Summary
Multiprocessors are basic building blocks for all computer systems. This course covers the architecture and organization of modern multiprocessors, prevalent accelerators (e.g., GPU, TPU), and datacenters. It includes a research project on multiprocessors and post-Moore era datacenters.

Content
- Introduction
- Metrics and methodologies
- Parallel programming models
- Communication models
- Applications and Workloads
- Cache hierarchies & memory models
- Memory & storage hierarchies
- Interconnects
- GPUs
- AI/ML/Analytic accelerators
- Near-memory computing
- Datacenters & cloud Computing
- Cloud-native CPU
- Cloud-native memory hierarchies
- Sustainable architecture

Learning Prerequisites
Recommended courses
- Advanced computer architecture
- Systems for data management and data science

Learning Outcomes
• Explore the development trend of computation systems and datacenters
• Establish the basic model to analyze the performance and characteristics of foundational workloads operating in cloud environments.
• Classify and describe the components of modern parallel systems, including multiple processors, cache hierarchies, memory systems, interconnects, and accelerators, and their roles in handling emerging workloads
• Define and clarify research questions and opportunities
• Interpret and critique research papers and extract insights for research questions
• Plan and conduct a research project
• Present research contributions

Teaching methods
Lecture, research paper retrieval, and a research project

Assessment methods
• Homework: 10%
• Research project (in group): 40%
• Midterm exam: 20%
• Final exam: 30%

Resources
Websites
• https://parsa.epfl.ch/course-info/cs471/