

This article was published in ASHRAE Journal, September 2018. Copyright 2018 ASHRAE. Posted at www.ashrae.org. This article may not be copied and/or distributed electronically or in paper form without permission of ASHRAE. For more information about ASHRAE Journal, visit www.ashrae.org.

Part 2

ASHRAE IT Power Trends

BY DONALD L. BEATY, P.E., FELLOW ASHRAE; DAVID QUIRK, P.E., MEMBER ASHRAE

The occupants (software) of the data center continues to change and present new changes to the power trends of a given data center. The previous column discussed how ASHRAE's Datacom Series' *IT Equipment Power Trends*, third edition, included a new section on workload (software applications).

Historically, software applications have been a function of people. In other words, the applications were primarily "on" when people were using the service or applications. Examples include:

- Social media;
- Financial transactions and investing;
- Internet searches; and
- E-commerce.

Because not everyone in the world was using every application at the same time with great intensity, the occupants (software) of the data center have had lots of diversity in their workloads.

However, the applications have now changed with the introduction of neural networks, artificial intelligence, and blockchain applications. These applications are additive and/or supplemental to most every other application we utilize today. Further, each of these applications can work in the background 24/7 with or without people using them.

In this column we'll discuss the disruptive nature of these new software applications and how they will change the power trends assumptions for

every given data center and all other software applications.

Workload Types

As discussed in the last column, the transition to using Workload type was brought about by the recognition that "software is the real load in data centers" (Beaty, 10/2015). These Workloads and their software drive server energy usage composed of a mix of both transient and steady-state components.

At the high end of the energy usage trends is the scientific workload. This is composed primarily of steady state server utilization at a high level since they are continuously "crunching numbers."

Toward the lower end is the business processing workload, where processor utilization is composed of longer idle periods during data entry followed by spikes after the data has been entered.

Donald L. Beaty, P.E., is president and David Quirk, P.E., is vice president of DLB Associates Consulting Engineers, in Eatontown, N.J. Beaty is a member and Quirk is research subcommittee chair of TC 9.9, Mission Critical Facilities, Data Centers, Technology Spaces and Electronic Equipment.

However, these new software applications (neural networks, artificial intelligence, blockchain, etc.) are changing the landscape to an environment where all workloads can continuously be high levels of utilizations.

Neural Networks

Artificial neural networks are a key stepping stone in the search for artificial intelligence. These digital networks are capable of learning from a combination of sensory inputs, outside stimulus, etc., thus allowing adaptive learning.

Artificial neural networks attempt to recreate human learning systems on computers by constructing a simple framework program to respond to a problem and receive feedback on how it responded. Unlike humans though, a computer can optimize its response by doing the same problem millions of times and adjusting its response according to the feedback.

There are many variations of neural networks, each with their own specialty purpose at solving various problems.

Neural Network Applications

Some of the recent neural network applications include:

- Color rendering of black & white images and videos;
- Enhancing low resolution images and videos;
- General adversarial networks;
- Lip reading;
- Process modeling and control;
- Machine diagnostics;
- Target recognition;
- Medical diagnosis;
- Plug & play generative networks;
- Object recognition;
- Image classification (e.g., face recognition);
- Automotive engineering;
- Applicant screening of jobs;
- Credit rating;
- Targeted marketing;
- Financial forecasting;
- Intelligent searching; and
- Fraud detection.

The key for data center planners and designers is to understand that these new applications are emerging and that they can work 24/7 on computers as overlays to virtually any other type of software.

Artificial Intelligence

Artificial intelligence (AI) is a much broader research field of computer science that uses many other software applications such as artificial neural networks, deep learning, and big data to produce “intelligence” in computer-based machines and robotics.

Artificial neural networks are actually a smaller subset of AI. The science of AI seeks to solve many other problems with reasoning, problem solving, knowledge representation, planning, learning, natural language processing, motion, social and general intelligence. The depth of the mathematics and computations that goes into AI applications is well beyond this column, however, the point to be made is that AI is relevant to all topics and all software applications for all humans and machines. The related impact on the workloads of the servers in the data centers will be correspondingly significant.

Artificial Intelligence Applications

Artificial intelligence applications include anything that requires intellectual tasks. Today’s modern AI techniques are found everywhere and are too many to list here.

High profile examples of AI include autonomous vehicles (such as drones and self-driving cars), medical applications, social media bots, playing games (like Chess and Go), internet search engines, online and mobile app assistants, spam filtering, and targeted online advertising.

Less publicized applications include:

- Financial institutions and investing companies;
- Video games;
- Military applications for lethal autonomous weapons, unmanned combat vehicles;
- Auditing tools (financial records, taxes, etc.);
- Advanced video surveillance security systems (facial recognition, social behavioral patterns, etc.);
- Medical diagnosis and surgical robots; and
- Automated online personal assistants.

There is virtually no topic or software application that will escape the AI reach and related impacts.

