- End-of-row
- In-rack
- {Water, Power, Carbon} usage efficiency
- Fittings, couplers, manifolds, feeds, returns

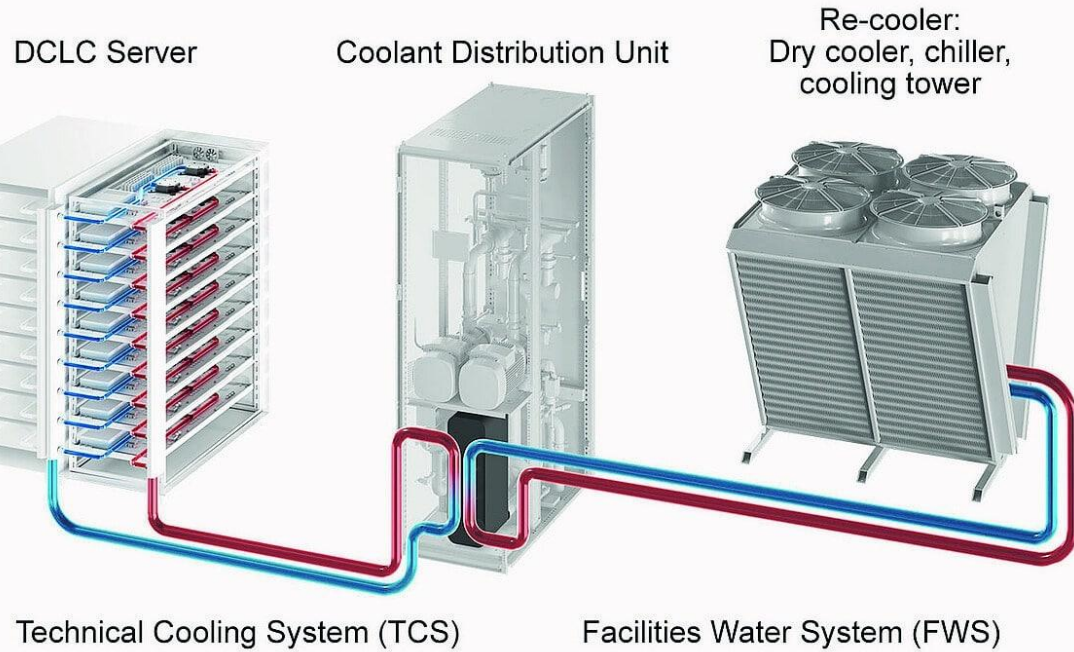


Immersion vs. Direct

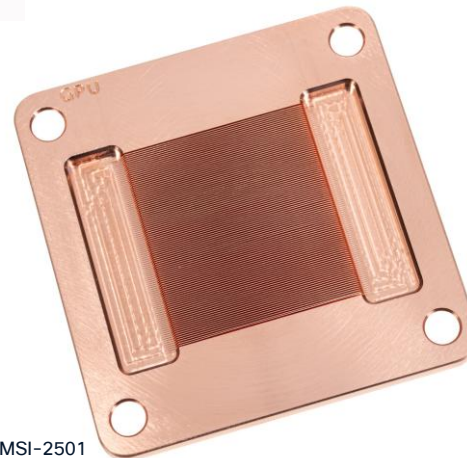
- Immersion was popular for crypto, some x86
- OCP has converged on single-phase Direct-to-Chip for next gen of GPUs



Key Bits



- Leak detection
- Flow sensors
- BMC control
- Secondary loop fluid (chemical engineers?)
- QD vs. Direct Mate



