SOLUTION BRIEF

Edge Computing

Lenovo Modular Edge Systems Target Enterprise, Telco Edge

Intel® architecture-based Lenovo ThinkSystem SE350 and ThinkSystem SE650 offer performance and flexibility for 5G, smart city, autonomous vehicles, and other emerging applications

Lenovo

Growing demand for low-latency access to cloud compute resources is driving the increase in demand for edge computing. This is a new computing paradigm where compute resources are located at various locations on the edge of the network for ultra-low latency response, geofencing for compliance, reduced cost, or improved data security. The compute performance of the edge server is suitable for supporting compute-intensive and latency sensitive applications such as artificial reality/virtual reality (AR/VR). Because of this compute power, edge computing use cases are growing and evolving, while seamlessly converging enterprise IT, operational technology (OT), and communication technology (CT) workloads. Working closely with Intel to architect this converged edge solution with a scalable, modular approach, Lenovo has developed the edge computing system that is optimized for a wide range of use cases, including 5G.

Growth of Edge Computing

Applications that use edge computing include but are not limited to traditional network functions, content delivery networks (CDNs), self-driving cars, security camera analytics, internet of things (IoT) analytics, location aware services, and others. Each application use case brings a need for different compute, storage, and networking requirements—challenging the use of a monolithic, one-size-fits-all system.

There is no single definition of the network edge, but rather there are several edge definitions—each with multiple applications that reside within them. There is a breadth of edge use cases/workloads that are spread across different locations and that share the same key performance indicators (KPIs) or value proposition:

- IoT sensors and other devices that are constrained for power, small size, environmental factors, and cost. Examples are IoT gateways or road side units (RSU), and applications include video capture, first-level video analytics, vehicle to vehicle (V2V), vehicle to everything (V2X), smart city, or gaming.
- Smart industry applications in enterprises that are the most remote edge location in the communication service provider's (CoSP) network. These applications have requirements for power, environmental factors, and physical security that drive the use of 1U/2U servers. Examples are cabinets or base stations, and applications include mobile network virtual radio access network (vRAN), private LTE, IoT, V2V or V2X, hosted networking services for retailers or other small businesses, and other low-latency workloads.
- Aggregating data from multiple IoT edge and enterprise edge systems to host more traditional servers in higher performance configurations. Applications include 5G base stations, CDNs, video analytics, virtual radio access network (vRAN) processing, and others.

Just as a monolithic edge server cannot serve the needs of the market, individual edge architectures for each application category do not serve the needs of CoSPs either (see Figure 1).

The management and maintenance costs increase total cost of ownership (TCO), while the potential interoperability issues could limit services or service features.

