



We power AI

Online Media Briefing

Infiniteon Technologies AG
Power and Sensor Systems Division
June 2024



Athar Zaidi, Senior Vice President & General Manager Business Line, Power ICs and Connectivity Systems



Athar Zaidi

was born in India in 1976.

He holds dual Masters in Mechanical and Computer Engineering from North Carolina State University and an MBA from University of California at Berkeley.

2001 - 2006

Various position in Analog Design at Semtech Corporation.

2006 - 2010

Business Management at Maxim Integrated

2010 - 2012

Director of Business Management at Semtech Corporation

2012 - 2016

Vice President at Dialog Semiconductor

2016 - 2023

Senior Vice President & General Manager Enterprise Power at PSS DCDC

Since 2023

Senior Vice President & General Manager Business Line, Power ICs and Connectivity Systems PSS PCS

Table of contents

1	AI is the next technological revolution	3
2	Implications and challenges of powering AI	6
3	How to make data centers green, robust and reduce TCO	11
4	Summary and Q&A	17

AI is a transformational technology

Every

3.4 months

doubling of the amount of computing power required to train cutting-edge AI models since 2012

5 days

Time it took for ChatGPT to reach 100 million users

\$196bn

Value of the global AI market

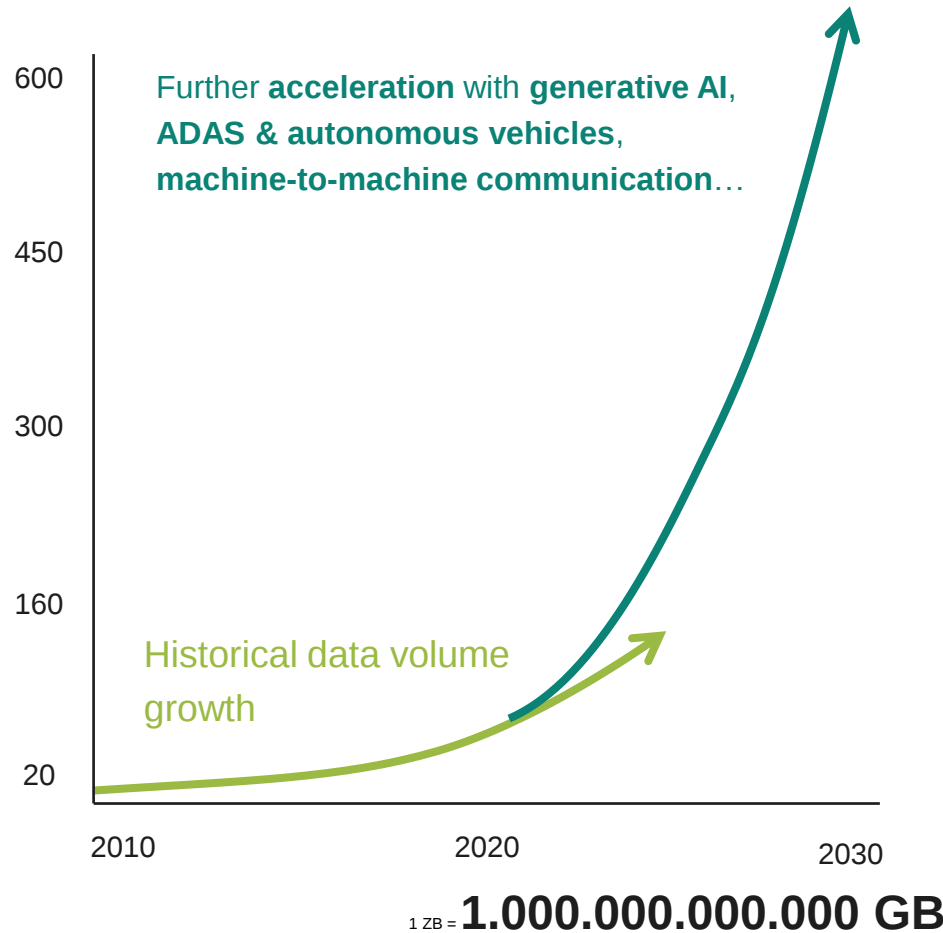
Already now 77% of the global population uses AI
Artificial Intelligence is here to stay

Sources: [BMZ](#), [Similiarweb](#), [openAI](#)