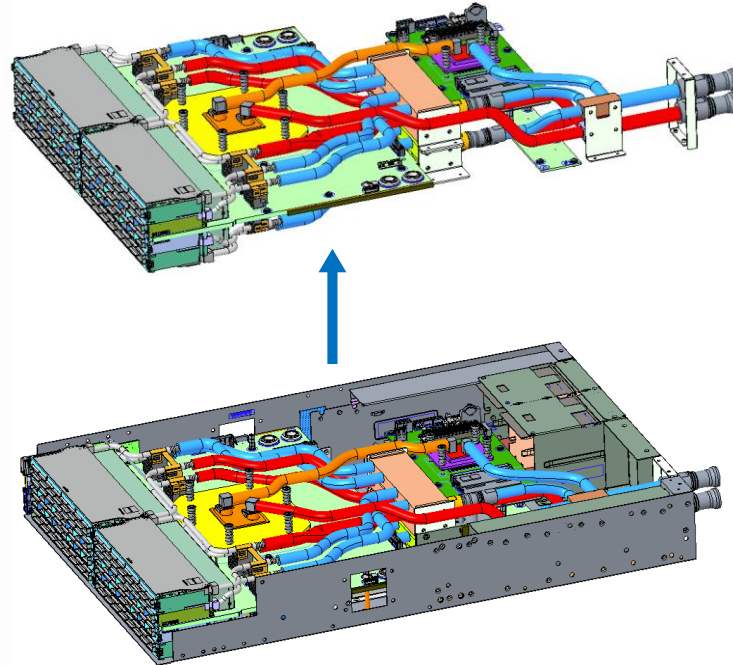
Trade Show Demos

OFC 51.2T Demo March 2025

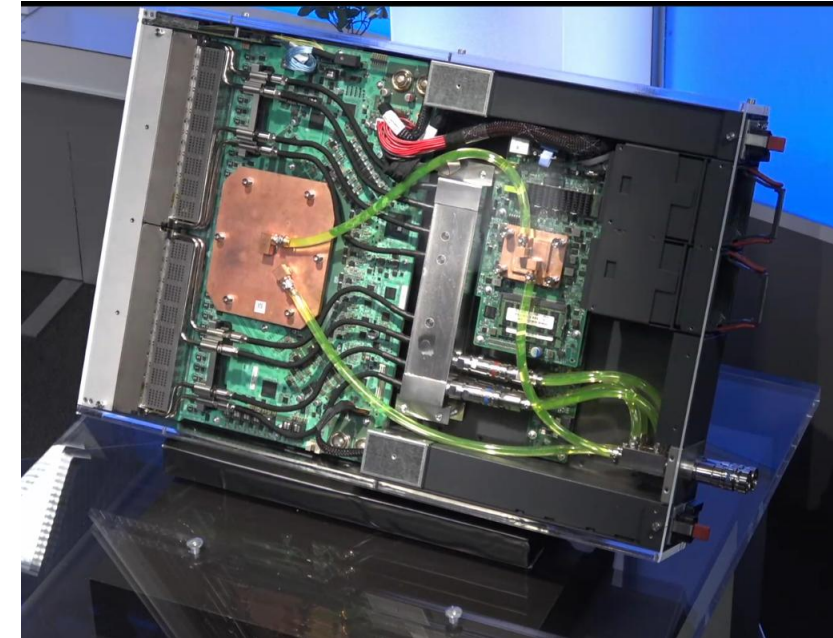


Key Features

- Liquid cooled components: G200 ASIC, CPU, 64xOSFP 800G optics
- Up to 90% of system heat is removed via liquid cooling
- Future proof design to support up to 40W of optics and over 1KW of NPU
- Compact 2RU platform



