White Paper

Rethinking Your Infrastructure for Enterprise AI

Sponsored by: IBM

Peter Rutten
June 2018

IDC OPINION

IDC strongly believes that the days of homogenous compute, in which a single architecture dominates all compute in the datacenter, are over. This truth has become increasingly evident as more and more businesses have started to launch artificial intelligence (AI) initiatives. Many of them are in an experimental stage with AI and a few have reached production readiness, but all of them are cycling unusually fast through infrastructure options to run their newly developed AI applications and services on.

The main reason for this constant overhauling of infrastructure is that most of the standard infrastructure that is being used in the datacenter for the bulk of workloads is not very suitable for the extreme data-intensive nature of AI. Not only is the performance and I/O of a typical server lacking for deep learning (DL) but the data lakes that are the breeding grounds for AI model development are unequipped for this critical task. The data lakes consist of slow monocultures based on traditional schemas that take weeks if not months to prepare for AI modeling. These data lakes are also considered noncritical for the business, whereas, once AI starts being developed on them, they will become hypercritical.

AI has thus become the lead actor in a play that tells the evolving story of emerging processor diversity in the datacenter — a diversity that manifests itself not only with the increasing presence of GPUs, FPGAs, many-core processors, and ASICs for specific workloads but also in a shift to other host processors and to better links between the host and the accelerator. Because, while accelerators can alleviate a lot of performance lag, it's in the interplay with a host processor that truly outstanding performance for a workload such as AI can be achieved.

This white paper discusses these challenges and looks at how IBM proposes to help businesses overcome them.

SITUATION OVERVIEW

We have moved from saying that "AI is coming" to "AI is here and now." AI is a cataclysmic shift that is causing not just IT in organizations but also the CIO, the CTO, and the CEO to rethink their business, urgently asking themselves: Is there a way that we can leverage AI? How do we use AI? How do we get there? While there is a lot of focus on the last 5–10% of AI, which is DL and training, there is less awareness about how to prepare for the AI revolution, that is, how AI is going to affect the way businesses manage their data and how it will affect their infrastructure.

There is plenty of data to work with. The data deluge is coming from multiple ever-larger spigots, including connected cars, wearable health devices, connected machines, and sensor-enabled devices.

This big data revolution has in a way been successful in that many businesses today have stood up a data lake. What they often have not yet figured out is how to get value out of their data lake. It can take weeks, sometimes months, before a line-of-business (LOB) manager gets useful information from the organization's data scientists out of the data lake.

For sure, today's data lakes are providing some insights, but they are difficult to leverage, and they are certainly not the best foundation upon which to build AI applications. As such, the AI revolution is an invitation to revisit big data and, this time, do it better to support new AI services. The data deluge is fueling capabilities around machine learning and especially deep learning to create and train intelligent systems that can discern patterns and reveal insights. The ultimate data-intensive workload today is deep learning at scale, and the question is: What server and software platform will businesses need to fully embrace AI?

Data acumen and expertise and the ability to extract value out of data enable a company to revolutionize the way it operates its business and interacts with its customers. AI is a way to truly comprehend the deeply hidden intricacies of the space that a company is operating in as a business or an organization. AI and especially DL represent an acceleration of traditional data analytics, providing a wealth of new value. But DL also presents numerous challenges. It is a rapidly evolving collection of technologies with very fast innovation, especially in open source, making it difficult to keep up with and to get all the pieces to work well together.

The challenges are with choosing the right infrastructure for DL, the right software — software that is not only easy to use and quick to deploy but also well integrated with the hardware and fully supported by the vendor — and with preparing the data.

This latter challenge, preparing the data for DL, is by far the biggest bottleneck on the path to AI. Businesses have found that by the time they get to the DL training and to optimizing the infrastructure around the training, months have been spent on getting the data ready. Much of the time spent in setting up an AI process is in transforming the data and getting it from point A to point B where it can be used. What's complicating things further is the fact that a truly effective AI application draws data from multiple sources — internal data, streaming data, and so forth — and combines that data in meaningful ways.

THE IBM AI INFRASTRUCTURE APPROACH

AI requires leveraging existing data stores and analytics that need to connect into a deep learning workflow. Understanding this, IBM has focused on those aspects that cause businesses to spend inordinate amounts of time building an AI environment, which, first and foremost, is about preparing the data.

