

DATACENTER-SCALE COMPUTING



..... Many of today's computing needs are fulfilled not from software installed in locally accessible desktop machines but from programs and services running in large computing clusters, accessible through high-speed Internet links. Personal data, such as photos and video, is also starting to move from home devices to Internet-accessible storage services due to the added convenience of access, and ease of manageability and sharing. Such trends bring a renewed spotlight to the architecture, design trade-offs, operation, and programming of the computing infrastructure behind these computing and storage services, referred to here as *datacenter-scale computing*.

Although computing systems based on networked servers or workstations has been an active area of research and development for decades, the architecture of the systems being built by the world's largest Internet services companies have evolved beyond the early clustered systems. When taken together, the scale of the hardware infrastructure, the size and complexity of the multitiered workloads, the approach to fault tolerance, the management and provisioning models, and the programming systems used in this domain, make it clear that these systems are a distinct computing platform.

This field is still in its relative infancy but it has received enough attention that a sizable body of work from both academia and industry is already available, and some consistent technological trends have begun to emerge. We therefore felt that it would be timely to compile a set of original articles with the objective of presenting a small sample of the work underway by researchers and professionals in this new field for the

IEEE Micro audience. We also decided not to restrict the topic areas to microarchitecture or hardware-level subjects. This being a new area, it was important to broaden the scope of submissions to include pertinent software and operational issues in order to provide *Micro's* readership with the best perspective. This decision also reflects the key role that hardware-software codesign plays in the development of effective datacenter-scale computer systems.

The articles

Consistent with these guidelines, the articles in this issue—which include peer-reviewed and invited opinion pieces—feature a rich cross-section of ongoing work and emerging trends in the areas of systems architecture design, programming abstractions, and manageability.

The first article, “Server Engineering Insights for Large-Scale Online Services,” by Christos Kozyrakis and his colleagues, discusses how traditional assumptions around system balance must be revisited for datacenter-scale computing, in the context of three large-scale production-class online services from Microsoft—Hotmail, Bing, and Cosmos. This is followed by “Challenges and Opportunities for Extremely Energy-Efficient Processors,” an interesting point-counterpoint discussion on an ongoing debate in the systems community around the role of “wimpy” low-power processors for datacenter computing. In “Extremely Energy-Efficient Processors,” Trevor Mudge argues that future near-threshold voltage processors offer significant energy efficiency improvements and, in conjunction with techniques such as boosting and other system

Luiz André Barroso
Google
Parthasarathy
Ranganathan
HP Labs

optimizations, represent an important opportunity for future data centers. Urs Hölzle, on the other hand, provides a cautionary note in “Brawny Cores Still Beat Wimpy Cores, Most of the Time,” arguing for consideration of software development overheads, response time constraints, increased infrastructure costs, and overall efficiency, in identifying the sweet spot for future energy-efficient processors. In their article, “The Case for Full Throttle Computing: An Alternative Datacenter Design Strategy,” Jose Moreira and John Karidis address a similar problem, but discuss how in a data center with diverse workloads, maximizing server utilization by consolidating latency-sensitive workloads with batch-like workloads can often provide the best way to minimize the total cost of computing.

In addition to the appropriate choice of compute elements, datacenter-scale computing requires an equally important emphasis on networking and storage. In “Scale-Out Networking in the Data Center,” Amin Vahdat and his colleagues provide an excellent overview of the key challenges in providing low-latency, high-bandwidth communication at the scale of tens of thousands of servers, focusing on network topology, network element design, network protocol considerations, and data forwarding. (Also in this issue, Sven-Arne Reinimo and his colleagues give an update on the IEEE datacenter bridging standards pertaining to future trends in Ethernet for high-performance data centers in “Ethernet for High-Performance Data Centers: On the New IEEE Datacenter Bridging Standards.”) David Anderson and Steven Swanson, in “Rethinking Flash in the Data Center,” discuss the role of emerging nonvolatile memories in the data center, focusing specifically on opportunities and challenges around flash.

The last two articles focus on the software and operational aspects of datacenter-scale computing. In “Transformer: A New Paradigm for Building Data-Parallel Programming Models,” Peng Wang and his colleagues describe a new programming framework to enable diverse data-parallel programming models, and present some early results based on a production cluster deployed at Tencent. The final article in our collection, “Google-Wide Profiling: A Continuous Profiling Infrastructure for Data Centers,”

by Gang Ren and his colleagues, spotlights manageability and performance monitoring, specifically focusing on the challenges, architecture, and use of the Google-wide profiling (GWP) solution.


We are fortunate to be able to present such a strong sample of ongoing work in the area of datacenter computing. Our thanks to all the authors of submitted papers, all the reviewers who provided valuable feedback to help improve the papers’ content and helped us choose the appropriate subset, and *IEEE Micro’s* editorial staff. This already being such a broad field of study, many important topics were inevitably left uncovered in this collection. They include power and cooling considerations for large-scale data centers, resource management, and datacenter scheduling. We believe that the opportunities in this space are immense, and that the next decade will likely see a transformation in the datacenter landscape, with new workloads and new technologies leading to a new systems stack and architectural optimizations targeted specifically at this space. We hope that the articles in this special issue provide broader visibility to, and initiate greater discussion of such topics.

MICRO

Luiz André Barroso is distinguished engineer at Google. His research interests include hardware and software design for datacenter-scale systems and energy-efficient computing. Barroso has a PhD in computer engineering from the University of Southern California.

Parthasarathy Ranganathan is a distinguished technologist at HP Labs. His research interests include system architecture and energy-efficient computing. Ranganathan has a PhD in electrical and computer engineering from Rice University. He is a senior member of IEEE and ACM.

Direct questions and comments about this article to Luiz Barroso, luiz@google.com, or Parthasarathy Ranganathan, Partharanganathan@hp.com.

 Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.