Solution Brief

5G, Edge, and IoT
Intel® FPGAs and Intel® Xeon®-D CPUs

Real-time Cloud Application Execution with Remote Data

Vcinity enables data and application-agnostic performance without using compression, de-duplication, or any other pre-processing of data.

> "Organizations that run data-intensive applications can only provide competitive advantage when end-users can access relevant, up-to-date data as quickly as possible. Any [delays] can lead to untimely insights and decisions that incur unnecessary costs or lost revenue. ESG has not seen any other technology that enables data access, regardless of where data is stored, as quickly as the Vcinity solution."
>
> - Enterprise Strategy Group

Humankind has experienced many technological transformations, from the stone, bronze, and iron ages to the industrial age, and from the pre-mechanical, mechanical, and electromechanical ages to the electronic age (as epitomized by computers and communication systems). Most recently, the world entered the data age, where the amount of data that is being created, captured, copied, and consumed is increasing at an exponential rate.

A zettabyte is 10^21 bytes. In the metric system, one zettabyte is 1,000 exabytes, where each exabyte is 1,000 petabytes and each petabyte is 1,000 terabytes. According to Statista[1], there were two zettabytes of data worldwide in 2010. This increased to over 64 zettabytes in 2020, and it is expected to rise to more than 180 zettabytes by 2025.

Unfortunately, where this data is generated is rarely where it's needed. Traditionally, data is moved or copied so that it's co-located with the applications and compute used to process it, but this takes time and costs money. Today, Vcinity has developed a solution that allows fast and secure access to data, no matter its location.

Problems with the Traditional Computational Paradigm

Data is created everywhere, primarily at the edge. In many cases—oil and gas exploration, medical imaging, artificial intelligence (AI), machine learning (ML), and high-performance computing (HPC), to provide just a few examples—data sets can be terabytes or more in size. The problem is that the software applications and compute engines required to do something with this data (e.g., in the cloud or in a central data center) are not usually located where the data is created (e.g., offshore or in a manufacturing facility).

In order to function efficiently, today's predominant computing paradigm demands that applications, compute, and data be resident in the same physical location on the same local area network (LAN). However, in many cases—such as media and entertainment, military, and scientific applications—the owners of the data (e.g., enterprise or commercial organizations) are reluctant to let the prime copy of that data leave their facilities, thereby forcing external users (e.g., editors or analysts) to access the data remotely.

If these users attempt to run applications on data that is located remotely—possibly on the other side of the world—via a wide area network (WAN) and/or a radio access network (RAN), they can end up spending hours watching a "busy icon" spinning endlessly on the screen while the data is being accessed and retrieved. As a result, and assuming the data's owners permit it to leave their facilities, these companies, enterprises, and government entities are obliged to spend vast amounts of time and resources moving and migrating enormous quantities of data between disparate locations (Figure 1).

Authors

Andrew Parsons
Chief Strategy Officer, Vcinity

Darren van Wageningen
Business Development Manager, Intel Corporation

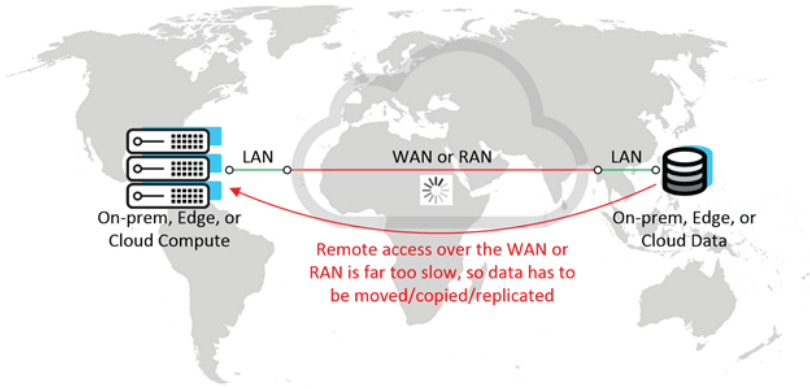

Figure 1. The traditional computational paradigm requires compute and data to be co-located.