IBM set out to streamline the process of preparing the data for AI applications and of building connections and synthesizing the various data sources for those AI applications. Data sources for AI include various types: those that are under a business' control (e.g., customer preferences) and that draw from existing patterns of customer behavior and those that are external, such as streaming data (e.g., social media behavior).

IBM takes the point of view that for AI infrastructure and for preparing for AI, businesses need to rethink the ways they are managing their data. To help businesses in this transition, IBM has introduced a concept called "AI Infrastructure," which consists of a single platform for data pipelines, services, and AI. Essentially it is an end-to-end server, storage, and software platform designed for businesses to stand up their version of the AI revolution.

This platform has as its foundation a modern data lake with improved storage, enterprise-grade Hadoop and Spark, and enhanced data management. This foundation supports data platforms that embrace multiple schemas (RDBMS, NoSQL, graph), multiple architectures (CPU, GPU accelerated, in memory), and uncompromising system performance. Together, the better data lake and the data platforms enable a flexible IT environment with dynamic storage, I/O, and memory that deliver services for real-time insights such as CRM, IoT, and fraud detection, to name a few. These building blocks then become the staging ground for deep learning training to infuse existing applications with AI and to develop new AI-driven apps.

At the Heart of AI Infrastructure: The IBM Power System AC922 and IBM PowerAI

IBM Power System AC922

At the heart of this AI Infrastructure is IBM's current accelerated POWER9 system, which IBM launched in December 2017. Branded as IBM Power System AC922, it is designed specifically for AI workloads. The Power AC922 is a two-socket server that has been built up with advanced I/O interfaces and with hardware and software that are optimized and fine-tuned for AI acceleration. It is a performance beast that is also the backbone of the CORAL Summit supercomputer. The system includes various extremely fast I/O architectures, including PCIe Gen 4, CAPI 2.0, OpenCAPI, and NVLink, all of which are ideal for very data-intensive workloads.

The two POWER9 CPUs each have four-way multithreading for host CPU parallel processing and from 2 to 6 NVIDIA Tesla V100 GPUs with NVLink, delivering extreme performance between CPUs and GPUs for HPC, deep learning, and AI workloads. NVLink is integrated into the processor (labeled "NVLink 2"), which further increases the performance throughput of the system. The system provides full memory coherency and enables accelerated applications to use system memory as GPU memory, overcoming the 16GB or 32GB memory limitations of GPUs. A Power AC922 configuration can be anywhere from 16 to 44 cores.

IBM PowerAI

IBM PowerAI is what IBM calls an "enterprise offering for distributed deep learning." Built on top of the IBM Power S822LC for HPC or the IBM Power AC922, PowerAI comes packaged with open source deep learning frameworks and tools such as Caffe, Torch, TensorFlow, Theano, and Chainer. Also included are supporting libraries such as DIGITS, OpenBLAS, distributed frameworks, Bazel, and NCCL. The purpose of this hardware/software package is to give businesses a fast way to get started with deep learning by providing enterprise-ready and -supported distributions of critical open source software, high-performance hardware, and integrated tools for data scientists that will boost their productivity.

What IBM aims to do with PowerAI is to help reduce the complexities of building a smoothly working AI infrastructure, ensuring that it is easy to get an optimized deep learning training framework in place — easier to consume so that a data scientist isn't wasting time trying to debug a TensorFlow package. The result is an approximately 45-minute installation of a complete DL environment from bare metal to something that you can begin training on.

IBM has also announced an update to its PowerAI framework that brings support to the entire stack — from bare metal all the way up to the deep learning environment. This full-stack support for the PowerAI platform is good news for enterprise customers that want to embrace AI in general and DL specifically but that are wary of taking on responsibility for supporting an environment that is primarily driven by open source project code.

Until recently, the PowerAI platform was built around IBM's accelerated system Power S822LC for HPC, which is also being used in two of the largest supercomputers currently being built. It is designed to allow data to flow as freely as possible. In the Power S822LC for HPC, acceleration with GPUs is at the heart of the server design. It connects every compute engine on the server with point-to-point high-speed NVLink connections.