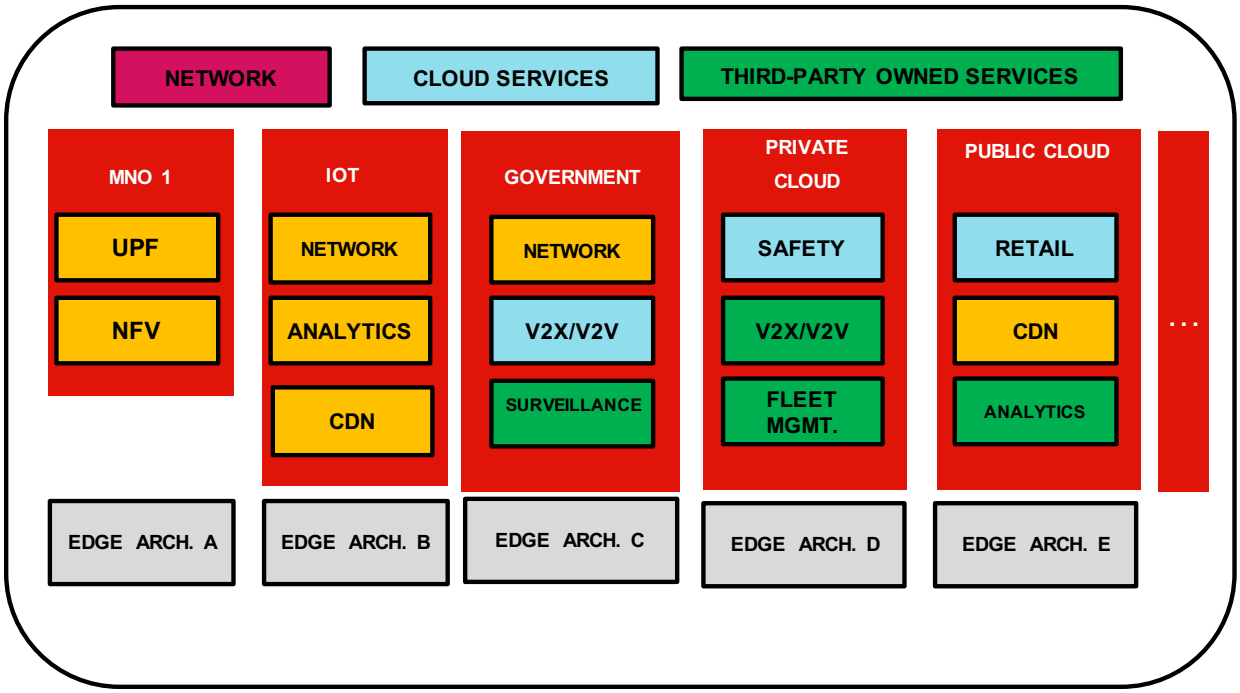

Figure 1. Multi-architecture approach to edge computing increases costs, and brings potential interoperability issues[1]

To address this evolving market, Lenovo has worked closely with Intel to develop several edge systems that embrace a converged edge architectural approach. Lenovo's modular edge servers are designed to be future ready so that once deployed they can support new use cases.

The Lenovo ThinkSystem SE350 is generally available; Lenovo ThinkSystem SE650 prototypes are currently in live trials with customers. These edge servers are inherently multi-tenant via a virtualization layer and CPU features that allocate compute resources per application and to different tenants. This enables multiple, independent services to run simultaneously. In addition, Lenovo sought to ensure the systems were manageable using automated infrastructure orchestration features that provide simplified, remote service deployment.

Modular at the Telco Edge: Lenovo ThinkSystem SE650

The Lenovo ThinkSystem SE650, which is a prototype, is a 2RU-high modular system powered by dual Intel® Xeon® Scalable processors designed for applications that need significant processing power such as vRAN, multi-access edge computing (MEC), and NFV infrastructure (NFVI). Physical and data security are important elements of the system, which features protections to help prevent unauthorized data access, a dedicated management port, and the ability to detect hardware tampering and unauthorized movement.

The two-socket server offers balanced I/O by using non-uniform memory access (NUMA) to pin a CPU to its corresponding memory resources and its accelerator. The system features optional slots for accelerators, offering more processing power for applications such as radio access network (RAN) or analytics. Other hardware features include:

- ETSI 600mm rack deployment
- Ruggedized for carrier class deployment (NEBS 3, ETSI)
- Remote management via LTE
- Six I/O slots
- DC and AC power
- Front access for service, power, and I/O

Lenovo MEC Software Stack

The Lenovo Research group has developed a MEC software stack for the ThinkSystem SE650 prototype to deliver a multi-tenant run-time and hosting environment for applications. For performance-optimized networking, customers can configure virtual networking functionality using Open vSwitch with Data Plane Development Kit (DPDK) for very high performance.

For applications that call on an accelerator, the stack features an accelerator abstraction layer (AAL) to manage hardware abstraction and make this processing capacity available to applications via an application programming interface.