Figure 1 reflects an extremely simplistic version of this scenario. In many cases, for example, an application may require working simultaneously on several datasets from diverse sources. Furthermore, if multiple people need to work on the same data using their own compute power, and these people are based in disparate locations, then—traditionally—they will each need their own copy of that data.

In addition to the costs associated with transporting the data and the increased storage footprint, the more copies of the data there are, the higher the risk is that one of them is going to be compromised. There's also a risk that one or more users will end up working with stale, out-of-date data.

One approach to alleviate this problem is to optimize the WAN/RAN by compressing and de-duplicating the data such that it can be copied or moved faster—but data transfer speeds are still slow and inefficient. Another technique is to use an enhanced caching strategy that monitors data usage patterns and moves whatever the system thinks will be required ahead of time. But this approach can fail if the data changes too quickly, the data set is too large to get where it's needed in time, or the system predicts incorrectly. Also, it does not solve the core issue of moving and storing multiple copies of the data.

Unless there is a paradigm change, it has been estimated that tens of thousands of hours and billions of dollars will be wasted on an annual basis moving data to the applications and compute required to process it.

A New Computational Paradigm

Vcinity has developed a revolutionary technology that introduces a new computational paradigm and completely changes the existing use model. This technology conceptually transforms existing WANs and RANs into Global LANs (GLANs), enabling instant, secure access to data located anywhere in the world (Figure 2).

Vcinity leverages key Intel architectures and accelerators based on Intel® Stratix®10 FPGA or Intel® Xeon®-D processor products, Vcinity enables the real-time, elastic, distributed enterprise, putting the power of data in users' hands when and where they need it. The ability to keep the data in-situ without downloading, uploading, copying, pre-caching, or pre-staging means that applications and compute no longer need to be collocated with the data with which they are working, thereby saving vast amounts of time and money.

Vcinity's Application Access Gateway Solution

As illustrated in Figure 2, one way to visualize the Vcinity Application Access Gateway solution is as "bookends" at either end of the WAN or RAN. As discussed later in this Solutions Brief (see: Deployment Options), this bookend technology can be implemented in software or hardware. A key consideration from a deployment perspective is that Vcinity's Application Access Gateway does not require any software or hardware to be installed on client machines. All that is required is a single software or hardware installation as part of the IT architecture, after which all of the users on that LAN can access this technology.

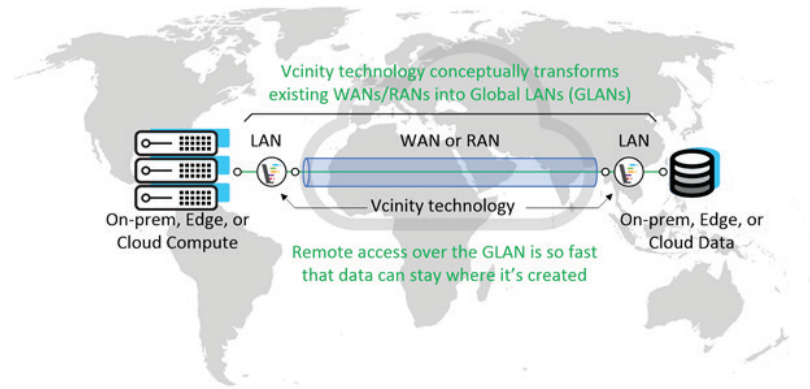

Figure 2. The Vcinity computational paradigm provides a local data experience regardless of the data's location.

A simple integration with existing storage and compute via Network File System (NFS), Amazon Simple Storage Service (Amazon S3), or Server Message Block (SMB) makes it easy for users and applications to access this technology. From the end user perspective, Vcinity's Application Access Gateway provides scalable, globally distributed file access across a managed, hybrid cloud enterprise. In the same way users and their applications see their local systems as hierarchies of folders and files, they can similarly view any Vcinity-connected remote data repositories.

