

# Server computing, a market propelled by AI and data-intensive<sup>1</sup>

## OUTLINES:

- Market forecasts:  
Yole Développement (Yole) expects the server computing market to reach US\$33 billion in 2025.  
AI<sup>2</sup> and HPC<sup>3</sup> dedicated products combined will reach 29Munits and US\$8 billion in 2025. Configurable SoC<sup>4</sup> dedicated to AI inference will be more used following AI's hype.  
The CPU revenue will reach US\$21 billion in 2025, driven by Standard Cloud servers and HPC.  
Yole expects a 11% CAGR<sup>5</sup><sub>2019-2025</sub> for the HPC segment.
- Market and technology trends:  
Social and media applications are growing strongly, more workload needs to be computed. Enterprise digitalization enhances the demand for hardware for virtualization and cloud computing for enterprises.  
5G arrival will likely lead to a bunch of all new applications.  
Limiting power consumption is critical to prevent an explosion in electricity consumption.
- Supply chain:  
Some big tech players are designing their own chips.  
Intel, Nvidia, AMD, and Xilinx, the few logic leaders, still by far owned the largest part.

*“Between new opportunities and requirements brought by 5G, high-end AI and HPC computing workloads, and growing pressure to limit and in fact reduce power consumption, the datacenter server computing market is facing many technical challenges.” asserts **Adrien Sanchez, Technology & Market Analyst, Computing & Software division at Yole Développement (Yole).***

---

<sup>1</sup> Extracted from: [Computing for Datacenter Servers 2021](#), Yole Développement

<sup>2</sup>AI: Artificial Intelligence

<sup>3</sup> HPC: High-Performance-Computing

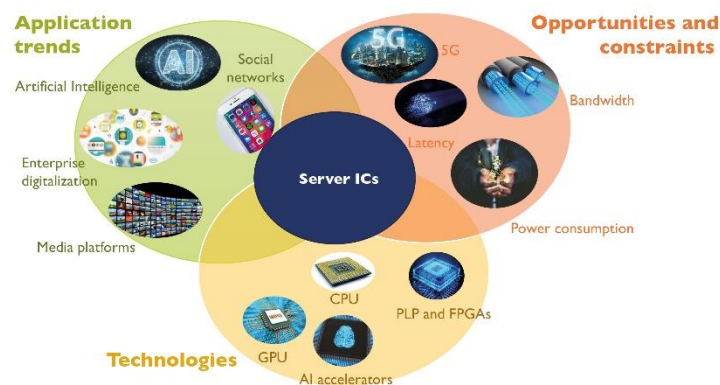
<sup>4</sup> SoC: System on a Chip

<sup>5</sup> CAGR : Compound Annual Growth Rate

What are the benefits and pitfalls of the various types of IC, with a focus on CPU<sup>6</sup>, GPU<sup>7</sup>, and PLP<sup>8</sup> such as FPGA<sup>9</sup>s and AI accelerators? What are the dynamics and outcomes of technology trends such as workload specialization and unit integration?

## 2021 Computing market for datacenter servers: three perspectives of server chips

(Source: *Computing for datacenter servers 2021 report*, Yole Développement, 2021)



In this context, Yole investigates disruptive technologies and related markets in depth, to point out the latest innovations and underline the business opportunities. Released today, the [Computing for Datacenter Servers 2021 report](#) contributes to the understanding of the impact on computing chips which can be found in datacenter servers, of key markets, such as cloud computing, enterprise digitalization, social and media, AI training and AI inference. This report provides a scenario for computing in this dynamic market and presents an understanding of the impact of both high-level workloads, including AI and HPC, and massive volume tasks such as video streaming. Including market trends and forecasts, supply chain, technology trends, technical insights and analysis, take away and outlook, this study also delivers an in-depth understanding of the ecosystem and main players' strategies. What are the economic and technological challenges of the computing for datacenter industry? What are the key drivers? Who are the suppliers to watch, and what innovative technologies are they working on? What are the impacts of technology transition in datacenter business such as 5G or rising constraints such as power consumption?

<sup>6</sup> CPU: Central Processing Unit

<sup>7</sup> GPU: Graphics Processing Unit

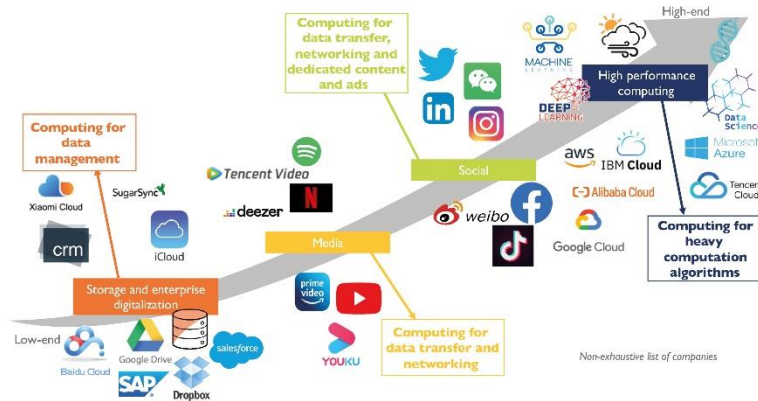
<sup>8</sup> PLP: Programmable Logic Processors

<sup>9</sup> FPGA: Field-Programmable Gate Array

Yole presents today its vision of the computing for datacenter servers industry.