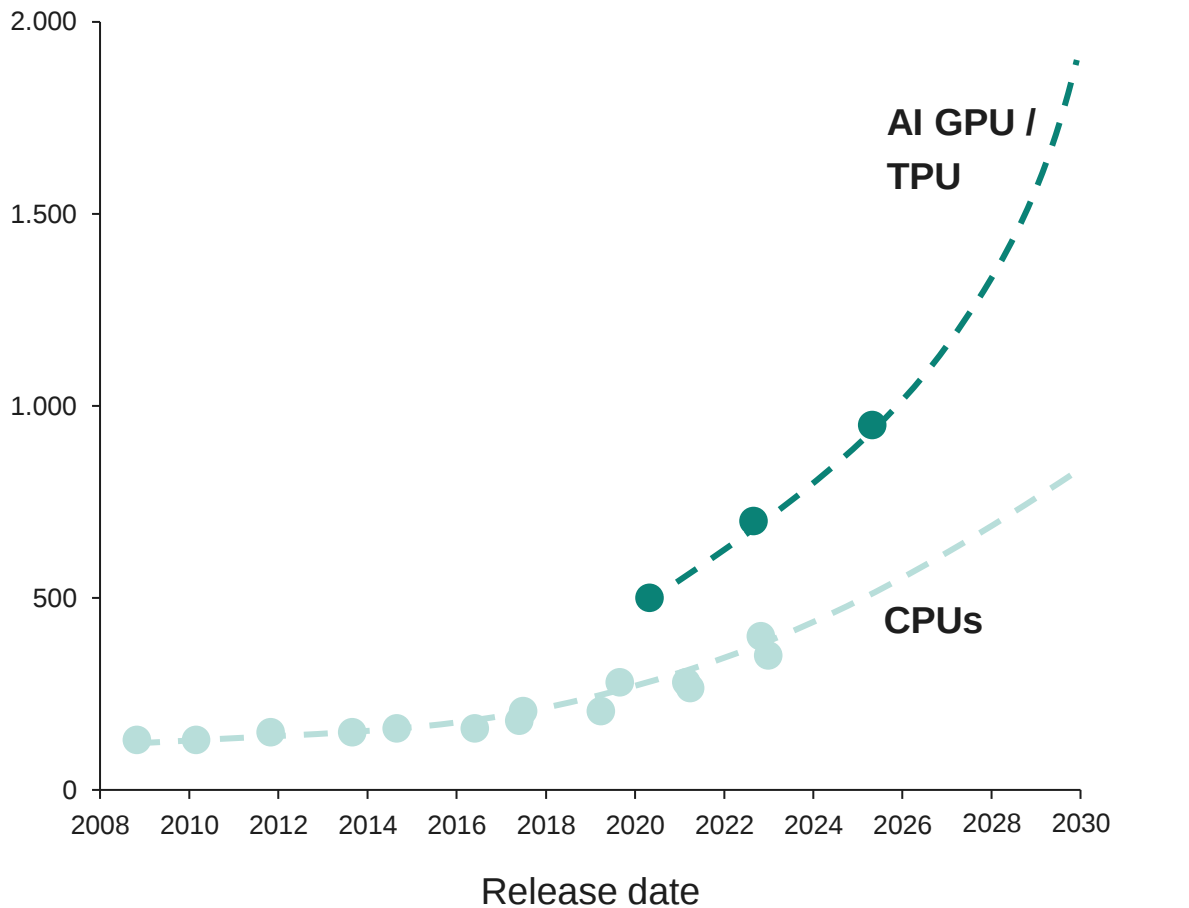
Exponential growth in global data and higher electricity demand of AI enabled systems increase the need for energy efficient solutions



Global data volume in zettabyte (per year)



Single processor electricity demand over time [Watt]



