Several ongoing developments necessitate a study of architectural and operational concerns in modern data centers, including data center fabric research, scale-up versus scale-out, resource consolidation, infrastructure management, and operational cost reduction.

Research focusing on data center networking has traditionally concentrated on an isolated set of issues such as load balancing and distribution, Web servers, databases, and application services. The reasons for not pursuing a more holistic view of data centers include the lack of visibility into commercial data center architectures and operation, use of proprietary fabrics and protocols, and the need to minimize any potential disruptions to the operations brought about by unreliable technologies. However, several ongoing developments necessitate a comprehensive study of the architectural and operational concerns in modern data centers. The main issues can be roughly classified into five categories.

**DATA CENTER FABRIC**

Currently, data centers use Ethernet for networking traffic and Fibre Channel for storage—perhaps in combination with directly attached storage. If clustered database or application servers need low-latency interprocess communication (IPC), they use specialized fabrics such as Myrinet, the InfiniBand architecture, or QsNet to carry such traffic.

With the availability of 10-Gigabit Ethernet and hardware implementations of basic networking and storage protocols, it is possible to use Ethernet as a unified fabric for all data center traffic types including networking, storage, IPC, and management. The main attractions of this approach are the reduced cost, widespread expertise in Ethernet-related protocols, and simplicity of using a single fabric. However, the transition to a unified Ethernet fabric requires resolving several major challenges, including efficient implementation of the dominant protocol layers—for example, iSCSI, TCP, RDMA, and SSL—at high data rates; advanced quality of service to ensure that a unified fabric works almost as well as multiple isolated fabrics; sophisticated high-availability mechanisms; and security mechanisms to address the exposure of storage/IPC traffic on the Ethernet.

Satisfying these and other demands of future data center applications could require examining new protocols, adding enhancements to existing protocols, and introducing new platform features. Developments in wireless and optical networking technologies provide another dimension for data center fabric research.

**SCALE-UP VERSUS SCALE-OUT**

The evolving economics of hardware technology has forced a rethinking of how to implement data center applications. In particular, with the availability of inexpensive but high-performance clustering solutions, using clustered application deployment instead of expensive multithreaded symmetric multiprocessors (SMPs) becomes more attractive.

A clustered solution offers many advantages over SMPs, including incremental upgradability, more flexibility, and the potential for higher fault toler-
Ideally, virtual clusters could support the illusion that each application is running on its own isolated cluster.

RESOURCE CONSOLIDATION
Data center servers typically run at a very low average utilization. Partly, this is necessary to provide sufficient reserve capacity to handle traffic bursts, but it is also a result of deliberate application segregation and isolation, which provides better availability and easier diagnosis, maintenance, and upgrade capabilities. If virtualization can provide the same attributes, it can significantly reduce the number of servers required, thereby reducing the costs associated with space, power usage, network infrastructure, maintenance, and hardware.

This potential for significant savings has resulted in interest in developing efficient virtualization technologies. The virtualization concept can be easily extended from the traditional virtual machines to virtualized I/O devices including storage and network interfaces and, ultimately, to multiple “virtual clusters” created dynamically using real or virtual components such as machines, devices, links, switches, and routers. Ideally, these virtual clusters could grow and shrink depending on the application needs and yet support the illusion of each application running on its own isolated cluster. There are, of course, substantial research challenges in making this vision work.

A somewhat related data center concept is utility computing. Instead of establishing its own moderate-sized data center, an organization could instead buy or rent resources from a large data center provider as a “utility.” This concept will be most useful if the customer can change the size, capacity, and capabilities of its resources on fairly short notice, perhaps even dynamically.

The virtual cluster concept, although not necessary for utility computing, makes the dynamic provisioning natural. Since the service provider’s real data center could be distributed, it is conceivable that a customer’s “virtual data center” could itself be distributed. However, issues such as service provisioning, service-level agreements, and security could be very challenging in such an environment.

INFRASTRUCTURE MANAGEMENT
As data centers increase in size and computational capacity, numerous infrastructure issues are becoming critical. For example, inventory management of hardware—including its physical location—becomes complex enough to require automated solutions. Similarly, in a data center full of heterogeneous hardware, software, and operating systems, operations such as backups, software upgrades and patches, and virus updates could require new technologies, algorithms, and associated management tasks.