## From Storage to High Performance Computing: application trends segmentation

(Source: Computing for datacenter servers 2021 report, Yole Développement, 2021)



As analyzed by Yole’s team in the new [Computing for Datacenter Servers 2021](#) report, from one side the share of specialized chips such as GPU, PLP and accelerators is increasing, and from another side more and more units are integrated into a single chip to address more workloads. Artificial Intelligence algorithms are driving both of these trends, with dedicated accelerators but also and dedicated units integrated into each type of chip. Dedicating the computing to a specific type of algorithm through faster and more energy-efficient chips is a central trend, which is leading either to the use of dedicated chips such as FPGAs and accelerators or to the integration of specialized units into other chips.

On the one hand, the IC<sup>10</sup> datacenter server market has been boosted by large-volume demand with the digitalization of companies, an existing trend which was boosted even more by COVID-19, as well as by consumer applications, such as social networks and media applications.

According to **Adrien Sanchez from Yole**: “These two data-intensive applications rely heavily on video streaming which requires intensive computing capabilities, while they also put pressure on bandwidth. This is leading progressively to a new network organization which is less centralized and closer to the edge”.

On the other hand, high-end applications using HPC and the spread of AI in every vertical are requiring more and more capabilities. And today most of the AI training and a still important portion of AI inferences are done in large datacenters. Yole’s analysts expect the associated revenues will grow with a CAGR of more than 10% in the coming years, reaching a total of US\$33 billion in 2025.

<sup>10</sup> IC: Integrated Circuit

All year long, Yole Développement publishes numerous reports and monitors. In addition, experts realize various key presentations and organize key conferences.

Throughout the year, discover the numerous battery-related reports. Make sure to be aware of the latest news coming from the industry and get an overview of our activities, including interviews with leading companies and more on i-Micronews. Stay tuned!

### Press contacts

**Sandrine Leroy**, Director, Public Relations, [leroy@yole.fr](mailto:leroy@yole.fr)

**Marion Barrier**, Assistant, Public Relations, [marion.barrier@yole.fr](mailto:marion.barrier@yole.fr)

Le Quartz, 75 Cours Emile Zola – 69100 Villeurbanne – Lyon – France – +33472830189

[www.yole.fr](http://www.yole.fr)- [www.i-micronews.com](http://www.i-micronews.com)- [LinkedIn](#) – [Twitter](#)

### About our analyst

As a Technology & Market Analyst, **Adrien Sanchez** belongs to the Computing & Software business unit at Yole Développement (Yole), part of the Yole Group of Companies. In collaboration with his team, Adrien produces technology & market analyses covering computing hardware and software, AI, machine learning and neural networks. Prior to Yole, he worked as an intern at AW Europe (Belgium), where he focused on image recognition & comprehension for ADAS. He also worked at ACOEM (France), where he focused on realtime sound classification using deep learning and edge computing. Adrien graduated with a double degree at Grenoble Institute of Technology PHELMA (Grenoble INP Phelma, France) and Grenoble Ecole de Management (GEM, France), and he earned an MSc on AI at Heriot-Watt University (Edinburgh, UK).

### About the report

#### **Computing for Datacenter Servers 2021**

*Popular data-intensive applications and the spread of AI applications are boosting server computing market to reach US\$33B in 2025.* – Performed by Yole Développement

#### **Companies cited:**

Achronix, Advanced Micro Devices (AMD), Alibaba Group, Ampere Computing, Apple, Akamai, ARM, Atos, Amazon Web Services (AWS), Baidu, Bull, Cerebras, Cisco, Compal, Cray, Deezer, Digital Reality, Dell, Electronic Arts, Equinix, Esperanto Technologies, Facebook, Fujitsu, Global Foxcom, GlobalFoundries (GF), Google, Graphcore, Groq, HiSilicon, Hitachi, Hewlett Packard Enterprise (HPE), Huawei, IBM, Intel, Inventec, Inspur, Kalray, Lattice Semiconductor, Lenovo, Lightmatter, Marvell, Mellanox, Microchip, Microsoft Azure, Mipsology, NEC, Netflix, NTT Communications, Nvidia, Oracle, Penguin Computing, Pezy Computing, Quanta Cloud Technology (QCT), Quality Technology Services (QTS), Qualcomm, Quanta Computer, Rackspace, Samsung, Sabey Data Centers, SAP, Semiconductor Manufacturing International Corporation (SMIC), Socionext, Sony, Spotify, STMicroelectronics (STM), Sugon, SuperMicro, Tencent, Taiwan Semiconductor Manufacturing Company (TSMC), Twitter, Xilinx, Weibo, and more...

#### **Related reports:**

- [Cameras and Computing for Surveillance and Security 2020](#)
- [Artificial Intelligence Computing for Automotive 2020](#)
- [\(x\)PU: High-End CPU and GPU for Datacenter Applications 2020](#)
- [Intel Foveros 3D Packaging Technology](#)

#### **About Yole Développement**

Founded in 1998, Yole Développement (Yole) has grown to become a group of companies providing marketing, technology and strategy consulting, media and corporate finance services, reverse engineering and reverse costing services and well as IP and patent analysis. With a strong focus on emerging applications using silicon and/or micro manufacturing, the Yole group of companies has expanded to include more than 80 collaborators worldwide... [More](#)

**For more information and images, please visit our website [i-Micronews](#)**

###