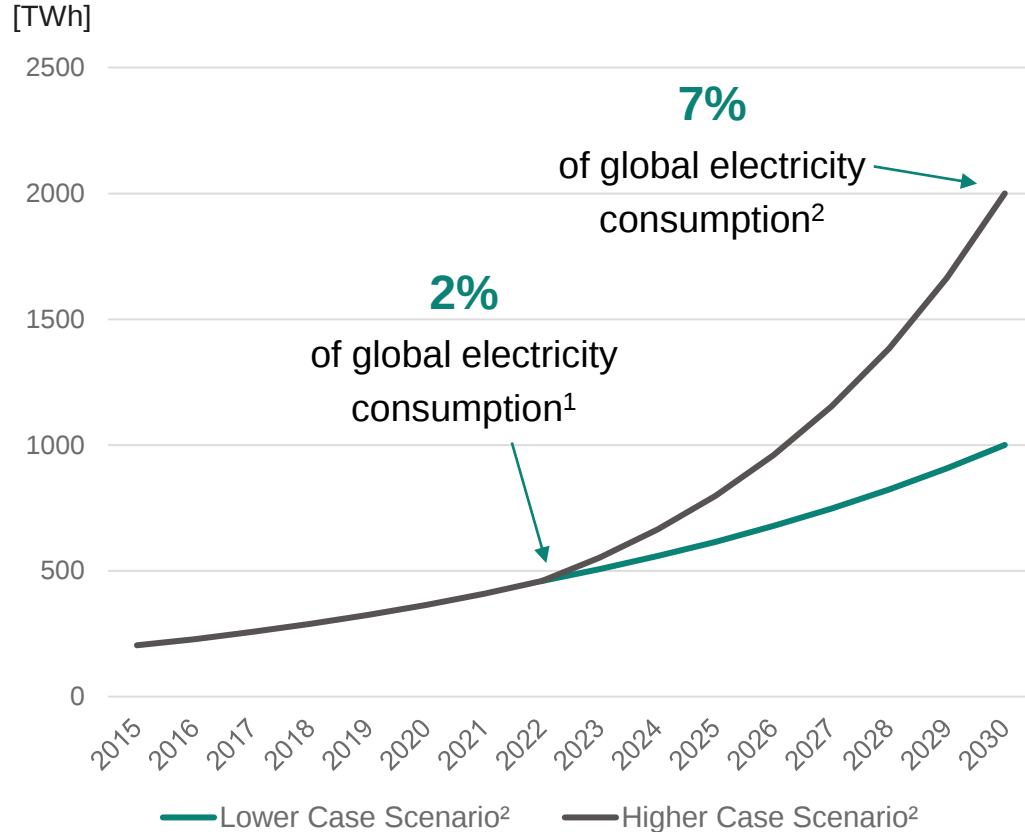
Sources: Omdia; Company information; Infineon

Table of contents

1	AI is the next technological revolution	3
2	Implications and challenges of powering AI	6
3	How to make data centers green, robust and reduce TCO	11
4	Summary and Q&A	17

AI accelerates power demand in data centers, increasing the need for energy efficient solutions

Projected electricity consumption of data centers^{1,2}



Sources

1 [IEA](#); including crypto mining energy use – 2015-2022

2 Infineon assumption and calculation

3 [McKinsey](#)

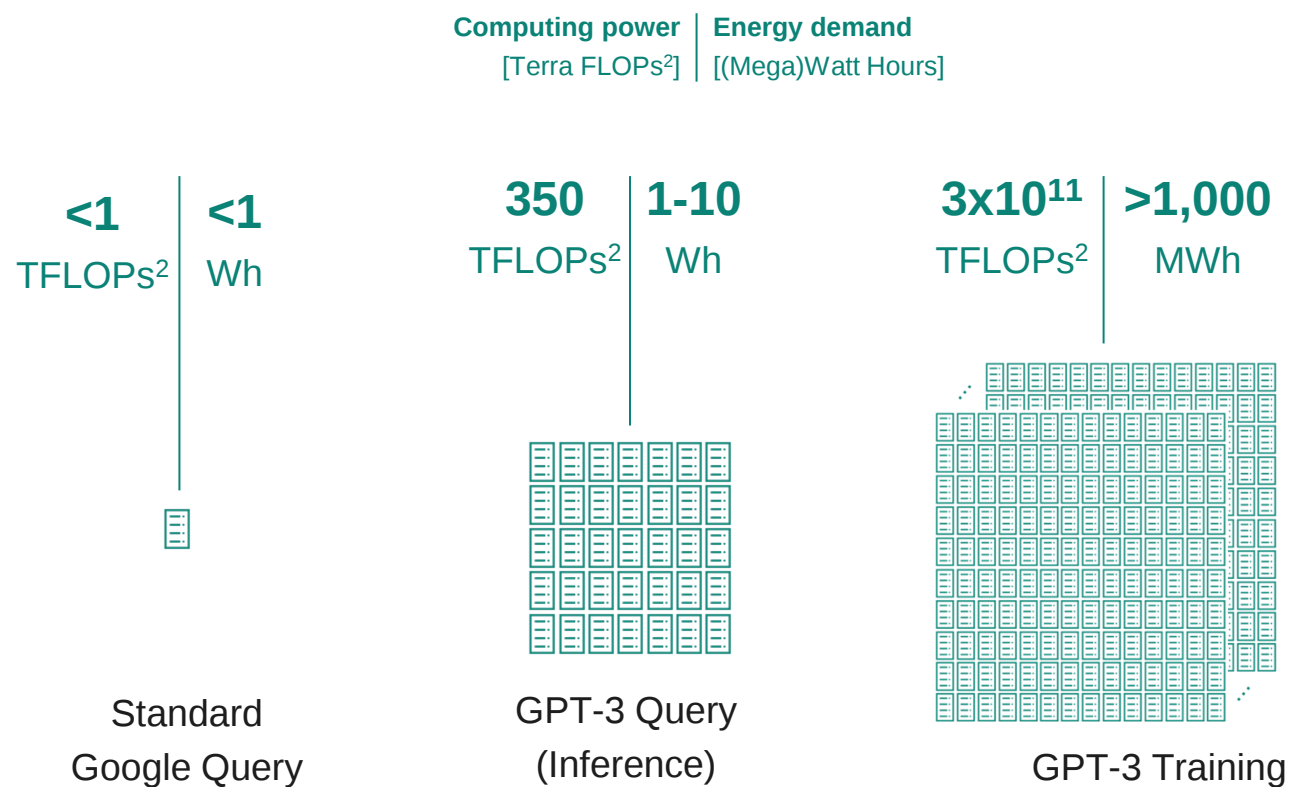
Data centers' share of global final electricity demand was 2% in 2022¹.