CiscoLive 51.2T Demo June 2025

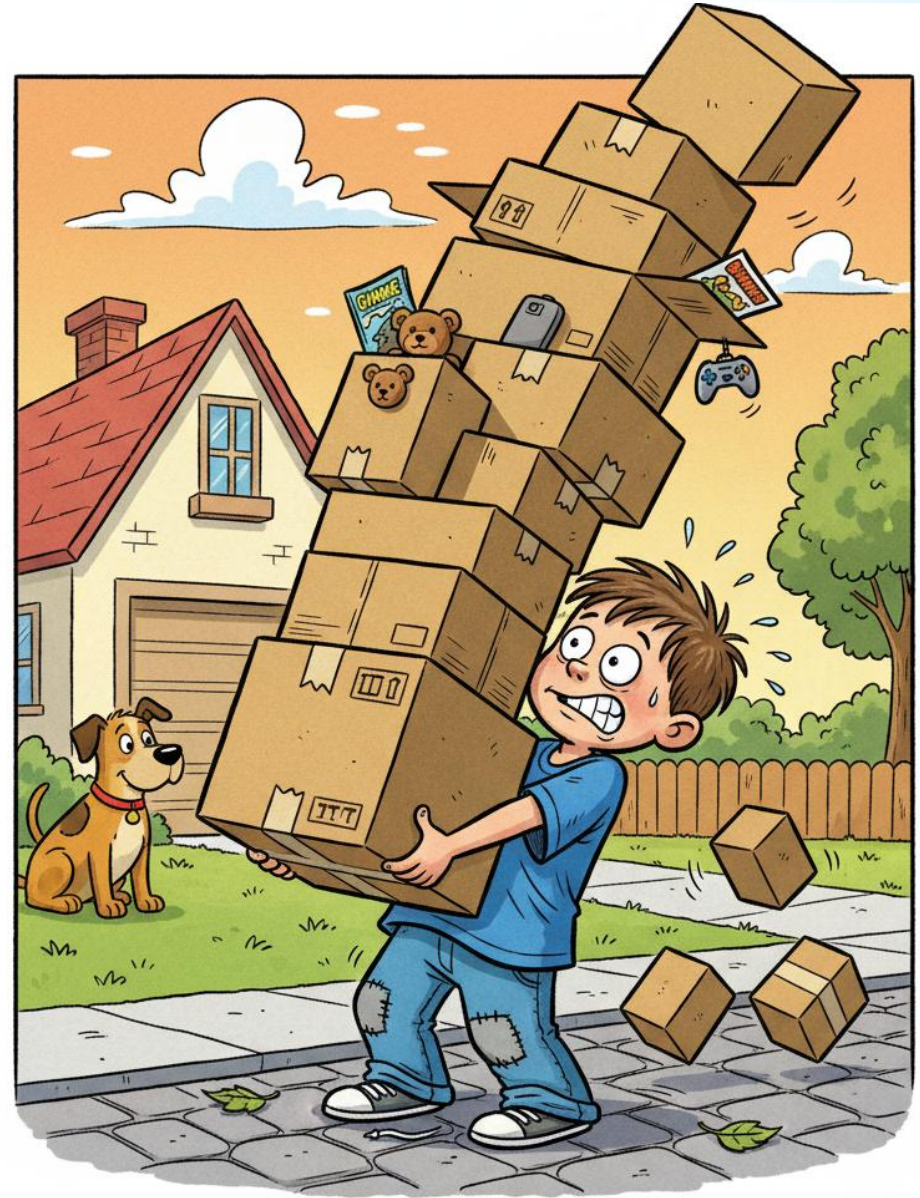


Key Features

- Full switch passing traffic
- Liquid cooled
 - NPU
 - Optics
 - CPU
- Industry standard UQD connectors

