# Power–Performance Tradeoff in a Federated Cloud

**Albert Zomaya,** *University of Sydney*

**This installment of *Computer*'s series highlighting the work published in IEEE Computer Society journals comes from *IEEE Transactions on Computers*.**

Cloud computing has received considerable attention and acceptance as a promising way to manage and improve the utilization of hardware, software, databases, and information resources. Improving energy efficiency and response time are key issues for large-scale server systems in current and future datacenters. In a recent issue of *IEEE Transactions on Computers*, Junwei Cao, Keqin Li, and Ivan Stojmenovic ("Optimal Power Allocation and Load Distribution for Multiple Heterogeneous Multicore Server Processors across Clouds and Data Centers," vol. 63, no. 1, 2014, pp. 45–58) discuss balancing these two conflicting goals and propose centralized management of cloud resources and user demands. A federated cloud of clouds could result in improved computing resource use across multiple datacenters, server consolidation, heterogeneous resource and computing power integration, and more powerful and flexible services for various scientific, business, and industrial applications.

Each cloud is represented and controlled by a multicore server, and core numbers and server speeds differ across clouds. Ideal for multitasking, multimedia, and networking applications, multicore technology delivers the ultimate computing solution: by incorporating multiple processor cores into a single package that delivers parallel execution of multiple software applications, multicore technology enables higher levels of performance *and* consumes less power than that typically required by a higher-frequency single-core processor with equivalent performance.

Software techniques for power reduction are supported by a mechanism called dynamic voltage (or frequency, speed, and power) scaling. A power-aware system-management algorithm can change supply voltage and frequency at appropriate times to optimize a combined consideration of performance and energy consumption. Dynamic power management at the operating system level provides supply voltage and clock frequency adjustment schemes implemented while tasks run. These energy conservation techniques explore the opportunities for tuning the energy–delay tradeoff.

The aggregated performance of a cloud of heterogeneous clouds, each controlled by a multicore server processor, can be optimized

 Published by the IEEE Computer Society 

through load distribution and balancing. Cao and colleagues aim to develop power- and performance-constrained load distribution methods for cloud computing in current and future large-scale datacenters. Their strategy is to formulate optimal power allocation and load distribution for multiple servers in a cloud of clouds as multivariable optimization problems—one power constrained and the other response time constrained. The authors also model a multicore server processor as a queuing system for multiple servers, and they describe new theoretical and practical insights into power management and performance optimization in cloud computing. Furthermore, they develop efficient methods to solve the optimization problems for two different power consumption and core speed models—the idle- and the constant-speed models—with consideration for both static and dynamic power consumption.

Load-balancing capabilities improve service quality (notably task response time) and are a source of revenue for cloud providers because they reduce power consumption and attract more end users. Hence, an efficient load-balancing strategy is a key component to building out any cloud computing architecture. The authors demonstrate the feasibility of studying the power–performance tradeoff for a cloud of clouds with heterogeneous servers in an analytic way. Their models, results, and algorithms are also applicable to other server systems and computing environments. **C**

***Albert Zomaya*** *is a chair professor of high-performance computing and networking at the University of Sydney. Contact him at albert.zomaya@sydney.edu.au.*