Expected to increase to 7% until 2030², which corresponds to the electricity consumption of India.

Example US: power consumption per Data Center is forecasted to grow by 10% a year until 2030³.

Generative AI exponentially increases electricity demand

Computing power and electricity demand in generative AI vs. a Google¹ query



Power supply of an existing data center is limited in the medium term

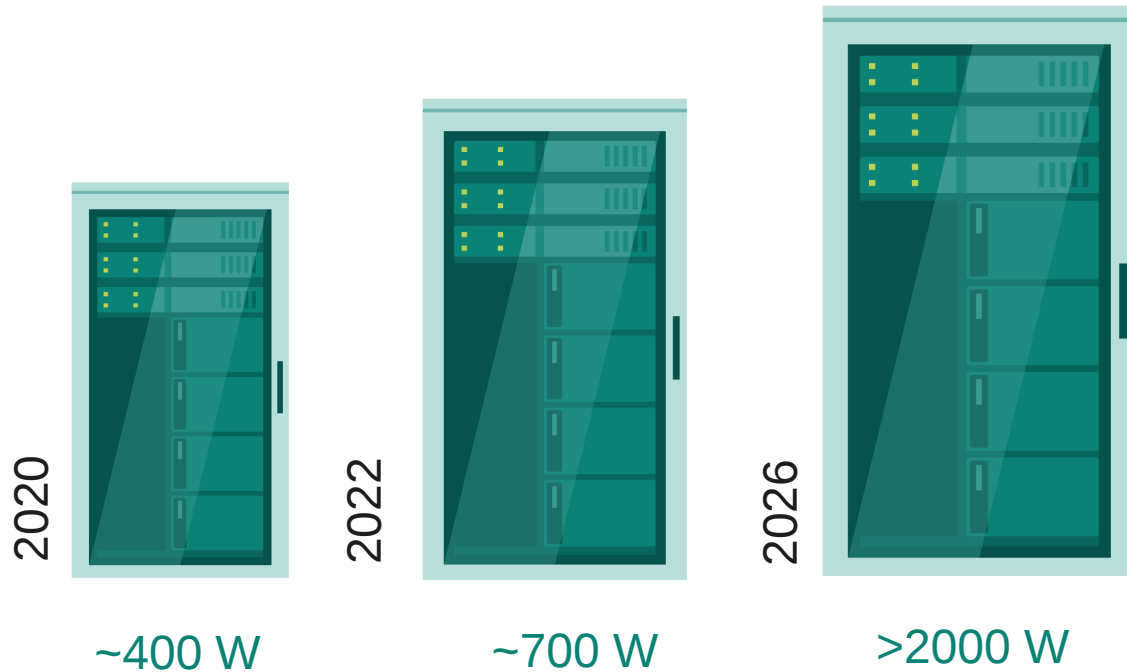
Focus is required on powering AI energy efficiently, w/o compromising on robustness and TCO