In the Power AC922, with NVLink, the physical connections are part of the processor die, making the GPU more like a peer to the CPU, rather than dependent on the CPU, in terms of how data is handled. The dual NVLink connections between GPUs and CPUs provide extremely fast direct and near-coherent access to the system memory (unlike designs that rely on a shared PCIe 3.0 bus or switch). These lower-latency connections allow for larger, more complex neural models and much larger training data sets.

The Foundation for AI Infrastructure: A Better Data Lake

While businesses can evaluate whether the Power AC922 should become the heart of their AI infrastructure deployment, they need to keep in mind that the foundation of this environment should be a better data lake, one that is suited for AI services running on top of it. To achieve this better data lake, they need to consider five critical aspects:

- An AI era data architecture
- High-performance servers powering fast and enterprise-grade Hadoop and Spark
- Scale-out storage that is integrated, compatible, and enterprise ready
- Better and easier data management compared with current practices
- Support for multiple data platforms — schemas and architectures

The sections that follow dive deeper into IBM's approach to building a data lake that is ready for AI around these five elements.

An AI Era Data Architecture

Today's data lakes are bottlenecked by a data architecture that predates the "big data" and software-defined infrastructure revolutions empowering them. This is because whether you're dealing with the software-defined storage that stores the data, the services that analyze the data, or the infrastructure that moves the data, almost all of it is run on servers. And the data architecture for those servers is PCIe 3.0, which was launched in 2010, a year before the initial release of Hadoop initiated the big data revolution and at least two years before the Storage Networking Industry Association (SNIA) attempted to define standards around software-defined storage. While the processors powering most of the world's servers have become faster at crunching numbers, they have remained largely unchanged with regard to data movement and management, often referred to as I/O.

Faster Data Servers

Due to their PCIe Gen 3 buses, most of the servers powering the data lakes and analytics needed to fuel enterprise AI are bottlenecked by I/O constraints. IBM Power servers have greatly improved I/O bandwidth and performance. IBM Power introduced PCIe 4.0 interfaces in IBM POWER9 servers that are twice as fast as PCIe 3.0. This improves data movement within the server by significantly improving the performance of PCIe-attached storage and FPGAs. It also improves the performance of the clusters by doubling the speed of the network interfaces connecting them. Furthermore, for data lakes leveraging GPU-accelerated analytics, the second-generation NVLink interface enables an approximately 5–6x bandwidth advantage over PCIe Gen 3.

Easier, Faster, and Enterprise-Grade Hadoop and Spark

In most businesses today, Hadoop and Spark are not considered mission-critical applications, and in many organizations, the data lake is still considered nonessential to business operations. Once it becomes a platform for AI, however, the data lake will quite suddenly and dramatically change from being a support system to becoming an environment that runs the organization's mission-critical applications such as ERP or CRM.

This means that businesses will need to build their data lake with enterprise-ready, mission-critical hardware and software. IBM says that it validates and supports the entire IBM Power hardware and software stack to guarantee reliability. In terms of the hardware, because of its objectively greater per-core performance than available with other processor architectures, IBM Power allows for a few nodes to perform as well as or better than larger clusters on alternative architectures.

With regard to software, enterprise-grade readiness starts with the IBM Elastic Storage Server (ESS) platform, which eliminates various problems with Spark and MapReduce that are HDFS related. IBM also provides enterprise support with Hortonworks, a massively scalable open source platform for storing, processing, and analyzing large volumes of data from many sources and in various formats. Hortonworks includes MapReduce, HDFS, HCatalog, Pig, Hive, HBase, ZooKeeper, and Ambari.

Better Storage

IBM believes that IBM ESS is an ideal storage approach for an advanced data lake. IBM ESS is a software-defined storage solution that combines IBM Spectrum Scale software with IBM POWER servers and storage enclosures. IBM Spectrum Scale, the parallel file system that is at the heart of IBM ESS, scales system throughput as it propagates while still providing a single namespace. This allows for high performance while avoiding the creation of data silos and makes storage management easier.

