WHITE PAPER
JUNE 2017

IT@INTEL

Disaggregated Servers Drive Data Center Efficiency and Innovation

Intel IT's breakthrough server design allows independent refresh of CPU and memory without replacing other server components, which results in faster data center innovation and 44-percent cost savings compared to a full-acquisition refresh.

Shesha Krishnapura
Intel Fellow and Intel IT
Chief Technology Officer

Shaji Achuthan
Senior Staff Engineer, Intel IT

Vipul Lal
Senior Principal Engineer, Intel IT

Ty Tang
Senior Principal Engineer, Intel IT

Executive Overview

The first groundbreaking server innovation in over a decade, Intel IT's disaggregated server architecture has the potential to dramatically change how data centers around the world perform server refreshes—leading to significant refresh savings and the opportunity to quickly take advantage of the latest compute technology. This technology is already being used in Intel's data centers in Santa Clara, California, which feature the world's lowest power usage effectiveness (PUE) rating of 1.06.

At the heart of the new design is the ability to independently refresh a server's CPU and DRAM, leaving the rest of the server enclosure untouched. This means it is no longer necessary to replace perfectly good fans, power supplies, cables, network switches, drives, and chassis.

Having already installed more than 40,000 disaggregated servers at Intel's data centers, Intel IT has found that the disaggregated design offers the following benefits[1]:

- Cuts refresh costs by a minimum of 44 percent
- Contributes to an extremely low PUE of 1.06
- Reduces technician time spent on refresh by 77 percent
- Decreases refresh materials' shipping weight by 82 percent

The ability to spend less time and money on refreshing servers means Intel IT can afford to refresh faster, bringing the most advanced Intel® Xeon® processor-based technology into Intel's data centers. We are excited about the resulting opportunities to boost data center efficiency and more effectively power Intel's silicon design jobs.

[1] Based on internal testing, March 2017.

Contents

Acronyms

CAPEX	capital expenditures
OPEX	operational expenditures
PUE	power usage effectiveness
QoS	quality of service
SLA	service-level agreements
TCO	total cost of ownership

Intel's compute, storage, and networking demands increase up to 40 percent annually.

The Challenge: Getting More Compute Power While Increasing Efficiency

At Intel, the data center is the heart of Intel's product design process—it is critical that we keep up with compute, storage, and network demand from Intel's business units. And yet, like many IT departments, we are pressured to meet growing compute and storage needs without increasing expenditures, and to increase the overall efficiency of our data centers.

Our business mandates are as follows:

- Meet up to 40 percent annual growth in compute, storage, and networking with a fixed physical space and power budget
- Take advantage of the latest compute technology (CPU and memory) without upgrading the entire data center infrastructure
- Continuously lower total cost of ownership (TCO) without negatively impacting service-level agreements (SLAs) and quality of service (QoS)

We have developed an internal IT strategy to accomplish these mandates that rests on three pillars:

- Providing Intel business units with the best possible SLAs and QoS
- Continuously minimizing IT infrastructure costs
- Optimally increasing the resource utilization of infrastructure assets

Ultimately, we strive to future-proof our data center investments by incorporating next-generation technologies; deliver more computing capacity with the same power-per-rack budget while maintaining or lowering data center power usage effectiveness (PUE); and continuously lower capital (CAPEX) and operational expenditures (OPEX) without decreasing SLAs and QoS.

Intel IT refreshes data center servers every four years to take advantage of new innovations and increased processor performance. Back in the late 1990s, we greeted the rackmount server with enthusiasm, as the new design contributed to maximizing use of space in the data center. When blade servers came along a few years later, we again embraced the new technology in our pursuit of data center space and energy efficiency.

But since that time, although Intel® processors, memory technology, and networking technology have continued to evolve, the server itself stagnated. The basic blade server design uses a shared power supply and shared network, but each blade has its own CPU with associated DRAM, along with direct attached storage (DAS) supported by either SAS or SATA drives and controllers.

We refresh our blade servers to take advantage of improvements in the Intel® Xeon® processor, with more cores, better performance per core, or more DRAM per core. But historically, we have had to replace the entire server—even though many components such as the chassis itself plus cables, power supplies, network switches, fans, and I/O components such as solid state drives (SSDs) and SAS drives still have many years of useful life remaining.

It seemed that this represented a terrible waste. Why replace so many server components that do not change from one processor generation to the next? Why fill the recycling center or landfill with perfectly good drives and components?

These questions led Intel IT to reimagine server design, leading to the first server innovation in more than a decade: the disaggregated server.