Sources: Company information; Statista ¹ Google BERT algorithm ² (Tera=10¹²) Floating Point Operations Per Second

Efficient AI is a multi-dimensional problem

Power management cannot be an after-thought

Exemplary development of power consumption of processors under maximum theoretical load

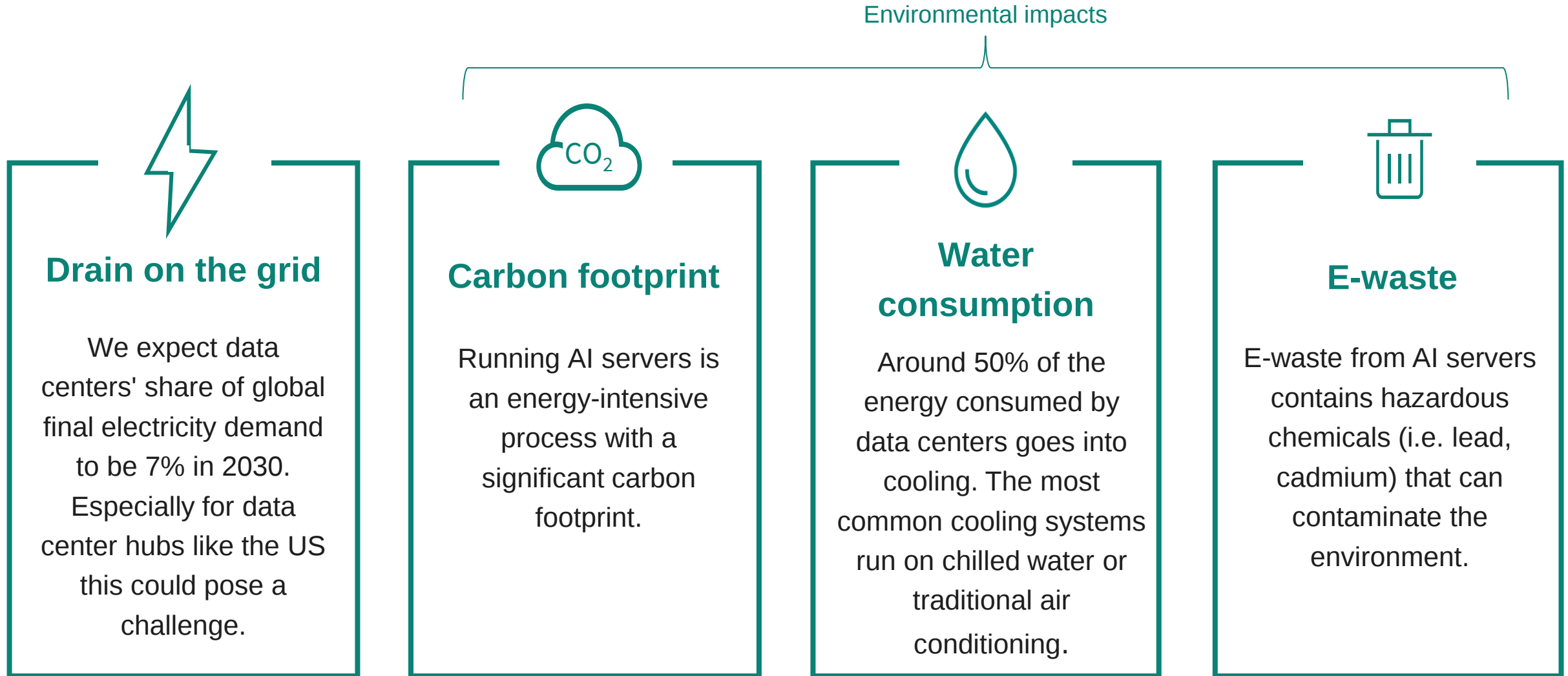


Concerns are emerging regarding the impact of escalating energy requirements linked to **newer chip technologies.**

Increasing compute is only one side of the coin, efficient power consumption being the other side.

We need to **prioritize increasing power efficiency** now to reduce the drain on the grid.