Power management, another fast-growing problem for data centers because of constantly increasing power consumption of CPUs, memory, and storage devices—includes two subissues:

- **power-in management**—intelligent power control schemes to maximize performance per watt both on individual devices and spanning the entire data center; and
- **power-out management**—intelligent controls to ensure that power dissipation, or the operating temperature, remains within limits while maximizing performance.

Power quality is also an increasingly important issue. A significant number of fault episodes in a data center result from unclean power, yet the transient power loads resulting from traffic variations and reconfigurations affect the quality of the power itself.

As data center complexity increases, so will the chances of hardware failures, operating system and application faults, operator/user errors, incompatible software patches and upgrades, thermal events, unclean power, and so on. With stringent system availability requirements—for example, five minutes downtime per year—data centers will require new techniques for effective fault tolerance and quick failure recovery.

OPERATIONAL COST REDUCTION
As data centers grow to handle more traffic, more complex services, and more business-critical functionality, the operation and management (O&M) costs become dominant in the total-cost-of-ownership (TCO) equation. Approximately 80 percent of the current TCO in large data centers relates to O&M costs, and the percentage is only likely to increase as the hardware costs decrease and the software/operational complexity increase. Ultimately, their impact on the TCO will determine the real worth of various emerging technologies.
This has given birth to the field of *autonomics*. IBM, the pioneer in this area, has espoused a vision of a fully automated data center that would provide four capabilities:

- self-monitoring/reconfiguration,
- self-optimization,
- self-healing, and
- self-protection.

However, realizing this vision is a long way off and would require unprecedented multidisciplinary collaborative research. As difficult as it might be, finding effective solutions to isolated problems in the field of autonomics addresses only a small part of the overall problem—the true challenge being integrating a variety of mechanisms and solutions and making them work collaboratively as an organism.

**IN THIS ISSUE**

This special issue includes four articles that address some of these important concerns. In “SoftUDC: A Software-Based Data Center for Utility Computing,” Mustafa Uysal and colleagues discuss how to bring disparate systems under a single umbrella of control, aggregating them into a centrally managed pool of server resources. SoftUDC relies on a virtual machine monitor that runs multiple virtual machines on each physical server, abstracts each virtual machine’s view of its storage and networks, and binds virtual machines across nodes into virtual farms. The authors discuss the SoftUDC architecture and its automated management features including support for online maintenance, automated deployment of applications, and automatic load balancing.

“TCP Onloading for Data Center Servers” by Greg Regnier and coauthors explores issues surrounding TCP acceleration at high data rates. The authors review the challenges of TCP/IP processing, discuss the TCP/IP acceleration solution space, and introduce the notion of TCP/IP onloading. They describe the overall philosophy of TCP/IP onloading and present ingredient technologies that can address the system and memory latency challenges of TCP/IP processing. They also describe in detail two key projects—ETA and MARS—that have experimented with onloading technologies.

In “Recovery-Oriented Computing: Building Multitier Dependability,” George Candea and coauthors explore building systems that can recover from failures using a range of options. At one extreme, they suggest microrebooting as a lightweight method for recovering software systems from transient and intermittent failures at virtually no cost in downtime and lost work. At the other extreme, they suggest a system-level undo that rewinds the entire system’s state, retroactively fixes the system, and then replays the state back to the present.

Finally, in “Power and Energy Management for Server Systems,” Ricardo Bianchini and Ram Rajamony address the increasing challenges of power and thermal management of servers in large data centers. The authors discuss the motivation for power and thermal research, survey previous work and ongoing efforts, and discuss the challenges that lie ahead.

We thank all the authors for their submissions to this special issue. In addition, we extend our thanks to the referees for their prompt and professional review of these submissions.

**Krishna Kant** is a senior engineer with the advanced enterprise architecture group at Intel’s Hillsboro, Oregon, facility and currently holds a visiting position at the University of California, Davis. His research interests include fabric issues in emerging data center architectures, hardware acceleration of networking protocols, performance modeling, congestion control, and traffic characterization. Kant received a PhD in computer science from the University of Texas at Dallas. Contact him at krishna.kant@intel.com.

**Prasant Mohapatra** is a professor in the Department of Computer Science at the University of California, Davis. His research interests include wireless networks, sensor networks, Internet protocols, and quality of service. Mohapatra received a PhD in computer engineering from Pennsylvania State University. Contact him at prasant@cs.ucdavis.edu.