IBM settled on ESS to provide an enterprise-grade alternative to HDFS, which is used for large, scale-out big data applications, such as Spark, MapReduce, and certain deep learning frameworks. HDFS can be inefficient because of its three-to-one duplication model and dedicated data silo. And while it is fast for simple things, it can be slow for large, complex processes. Also, HDFS has limited standard protocol support and is not easy to integrate with other enterprise infrastructure.

IBM ESS is powered by IBM Spectrum Scale, a software-defined storage often deployed in various industries to support compute-intensive parallel workloads. IBM ESS looks and feels like HDFS and performs HDFS very well; it supports multiprotocol access to data including Common Internet File System (CIFS), NFS, Object, block storage, and more. IBM Spectrum Scale has policy-driven data movement that can be used to tier flash, disk, tape, and cloud based upon data usage or other defined criteria. IBM ESS provides businesses with a storage platform that can execute big data analytics as well as other workloads while leaving data in place for use with other applications or for cost-effective archiving.

Better and Easier Data Management

Spark is an excellent open source big data analytics framework for deep learning, but implementing Spark is not a trivial exercise. Businesses need the right tools, skill sets, and workflow, as well as integration with other frameworks, to make it work efficiently and securely and ensure its manageability. IBM designed Spectrum Conductor to help businesses overcome these hurdles. Rather than putting together the various components of the environment themselves, businesses get an integrated and fully supported solution that includes a Spark distribution, supports multitenancy for Spark and other frameworks, and enables the kind of dynamic resource allocation on which HPC infrastructures thrive.

IBM Spectrum Conductor uses Spark as the compute and transport layer to pull data from the variety of data sources in the data lake. The solution subsequently adds vector information and metadata to the data based on user-provided definitions and then executes the entire data transformation. It will germinate hundreds of different Spark instances to go through the data from the various sources and transform the data into a data set that is ready for, say, Caffe or TensorFlow.

IBM Spectrum Conductor is a workload and resource manager optimized for analytics and deep learning workloads. IBM suggests that businesses use IBM Spectrum Conductor rather than YARN to process these workloads so that jobs are scheduled effectively with predictable runtimes and that they get the required resources, such as the right number of GPUs and CPUs, the right amount of in memory, etc. Based on the characteristics of the workloads, IBM Spectrum Conductor will determine which resources are available, how many of the resources are available, where the resources are available, and what the queue order should be. This maximizes utilization, which IBM states can reach 40% or better, significantly higher than the average datacenter utilization of approximately 20%. The result is a significant performance boost that allows Spark to run more effectively and efficiently.

The multitenancy feature of IBM Spectrum Conductor allows for multiple instances of Spark to be deployed for optimal resource utilization and greater scale and performance while eliminating resource silos that are tied to separate Spark implementations. The solution also facilitates integration of Spark with such application frameworks as Hadoop, MongoDB, or Cassandra. IBM Spectrum Conductor is a licensed, supported software package that end users can build into an existing or on a new cluster.

IBM Spectrum Conductor Deep Learning Impact is a deep learning environment designed to accelerate and simplify the life of the data scientist. It can run on any infrastructure, and IBM has developed GPU optimization within Spark, accelerating the compute inside the Spark instances. Given its higher CPU-GPU bandwidth than competing infrastructure, PowerAI can take advantage of this GPU optimization in a truly noticeable way. Hence the same PowerAI cluster used for deep learning training can be used for data preparation on IBM Spectrum Conductor in a kind of ebb and flow of preparation and training. On the IBM PowerAI Enterprise road map, the IBM Spectrum Conductor Deep Learning Impact features will become integrated in the PowerAI Enterprise package in 2Q18.

A Better User Interface

Most data lakes today are the domain of data scientists — requests for reports from the line of business require specialized skills and may take weeks or months to be generated. It is not unreasonable to assume that today's data lakes will evolve in the same fashion as yesterday's data warehouses. Twenty years ago, it took months to get useful information from a data warehouse. Then, however, tools emerged that enabled developers to gain access to the data warehouse, allowing them to develop applications that leveraged its data. Next, line-of-business people started using Excel plug-ins to access the data warehouse to generate business reports.