Challenges we can address by focusing on powering AI data centers more efficiently



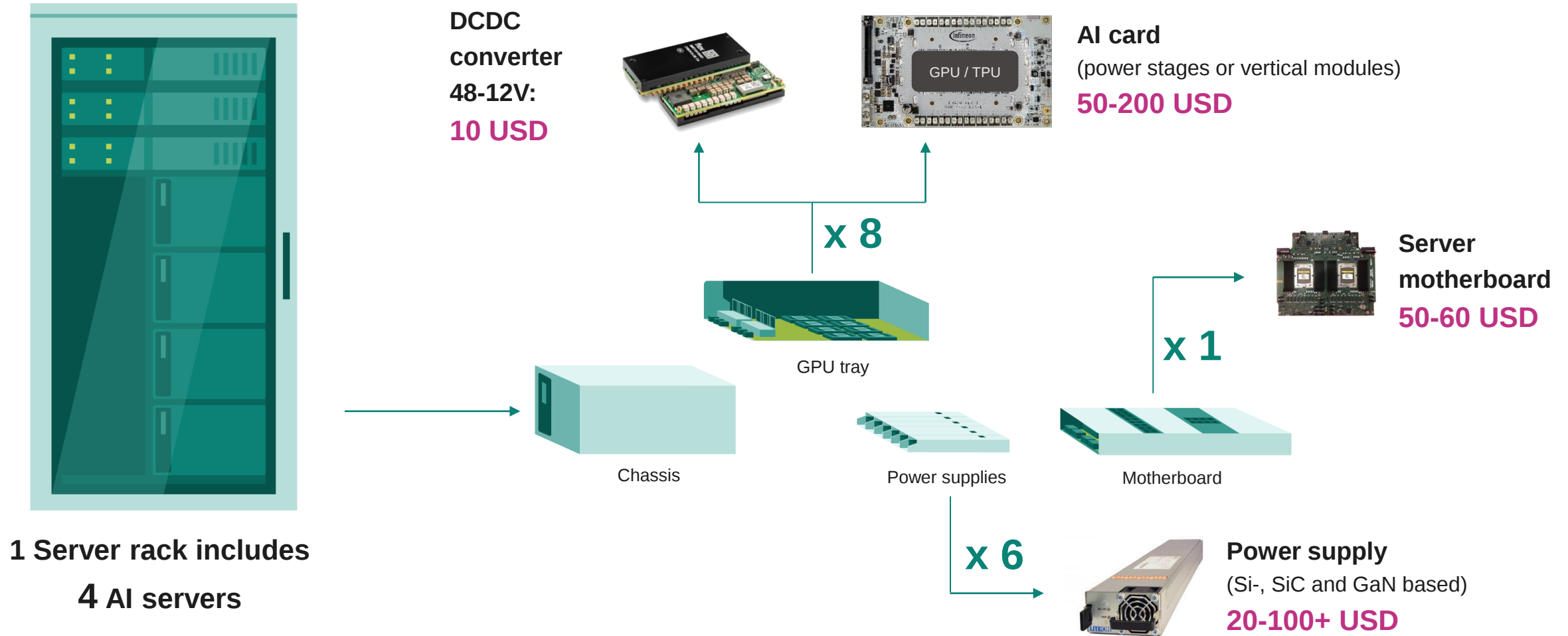
Sources: Infineon, [Earth.org](#), [Study Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models](#)

Table of contents

1	AI is the next technological revolution	3
2	Implications and challenges of powering AI	6
3	How to make data centers green, robust and reduce TCO	11
4	Summary and Q&A	17

We power AI from the grid to the core

Average Infineon BOM per AI server about 850 to 1800 USD



1 Server rack includes
4 AI servers

USD = potential Infineon content per AI server

Infineon improves current existing solutions at all fronts to increase power efficiency and robustness and minimize e-waste



Innovation fronts to improve how we power AI

- **Rearchitecting** power from the grid to the core- **48V** systems, vertical power delivery
- Designing both **Silicon and wide-bandgap** based efficient power supplies
- Make use of **advanced packaging** for **density and cooling**
- Enable **smart control & software**