At its heart, Vcinity's Application Access Gateway employs a remote direct memory access (RDMA) protocol—as seen in HPC deployments like Infiniband and RDMA over Converged Ethernet (RoCE)—that bypasses the common slowdowns, like bandwidth constraints, associated with traditional networking protocols like Transmission Control Protocol (TCP) or User Datagram Protocol (UDP). This solution takes the RDMA protocol and extends it to a global scale. In addition to features like intelligent flow control, buffer crediting technology, and packet loss recovery technology, this solution provides sustained performance at 90 percent LAN speed or more when running on Intel's Xeon-D class CPUs or Intel® FPGAs.

In the same way an application accessing local data requests only those portions of the data in which it is interested, Vcinity's solution supports similar functionality over distance. Putting this another way, this solution intelligently moves only the data the application is requesting and moves it with near line rate speed and low latency.

With respect to latency, note the very first bit of data moved from the remote data repository into the application memory space on the compute platform will always incur the physical latency associated with the link. If the link latency is 100 milliseconds (ms), for example, then the application will have to wait 100 ms until it sees the first bit. With Vcinity's solution, however, everything following that first data bit transmits at 90 percent LAN efficiency (Figure 3).

From the IT perspective, Vcinity's Application Access Gateway facilitates administration via standard graphical user interfaces (GUIs), command line interfaces (CLIs), and application programming interfaces (APIs) like REST and SMP. Vcinity's Application Access Gateway also incorporates a globally distributed, high-speed parallel file system that intelligently manages user access, like locks and permissions.

Vcinity's Application Access Gateway Solution at Work

An oil and gas company collects seismic survey activity on ships, for which the data sets captured during these voyages range from multiple petabytes to exabytes. Once captured, the organization's data scientists need access to these data sets to determine where to drill for natural resources. The company's analyst, who is based in Houston, Texas, needs to use a terabyte data file residing in Rio de Janeiro, Brazil for a 3D model. Historically, when an analyst attempted to remotely access the data set using the traditional method, network delays made the process untenable. As a result, the company would ship the data by boat, which introduced a week or more day delay before the analysis could begin—all while the Rio-based crew remained idle awaiting direction from Houston.

To solve these production delays, the company evaluated using Vcinity for two key purposes: #1 accessing remote data and #2 copying and moving data.

Example #1: Accessing Data Remotely

To assess the impact of Vcinity's Application Access Gateway solution to improve remote data access, the company set up two AWS clients located in the US West Region. These clients employed the exact same WAN at the same time to access the same data repository located 2,800 miles away on the US East coast (Figure 4).

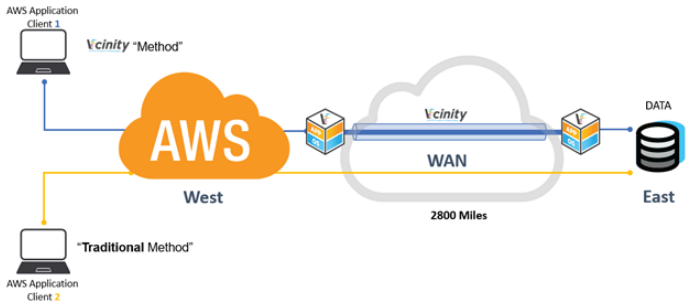

Figure 4. Demonstration setup to compare traditional and Vcinity methods.

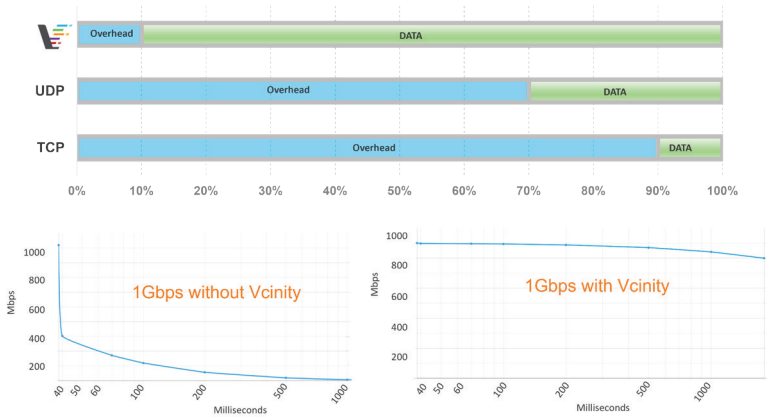

Figure 3. Comparing traditional WAN protocols with Vcinity's GLAN protocol.