A similar "opening up" will likely happen for today's data lakes. Enterprise developers today will be looking at AI to revolutionize their apps. For that, they need access to the data in the data lake and to the organization's AI capabilities. Line-of-business users will not be far behind. The bottom line is that access to critical data in the data lake and the capabilities to leverage that data for AI will be democratized. This means, however, that the tools to achieve this will need to be improved — they will need to be easier to use and more intuitive.

To that end, IBM has developed DSX Local, which stands for Data Science Experience. DSX Local simplifies the ETL portion of using data lake data and connects to the appropriate deep learning framework. DSX Local works on-premises and in the cloud and not just on IBM Power but also on other processor architectures. DSX Local is an out-of-the-box on-premises enterprise solution for data scientists and data engineers that offers a collection of data science tools such as RStudio, Spark, Jupyter, and Zeppelin notebooks integrated with IBM technologies. The user interface has been designed to be as intuitive as possible, and the tool provides a collaborative projects space for data scientist and developer teams.

Support for Multiple Modern Data Platforms

When it comes to databases, one size doesn't fit all. While there are many tasks that can still be executed with a traditional relational database (e.g., CRM and ERP), an increasing number of challenges require modern schemas. Businesses that intend to build an AI infrastructure — executing AI training and inferencing on an improved data lake — will need to embrace more schemas than just SQL, specifically NoSQL (e.g., for IoT or content workloads) and graph (e.g., for fraud detection).

Also, if the businesses are investing in a new platform for their data, they should investigate modern architectures such as in-memory databases and GPU-accelerated databases for graph and NoSQL. Such data-intensive schemas require many threads, some of which will be processed in the CPU only, some in memory, and others in a GPU, depending on the schema and the database. When using a graph database, for example, the entire database needs to be in memory.

Enabling multiple data platforms will result in a powerful platform for AI. On top of the data lake, businesses can build a new class of accelerated open source databases such as Redis, MongoDB, and EDB Postgres; graph database Neo4j; and Kinetica, a distributed in-memory database for GPUs. IBM claims distinct performance advantages with POWER9 on these databases and goes as far as providing price-performance guarantees for them. These databases leverage Power's I/O, fast interconnect, built-in RAS, and large memory. Many of these new services will be in memory, requiring large amounts of memory, high performance, and flexible access to I/O as data is being moved back and forth at unprecedented levels.

Large Model Support

IBM believes that Power Systems is an ideal platform for running these modern schemas thanks to its four-way multithreading, per-core performance, I/O capability, and built-in NVLink that provides full coherence. NVLink allows a process on the GPU to leverage system memory with almost no performance penalty, which is a critical capability to overcome the maximum memory of 12GB, 16GB, or — since recently — 32GB that GPUs typically come with. This memory is not enough for AI vision, 4K video, or more complex AI models with multiple layers of matrices. As a result, businesses have to make trade-offs, such as using lower-resolution images, working with web-scale rather than high-definition video, or developing networks that are less deep than desirable.

With the new release of PowerAI, IBM is therefore introducing what's called "Large Model Support." On PowerAI, the entire model can be loaded into CPU memory. For example, in a four GPU system there can be four instances of a 230GB model instead of four instances of a 16GB or 32GB model, allowing for an entire data set to be loaded into the model rather than just subsets, problems to be resolved more accurately, or video files to be high definition rather than web scale. In other systems, a PCIe card connects the GPUs to the system memory, which causes a performance penalty because of PCIe's limited bandwidth.

Even sharing flash memory will become achievable because IBM Power features CAPI, with significantly higher bandwidth than PCIe 3.0, to enable connections to very fast SSDs that can then be used as memory at near-memory performance. This makes it feasible to put a significant portion of the data lake in memory on the SSDs. IBM announced Large Model Support as a technology preview as part of its Caffe release in August 2017 and is now starting to open source it to various framework providers, the first one being the open source deep learning Chainer framework from Preferred Networks in Japan.

FUTURE OUTLOOK

IDC expects that AI will develop very rapidly and that most businesses will need to adopt an AI culture in the next 12–24 months or risk being outsmarted by their competition. AI will not only develop fast but also infuse every workload in the datacenter or cloud. Therefore, a long-term strategic approach to AI is essential: How can AI improve my business? How can AI defend my business? How can AI make my business more competitive? How can AI make my business more efficient? There are dozens of such questions that need to be asked and answered. At the same time, AI development by data scientists and app developers needs to go hand in hand with a long-term infrastructure plan for AI.