Improve energy efficiency at least by 8-10%



Increased power density by 30-60%



Best-in-class robustness



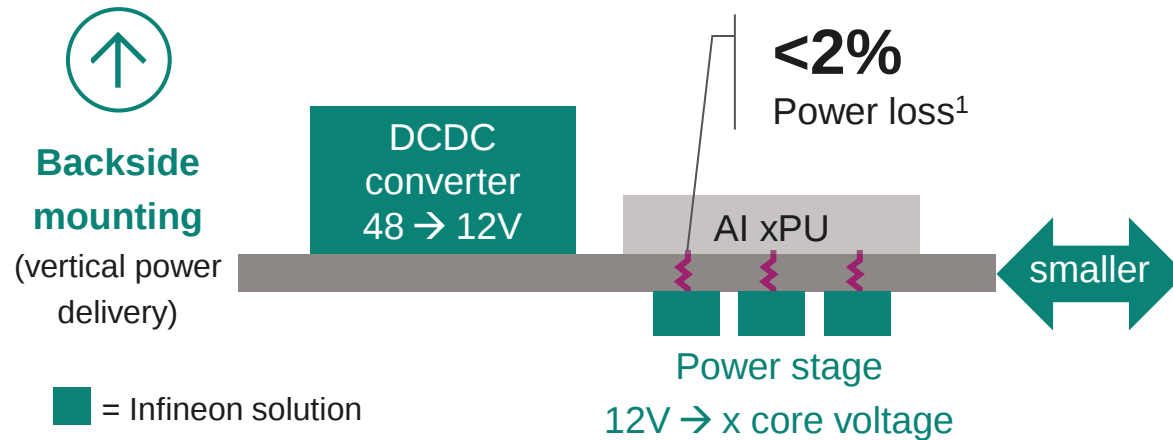
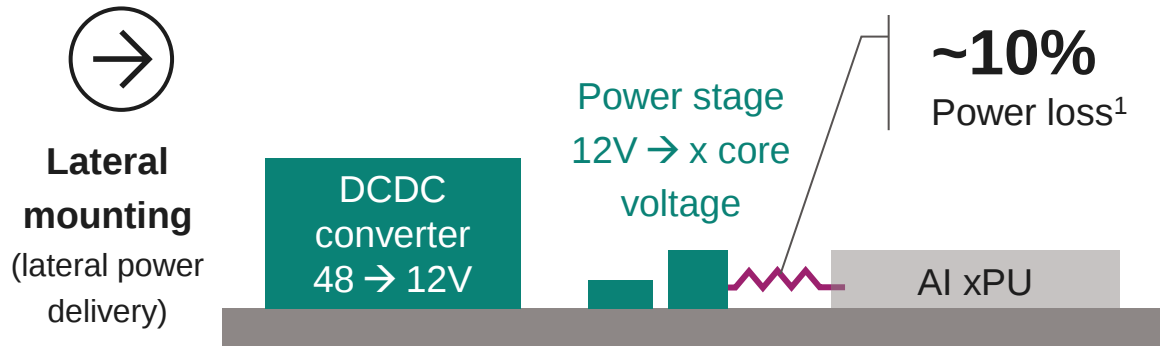
Best-in-class TCO



22 million metric tons CO₂ equivalent could be saved by using Infineon products in all data centers worldwide

Example 1: How does this look on a product level?

Infinion power modules on the AI accelerator card, powering the xPU



Source: Infineon calculation 1 Power delivery loss in % of xPU power

Customer benefits of vertical power delivery enabled by Infineon's vertical power modules



Increase power density via smaller size to enable further increase in compute power



Reduce power losses by >7MW for an average data center (100,000 CPU nodes)