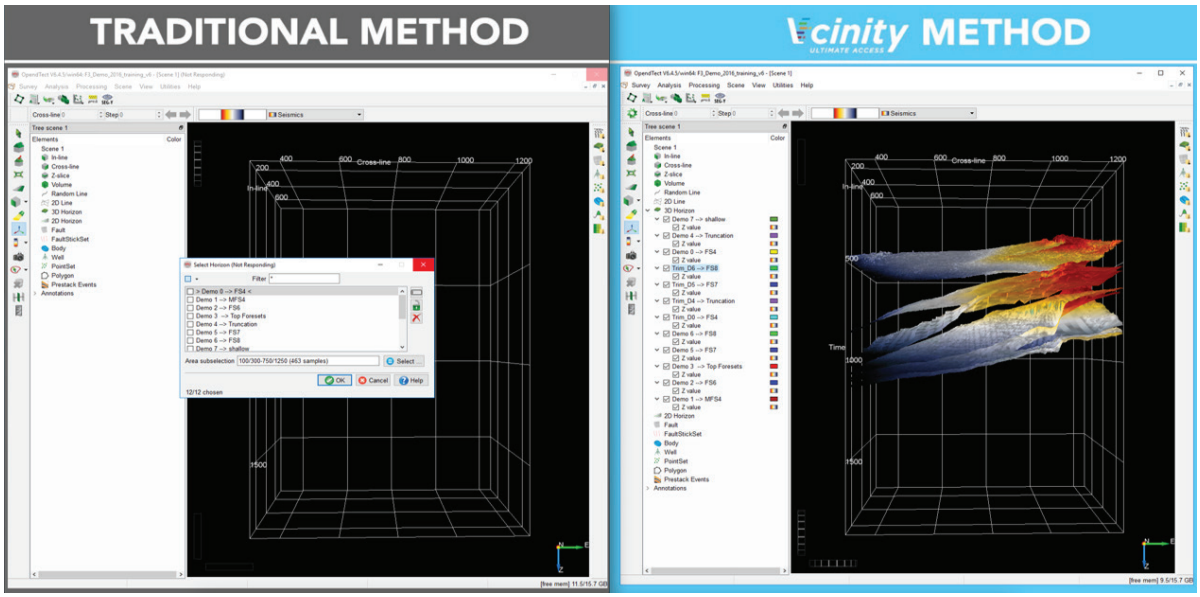

Figure 5. Comparison of remote data access using the traditional method (left) and the Vcinity solution (right).

One of these clients used the "Traditional Method" (remote access via WAN) to access the data repository, while the other used Vcinity solution. In this case, the Vcinity "bookends" were software implementations.

To simulate this real-world experience, the first test involved running the same 3D modelling application on a terabyte of oil and gas data. Even though it was launched second, the application on the client using the Vcinity solution loaded within seconds, allowing the data to be viewed and visualized. By that time, the client using the traditional method (WAN) was still waiting for the application data to start loading (Figure 5).

The organization took the evaluation even further by comparing Vcinity's remote data access capabilities to in-situ compute. Once the data file was delivered via boat to Houston, the analyst accessed it through a 10 Gpbs LAN and recorded the time required to open all the model's layers. The analyst then performed the same action but accessed the original file in Rio de Janeiro over a 1 Gbps WAN connection—and was able to open the full model in 0.25 seconds. Now, using Vcinity, the company can monetize its data in minutes instead of days

Example #2: Copying or Moving Data

Most of this Solution Brief has focused on the value of being able to access remote data in-situ, which offers significant benefit with respect to eliminating the need to copy or move data.

It has to be acknowledged, however, that there are occasions when it is required to move or copy data, such as for backup purposes. Once again, the Vcinity solution offers overwhelming advantages. Using the same demonstration setup illustrated in Figure 4, a raw data file was copied from the data repository on the East Coast to two clients in the West Region. This data file, which was 0.5 terabytes in size, was composed of random numbers, which made it unsuitable for compression. For both of the clients—one using the traditional method and the other using the Vcinity method—the file was copied using a drag-and-drop technique.