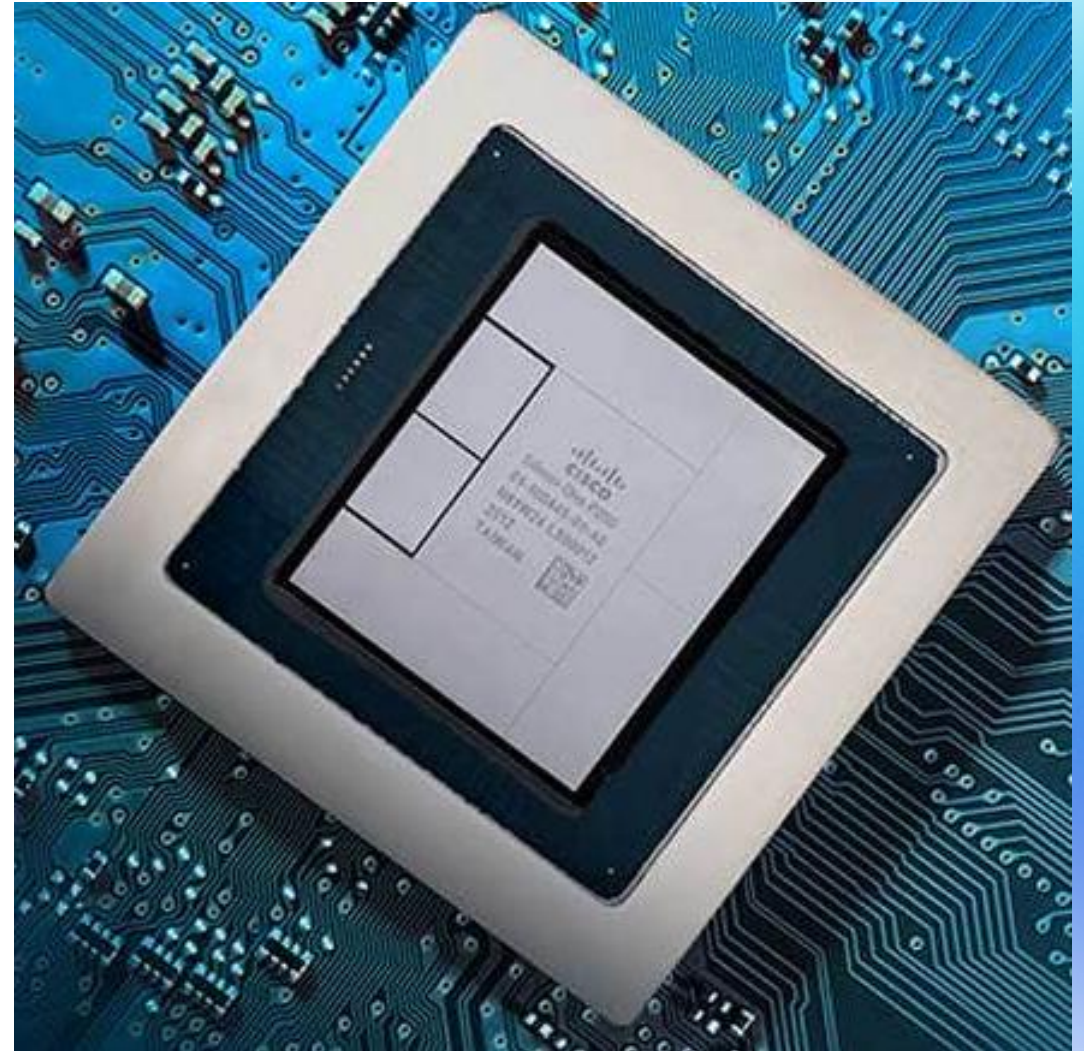
“Advanced Packaging”

- Not this.



“Advanced Packaging”

- Substrates
- Chiplets
- External memories
- 2.5D / 3D packaging
- PVT sensors (Process, Voltage, Temp)
- Packaging and test costs increasing



Wrap Up

Really Big Themes

- Hardware is ... Hard.
- It takes a long time to change things.
- It costs a lot to change things.
- Tradeoffs, tradeoffs... Tradeoffs!
- All designs are a tug-of-war



Really Big Themes

- Tradeoffs, tradeoffs... Tradeoffs!
- Stuff becomes nonlinear.
 - Sometimes mathematically, sometimes non-technically
- “Exchange rates” shift over time
- There are always reasons.
 - Sometimes, there are even good reasons.



Q & A (also webex space)

Complete your session evaluations



Complete a minimum of 4 session surveys and the Overall Event Survey to claim a Cisco Live T-Shirt.



Earn up to 800 points by completing all surveys and climb the Cisco Live Challenge leaderboard.



Level up and earn exclusive prizes!



Complete your surveys in the Cisco Live Events app.

Continue your education



Visit the Cisco Stand for related demos



Book your one-on-one Meet the Expert meeting



Attend the interactive education with Capture the Flag, and Walk-in Labs



Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Thank you

CISCO Live !