The Solution: Decouple CPU/DRAM and NIC/Drives Modules from Other Server Components

According to a recent IDC data center research report,[2] two-thirds of U.S. enterprise data center facilities have a PUE over 2.0, wasting money on uncontrolled cooling and power costs. The PUE measure divides total power delivered to the data center by the actual power the IT equipment consumes. An ideal PUE is 1.0, meaning that all of the energy needed for a data center facility goes to the computing devices instead of overhead such as cooling or power conversion. Intel is committed to operating efficient facilities, including the world's most efficient data centers located in Santa Clara, California, with a PUE rating of 1.06.[3]

To further optimize the world's most efficient data centers, Intel IT searched for ways to maximize the number of servers that can be fitted in a nine-foot rack while consuming a minimum amount of power.

Concept Engagement to Full Large-Scale Production Delivery in Five Weeks

When Shesha Krishnapura, Intel Fellow and Intel IT Chief Technology Officer, first presented his idea for the disaggregated server in 2016, his idea was met with skepticism.

"An organized skepticism is part of the process of innovation," said Krishnapura. "Others must try to disprove that your idea is not good enough."

But when colleagues looked closely at his design, they became convinced of its worth. With this backing, in June 2016 Krishnapura approached one of Intel's suppliers and told them he had a simple—but groundbreaking—idea.

"It is rerouting the motherboard, moving some of the components from the left side to the right side and adding this connector," Krishnapura told the supplier. "The cost should be minimal and we should be able to do this very fast."

"Very fast" does not even begin to describe the pace at which the new design came to life. Krishnapura says the supplier used its vertically integrated full-service capabilities and collaborated closely with Intel IT to deliver a solution—an optimally tuned, high-quality product with full supply chain and large-scale delivery support—in five weeks. Within a few more weeks, several thousand of the new servers were installed and running Intel® silicon design jobs.

"With 280 Intel® Xeon® processors-based server blades packed into a nine-foot rack, the high-density, high-efficiency, and disaggregated architecture is a game changer," said Krishnapura. "For the first time it allows for the independent refresh cycles of the server compute modules. This will unleash a new wave of disaggregated hardware architecture."

[2] Quinn, Kelly. "Power Issues in the Datacenter: IDC Survey Results". IDC Doc# US40885516. March 2016.

[3] King, Rachael. "Intel CIO Building Efficient Data Center to Rival Google, Facebook Efforts". Wall Street Journal. November 9, 2015.

"The disaggregated server architecture is a perfect fit for our data centers. Just like when a homeowner upgrades lighting, replacing only the bulbs with the most energy-efficient ones without replacing the entire lighting fixture, Intel IT prefers to upgrade just the compute modules with the latest technologies without replacing the entire server infrastructure."

–Shesha Krishnapura
Intel Fellow and Intel IT CTO

The answer is disarmingly simple: separate the CPU/DRAM module and the NIC/Drives module on the motherboard. Redesigning the server to be modular enables us to upgrade the CPU/DRAM module while retaining the other components that are not ready for end-of-life.

We designed and built a patent-pending new approach to server hardware (Figure 1): a disaggregated server that dovetails with our commitment to deploying Intel® Rack Scale Design (Intel® RSD) throughout our data centers (see the sidebar, "Disaggregated Server Architecture Complements Intel® Rack Scale Design"). This innovative approach to refresh enables us to affordably increase compute and storage performance and/or capacity with the latest generation of Intel® processors, without replacing reusable components.

The new disaggregated design makes server refresh a whole new experience. Instead of spending many hours on a refresh, we can now simply remove four screws, slide the CPU/DRAM module out, and install the new CPU/DRAM module. This module connects to the PCIe slot, which supports multi-generational drives (SAS drive, SATA drive, or Intel® Solid State Drives, including NVMe drives).

As described in detail in "The Result: Faster, More Efficient Refresh" section, we estimate that replacing only the CPU/DRAM module cuts our refresh costs by at least 44 percent (based on internal testing). Spending less on refresh means we can refresh more often. And having the latest generation of processors in our data centers means we can keep pace with compute demand and meet our SLA and QoS goals. While we will still need to balance the benefits of refresh against IT budget limitations, the refresh savings enabled by disaggregated servers makes that balancing act far easier.