Once again, the transfer using the traditional method was launched first, followed by the transfer using the Vcinity method. The traditional method achieved a transfer rate of only around 6 to 7 MB/s, while the Vcinity solution sustained a transfer rate of more than 130 MB/s (Figure 6).

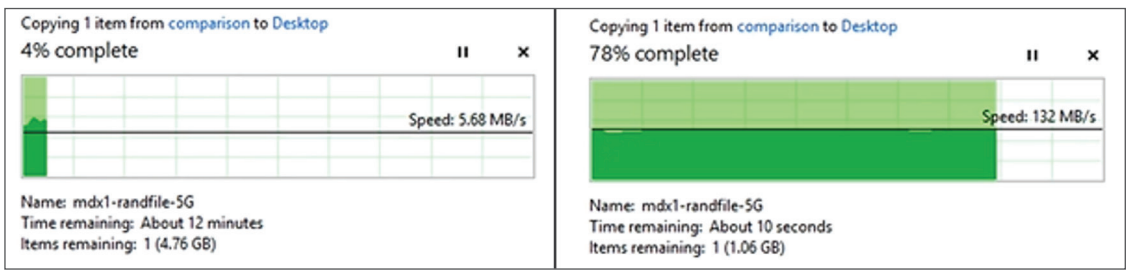

Figure 6. Comparison of copying data from a remote repository using the traditional method (left) and the Vcinity solution (right).

Deployment Options

As was discussed earlier, Vcinity's Application Access Gateway solution can be deployed via either software or hardware implementations. In conjunction with Intel, Vcinity offers a suite of software and hardware options. In the case of software, Vcinity can be deployed as a virtual machine (VM) running on an X86-based Intel® Xeon® processor platform, for example. Such a software solution can support sustained bandwidths from 2 to 5 Gbps depending on the hardware backend.

Users requiring 10 Gbps or higher will require some form of hardware-assist, such as an FPGA-based accelerator presented in a PCI express card form factor. In the case of hardware-assist, Vcinity can be deployed on the Intel Stratix 10 FPGA-based Silicom SmartNIC N501x in an Intel Xeon-D-based server.

Meanwhile, users requiring multiples of 10 Gbps, such as 100 Gbps or higher, may take advantage of a 1U Intel FPGA-based appliance that can be mounted in the server rack.

Getting Started

Vcinity's Application Access Gateway solution runs on a continually growing and broad portfolio of Intel products, from Intel Xeon-D class CPUs to FPGA-based accelerator cards. Integrating seamlessly with hybrid or multi-cloud environments, Vcinity allows applications, whether in a public or private cloud, or on prem, to instantly access data, no matter where it resides—accelerating time to insights, increasing productivity, reducing data transfer and storage costs, and improving business outcomes.

Ready to tap into the power of data when and where you need it? Contact Vcinity at: info@vcinity.io

To learn more about Intel Xeon-D, visit: intel.com/xeond

To learn more about Intel FPGA Programmable Acceleration Cards, visit: www.intel.com/content/www/us/en/products/details/fpga/platforms/pac.html.

To learn more about the Silicom FPGA SmartNIC N5010, visit: www.silicom-usa.com/pr/fpga-based-cards/fpga-intel-based/fpga-intel-stratix-based/silicom-fpga-smartnic-n5010_series/

About Vcinity, Inc.

Vcinity increases the agility and velocity of digital transformation for enterprises by enabling applications to instantly access and operate on data regardless of where it exists. Vcinity solutions eliminate the necessity to move, copy, or cache data near compute. Applications process data at the point and time of creation with LAN performance. To find out more about Vcinity, visit vcinity.io or check us out on Twitter and LinkedIn.

References

1 Statista: Amount of data created, consumed, and stored worldwide 2010-2025: www.statista.com/statistics/871513/worldwide-data-created/

SS-1144-1.0