Blockchain

Shortly after the 2008 financial collapse, the first blockchain was conceptualized in a white paper entitled, “Bitcoin, A Peer-to-Peer Electronic Cash System,”

published by a mysterious entity that called themselves Satoshi Nakamoto. Though Bitcoin has received a lot of the headlines, blockchain is far more than this single cryptocurrency application first described back in the 1990s.

Blockchain, in simplest terms, can be summed it up in one of the following ways:

1. Peer-to-peer exchanges of goods, services, contracts, anything without middle-men and verification systems and processes.
2. It's the optimization and securitization of the exchange of everything.
3. It's a continuously updated database of who holds what and it's secured in the blockchain using cryptology.
4. It's the next evolution of a payment processing system (likely to replace all other forms of payment we use for any transactions today).

This software application (and distributed database) is projected to disrupt every industry and everything about our lives in the current and coming decades.

Thus far, blockchain has been called many things:

1. Trustless system of value;
2. The next stage of the internet;
3. The next evolution of payment; and
4. The most important technology ever invented.

For the purposes of the data center industry, the implications may be far-reaching. Some of the basic implications include:

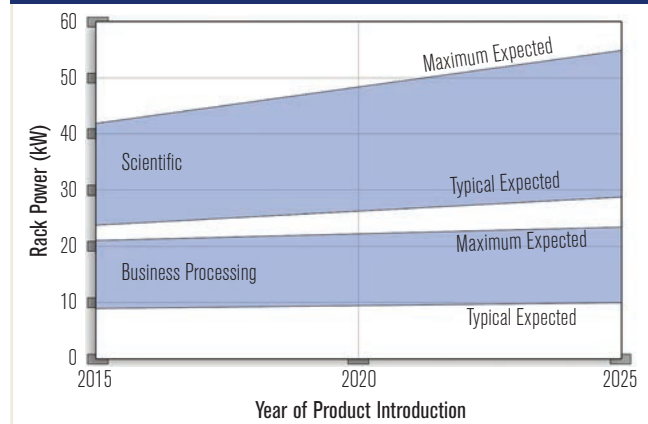
1. A significant shift in how data is stored and transactions are processed; this will shift who, what, when, how, and where data is managed.
2. Elimination of digital copies of anything on the internet (the current day approach of information sharing and exchange of goods and services on the internet). This could be transformation on the needs for storage vs. compute power in data centers.

Since it is so early in the evolution of blockchain, predictions on the implications are also premature.

Blockchain Applications

The applications for blockchain are seemingly endless. The most common applications that are emerging fit under the heading of smart contracts. Basically, any goods or services exchanged today but without any middlemen in the transitions, at all. This can include money, real estate, goods, services, music, and just about anything else.

FIGURE 1 Scientific and business processing workloads 42U rack power from *Power Trends*, Third Edition.



Blockchains have been categorized into public, private, and consortium. Regardless of these types, the underlying technology and approach is the same.

Software Workload Types

As covered in the previous column, trends by workload type are provided with boundaries for maximum expected and typical expected rack power levels based on the year of project introduction (Figure 1).

Historically, the biggest misconception of *Power Trends* was that the forecasted maximums were the typical. Arguably, this led to many over-provisioned data centers as many engineers designed to these forecasted maximums. The delta between the typical expected and maximum expected is significant as shown in Figure 1.

The range between typical expected and maximum expected requires some iteration and understanding of how a client configures their equipment in a rack, how much utilization is expected, and the workloads of the software on that configured hardware.

What's important to understand is that the historical perspective is about to change yet again with these new software applications. In many instances, it can be expected that the maximums will be achieved and sustained.

Aside from those applications covered in this column, there are many others under the heading of machine learning and AI, such as:

- Pattern recognition;
- Reinforcement learning;
- Statistical inference;
- Probabilistic machine learning;
- Supervised and unsupervised learning;

- Ensemble algorithms;
- Structured prediction; and
- Clustering.

Conclusions

As covered in the last column, the third edition of *Power Trends* provides much more detailed information that can be utilized by data center owners and design engineers to develop realistic power growth projections for use in planning their data centers. The software *applications* running on the equipment will dictate power densities and they will *never* be the same as these published *hardware* trends.

More than ever, it is important to understand the potential workloads in the data center. New software applications such as neural networks, artificial intelligence, and blockchain are going to push the boundaries of the power densities in data centers.

These new “machine-based” software applications consume far more data than prior human-centric applications. Further, they are not bound by time of day or day of the week for their processing needs. These new applications can work at 100% 24/7 and will begin to push the boundaries of the hardware much further.

As the IT infrastructure is pushed further on the boundaries, so will the MEP infrastructure that supports it. The ability for the MEP infrastructure to support higher densities and more transient workloads will become more prevalent.

Armed with a deeper understanding the software applications

in a given data center, data center owners and design engineers are better equipped to plan their next data center growth project.

References

1. Beaty, D.L. 2015. “Data center loads – unknown, unseen, unheard. why you must know what your software is doing.” *ASHRAE Journal* (10). ■

Advertisement formerly in this space.