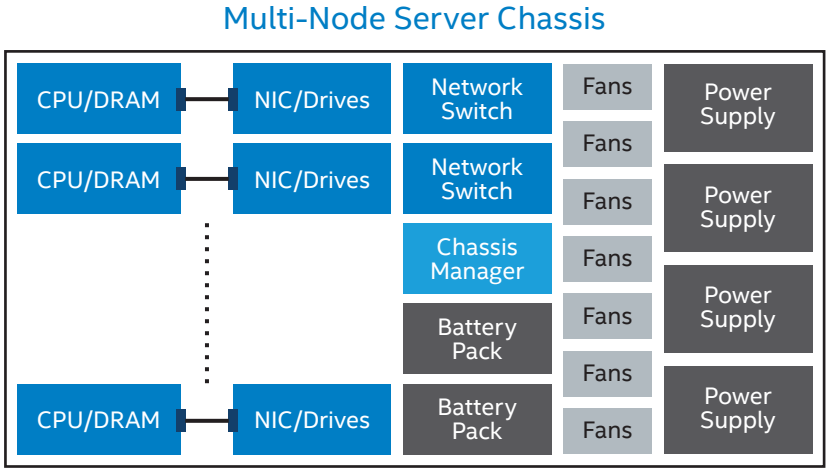

Figure 1. The disaggregated server architecture is characterized by a CPU/DRAM module and a NIC/Drives module that can be refreshed independently of each other and of the rest of the server components.

Intel's Santa Clara data centers feature the world's lowest PUE of 1.06.

Future-proof, disaggregated Intel® Rack Scale Design-ready architecture: incorporate the next-generation CPU/DRAM module without changing the rest of the design.

The Result: Faster, More Efficient Refresh

Electronic design automation (EDA) workloads are compute-intensive and require many servers to rapidly complete complex simulations. Shortening the design cycle reduces time to market and therefore creates a competitive advantage for Intel.

To support the business mandates discussed earlier while minimizing IT infrastructure costs, Intel IT has rapidly adopted the disaggregated server architecture, deploying more than 40,000 server blades in its Santa Clara data centers, two of which feature the world's lowest PUE of 1.06. In addition to industry-leading server density and power efficiency, the new innovative architecture enables the independent upgrade of the compute module without replacing the rest of the server enclosure including networking, storage, fans, and power supplies, which refresh at a slower rate.

By disaggregating CPU and memory, each resource can be refreshed independently, allowing data centers to reduce refresh cycle costs. This is similar to how a homeowner who needs a more efficient and powerful lightbulb does not have to change the entire light fixture, switch, and wiring—the homeowner simply installs the latest lightbulb technology.

When viewed over a three- to five-year refresh cycle, the disaggregated server design can deliver, on-average, higher performance and more efficient servers at lower costs than a traditional rip-and-replace model by allowing data centers to independently adopt new and improved technologies. Also, the disaggregated servers we have installed are designed for advanced airflow and cooling. The ambient temperature for these servers can be as high as 40°C (104°F). Green computing features such as this give Intel IT the opportunity to operate their data centers more efficiently.

Refresh Cost Savings

We estimate that using disaggregated servers can cut refresh costs by a minimum of 44 percent. We will now be able to refresh more frequently, putting the latest, most advanced Intel Xeon processor-based technology to work for Intel's design teams.

Figure 2 illustrates how disaggregated servers can cut refresh costs. Consider a 3U chassis with 14 blades. Refreshing that chassis by replacing all the blades but keeping the chassis itself along with the networking switch, power supply, and fan modules, saves 17 percent compared to a full-acquisition (rip-and-replace) refresh. But with disaggregated servers installed in the data centers, it is possible to refresh only the CPU/DRAM module, saving 44 percent compared to a full-acquisition refresh. (These results are based on internal testing at Intel and serve as an example only.)

In addition, there is no need to reinstall the OS or spend time replacing parts unnecessarily. In our internal tests, we determined that disaggregated servers represent a 77 percent reduction in technician time due to far fewer handoffs and required skill sets (see Table 1). Faster refresh of CPU and memory is also expected to reduce maintenance and downtime issues. We anticipate about USD 1 million per year in OPEX savings due to the 40,000 high-density disaggregated servers already installed in our data centers.

Table 1. Faster Refresh Is Now Possible

Old Method	New Method
Six different technician skills, five handoffs: • Data center manager • Physical rack and stack technician • Network cabling technician • Network configuration engineer • Server/OS configuration engineer • Batch clustering administrator (for new system name configuration)	Two technician skills, one handoff: • Board replacement technician • Server/OS configuration engineer
35 hours of work time per rack[1]	**8 hours of work time per rack[1]**

[1] Based on internal testing.