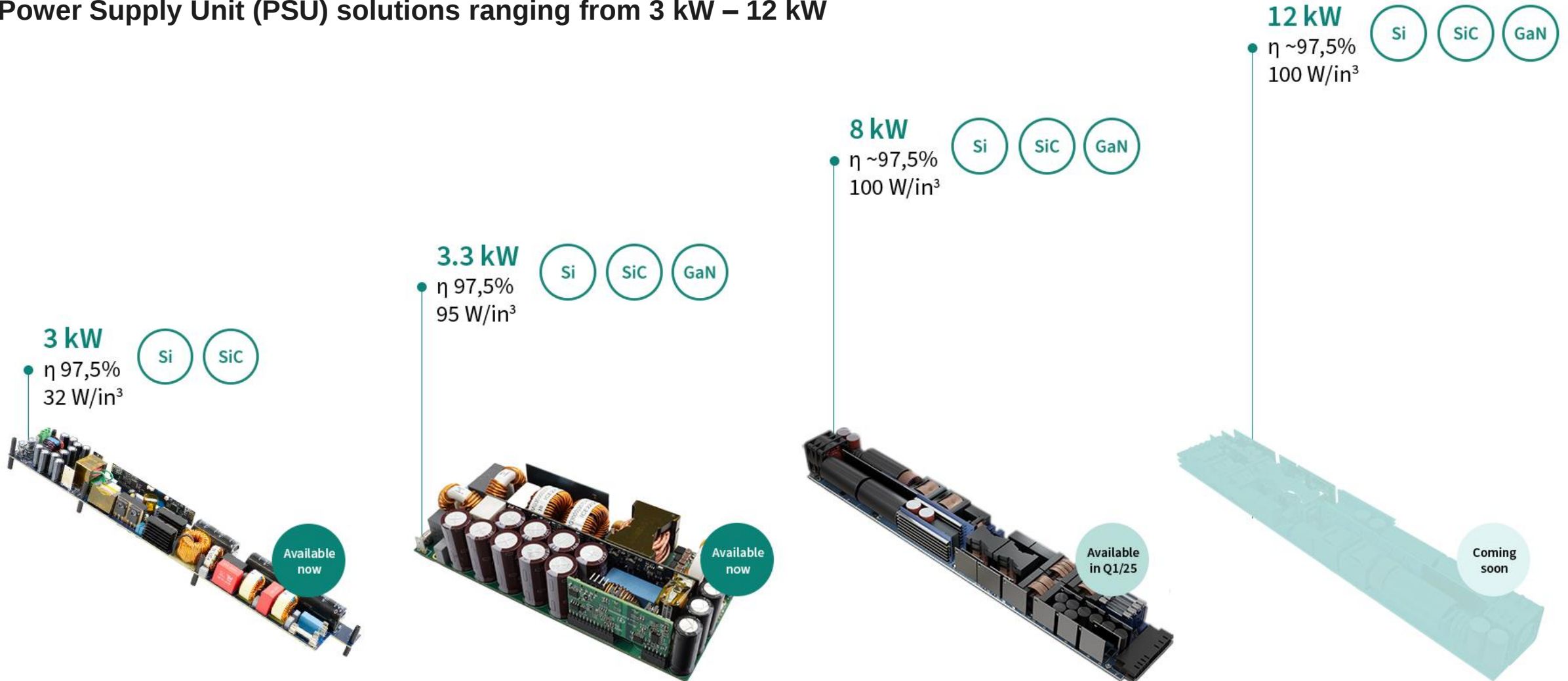
>12% total cost of ownership saving compared to lateral power delivery networks

Example 2: How does this look on a product level?

For AC/DC, Infineon is addressing the growing power demand of AI



Power Supply Unit (PSU) solutions ranging from 3 kW – 12 kW



AI will not only be a strong driver of revenue for Infineon's server business but offers also an important lever to drive decarbonization



We have a clear differentiation potential...

- **System innovation** with **leading companies**
- Industry-leading **system and innovation expertise**
(i.e. digital power control, chip embedding, mastery of Si, SiC and GaN power technologies, our novel vertical power supply architecture)
- Best-in-class in **efficiency** and **lowest cost of ownership**



**Expected FY24
AI server revenue**

Low **triple digit million** amount
(~1bn expected in the next few years)



**CO₂ saving potential by using Infineon
products in AI servers**

22 million metric tons CO₂

Table of contents

1	AI is the next technological revolution	3
2	Implications and challenges of powering AI	6
3	How to make data centers green, robust and reduce TCO	11
4	Summary and Q&A	17

Summary

We power greener AI, shaping the future with our solutions



AI is the next technological revolution and characterized by an **astonishing pace of adoption**¹



Behind the brilliance of AI lies an **energy-intensive** process with a substantial **carbon footprint**²



We are at the forefront of innovation offering energy-efficient and robust solutions **powering AI from the grid to the core!**

¹ Omdia

² [IEA](#)

Hungry for more information?

Listen to our new “We power AI podcast”

Dive into facts, challenges and innovations to power the brilliance of AI



Introduction to powering AI

Adam White, President Power & Sensor Systems



The power struggle: tackling AI data centers' energy requirements

Athar Zaidi, Senior Vice President & General Manager Power ICs and Connectivity Systems



Optimizing AI server power flow with 48 V architectures and vertical power delivery

with Carl Smith, Global Senior Director Datacenter & Computing Application

Upcoming



The power of efficiency: the sweet spot for Si, SiC, and GaN in data centers

with Gerald Deboy, Head of Innovation lab

Upcoming



The importance of robustness, high efficiency & quality in data centers

with Danny Clavette, Distinguished System Architecture Engineer

Upcoming



48 V architecture: topologies, benefits and applications

with Roberto Rizzolatti, Principal Engineer

Upcoming



... or visit our [Infineon promo page](#)