When businesses design or redesign their data lakes for intelligent services, they should assume that they don't yet know where they are going to be implementing AI. There are many potential scenarios, from training across the data lake while correlating with real-time services to running inference models that are correlating with social media. Businesses should design with the assumption that it is unclear how fast they will be adopting AI in the next 12, 24, or 36 months and build the system in such a way that they will be able to surround everything with AI without suffering in performance. The idea is to have a platform that is like a sandbox and that allows businesses to play with all the different AI approaches.

As part of this forward-looking approach, businesses should also assume that the demand that AI will impose on their infrastructure is going to increase distinctly over the next two to three years. The AI models will become more complex, the data used for training more voluminous, the applications leveraging AI more numerous, and the inferencing load, by most accounts, potentially astronomical. To avoid the situation that the earliest adopters have found themselves in — switching infrastructure three times, in some instances over the course of two or three years — they should use an infrastructure solution that can handle these extremely data-intensive AI workloads not just today but in the next three years.

CHALLENGES/OPPORTUNITIES

For Businesses

For today's businesses the first challenge is to bring AI into the organization in a planned, streamlined fashion. A few years ago, haphazard experimentation with AI was the order of the day and generally accepted as necessary, albeit at a price. Today, that scenario is avoidable. There are various solutions in the market to implement a coherent, efficient, and functioning AI road map for the enterprise. As such, AI is shifting from being a challenge to becoming an opportunity for the business. There are still hurdles, of course, not in the least with attracting professional AI data scientists, but in terms of preparing the data and building the right infrastructure, no business needs to reinvent the wheel anymore.

Another challenge for businesses is the fact that we are moving toward heterogeneous datacenters. After years of running their workloads on standard architecture servers, businesses need to not only comprehend how to integrate, program, and scale accelerators such as GPUs, FPGAs, ASICs, and many-core processors but also start getting used to different architecture host processors that do

certain things better and more efficiently. There's no reason for alarm, though. Thanks to Linux, virtualization, and containerization, a different processor is hardly noticeable, except in the way it performs better on a certain workload.

For IBM

For IBM, the greatest challenge remains closely related to the second challenge for businesses mentioned previously. Overcoming the inertia that businesses suffer from with regard to accepting that some workloads objectively run better on systems with different processors, such as Power Systems, continues to be IBM Power Systems' greatest hurdle toward breakthrough market success. The notion of "standard" architecture is so deeply embedded in the IT culture that most people forget that "standard" is born from common denominator design. In other words, if it's standard, it does everything well and nothing exceptionally well.

AI is therefore a once-in-a-generation opportunity for IBM Power Systems. It is the ideal workload for a Power System, and it is quickly evolving into a huge market opportunity. That opportunity is IBM's to lose, and the most important message for IBM to succeed is: keep it simple, no matter how complex the underlying designs; keep it affordable; and market it with a splash toward developers, who are the prime influencers in the datacenter.

Finally, one opportunity that IBM should not overlook is the brand recognition it has with Watson. Connecting AI Infrastructure with Watson, actually or conceptually, is not such a stretch and could prove advantageous for IBM Power Systems.

CONCLUSION

Businesses that are still experimenting with AI infrastructure as they try to bring AI to their operations, products, and services can alleviate that burden. Unlike just 12 months ago, there is a growing understanding today of what type of infrastructure deep learning as well as AI inferencing require. Vendors are putting together packages of hardware, software, and services for AI that lower the hurdles toward AI applications. IBM, specifically, has leveraged its Power Systems processor and I/O advantages as well as the custom development it did to integrate the new POWER9 processor with NVIDIA V100 GPUs via NVLink 2.0 to provide an extremely performant server that is the foundation for its PowerAI system. The company also, correctly, saw that businesses are struggling with upgrading their data lakes to allow for fast and comprehensive preparation of data of any kind. This is a critical capability that is often overlooked. IDC believes that AI requires different infrastructure than what has been the standard in the datacenter for many years, and it is increasingly apparent that businesses are aware of this. They are embracing different host processors, various acceleration technologies, and large memory footprints as the way forward.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com