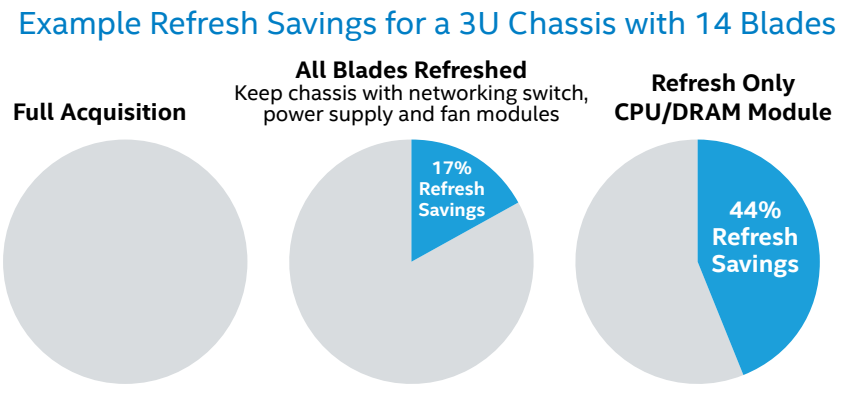

Figure 2. Refreshing the CPU/DRAM module in a disaggregated server saves at least 44 percent compared to a full-acquisition (rip-and-replace) server refresh. Based on Intel internal testing, March 2017.

Disaggregated Server Architecture Complements Intel® Rack Scale Design

Intel® Rack Scale Design (Intel® RSD) is the blueprint for the software-defined hyperscale data center. It is a logical architecture that disaggregates compute, storage, and network resources, and introduces the ability to more efficiently pool and utilize these resources. This approach enables dynamic composition of data center resources based on workload-specific demands. A common management framework exposes resources to an orchestration layer, which makes the data center infrastructure more flexible, simpler to manage, and easier to scale out as required. Pooled resources can deliver increased workload performance, while data center operations benefit from analytics-based telemetry.

Intel RSD provides a computing, storage, and network backbone that combines with virtualization and cloud-based computing to usher in an age of truly agile digital infrastructure.

See more at intel.com/content/www/us/en/architecture-and-technology/rack-scale-design-overview.

Environmental Considerations

Another benefit from using disaggregated servers relates to environmental conservation. Our internal testing indicates that ordering refresh components for a disaggregated server can save 86 percent in volume (meaning fewer boxes to ship and to store and stage) and 82 percent in shipping weight (meaning less shipping costs). Newer processors generally have lower power and cooling requirements than previous generations of processors. Avoiding recycling still-useful components such as SAS drives and fans, combined with fewer shipping materials and less fuel and time spent transporting the new parts, can contribute to less waste and a smaller carbon footprint.

Conclusion

Just as it makes little sense to replace an entire light fixture when all that is needed is a more energy-efficient and powerful light bulb, replacing an entire server does not make sense if all that is needed is a more advanced CPU and DRAM. The future-proof, disaggregated server architecture gives Intel IT the flexibility to upgrade the CPU and DRAM more quickly while preserving the existing investments made in the networking, drives, power supplies, and cables. This disaggregated approach to server refresh results in the following:

- Lower CAPEX (IT spends only a fraction of what otherwise would be spent on refresh)
- Lower OPEX (replacing a module involves less work and manpower than replacing the entire server)
- Overall lower data center TCO

For CAPEX alone, we estimate that the savings is at least 44 percent.

The development of the disaggregated server is poised to bring huge advantages to the IT industry. Intel IT is already reaping the benefits associated with cost efficiency, material savings, environmental responsibility, shipping costs, supply chain efficiencies, and more. Our end customers—Intel's business units—will be thrilled to have the most advanced processor and memory technology at their fingertips. Also, server vendors, suppliers, and the rest of the ecosystem will benefit as well.

For more information on Intel IT best practices, visit **intel.com/IT**.

Receive objective and personalized advice from unbiased professionals at **advisors.intel.com**. Fill out a simple form and one of our experienced experts will contact you within 5 business days.

IT@Intel

We connect IT professionals with their IT peers inside Intel. Our IT department solves some of today's most demanding and complex technology issues, and we want to share these lessons directly with our fellow IT professionals in an open peer-to-peer forum.

Our goal is simple: improve efficiency throughout the organization and enhance the business value of IT investments.

Follow us and join the conversation:

- Twitter
- #IntelIT
- LinkedIn
- IT Center Community

Visit us today at **intel.com/IT** or contact your local Intel representative if you would like to learn more.

Related Content

If you liked this paper, you may also be interested in these related stories:

- Interview with Shesha Krishnapura on Inside IT
- Intel® Rack Scale Design web site

Printed in USA Please Recycle 0617/DPEA/KC/PDF