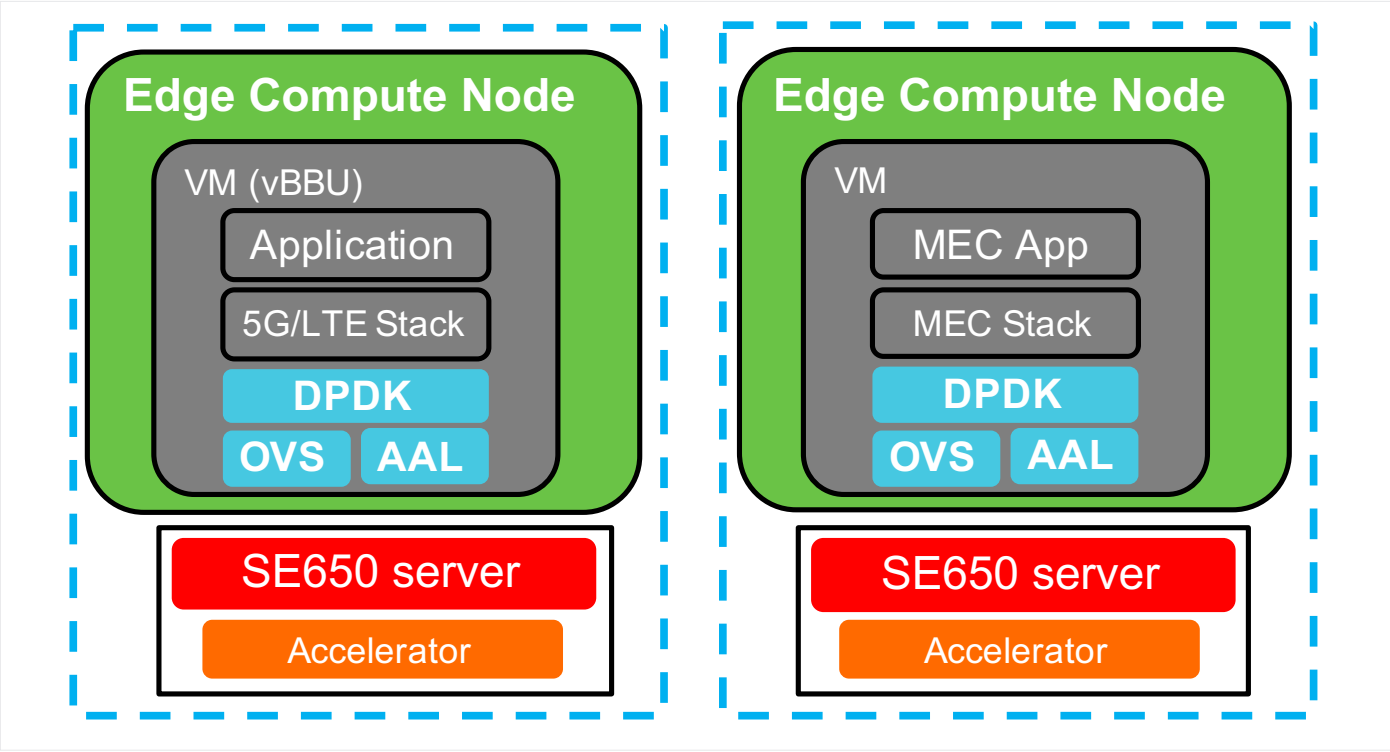

Figure 2. Block diagram to the left is an example stack for a 5G application and on the right is the MEC software stack

On top of this layer is a customized MEC stack designed for workload convergence that could include either open source Akraino, Open Network Edge Services Software (OpenNESS) toolkit, or third-party MEC software. A MEC application enabler can also be utilized: for example, the Intel® Distribution of OpenVINO™ toolkit for video analytics processing or software for augmented reality/virtual reality (AR/VR). This infrastructure is in place for an application such as a 5G open RAN small cell, which is shown in Figure 2.

Modular at the Enterprise Edge: Lenovo ThinkSystem SE350

The Lenovo ThinkSystem SE350 is a compact, powerful, edge server designed for rural edge, smart city, enterprise edge, or IoT edge applications. It features Intel Xeon D processors to deliver density-optimized processing power for applications including autonomous vehicles (V2V/V2X), video analytics, CDNs, and a variety of IoT/smart city applications.

The ThinkSystem SE350 is small-footprint server measuring 1.75 inches high, 8.1 inches wide, and 14.9 inches deep that can be mounted on a wall, stacked on a shelf, or installed in a rack. The server is hardened in a shock-resistant and dust-resistant case and can handle temperatures from zero degrees Celsius to 55 Celsius.

This server uses the Intel Xeon D-2100 processors, which feature a data center processor architecture that is optimized for network, storage, and cloud-edge systems. Intel Xeon D-2100 processors have up to 16 cores with a performance-per-watt of power that is optimized for enterprise edge applications.

The ThinkSystem SE350 also features 256 GB of RAM and 16 TB of internal solid-state storage with optional support for additional accelerators. Multiple connectivity options are offered, including wired Ethernet, Wi-Fi, and 4G/LTE. Because it will be deployed in the remotest of locations, security is a key consideration. The ThinkSystem SE350 includes encrypted storage and physical security features, such as a locking bezel, along with intrusion and tamper-detection mechanisms.

Figure 3. Lenovo ThinkSystem SE350

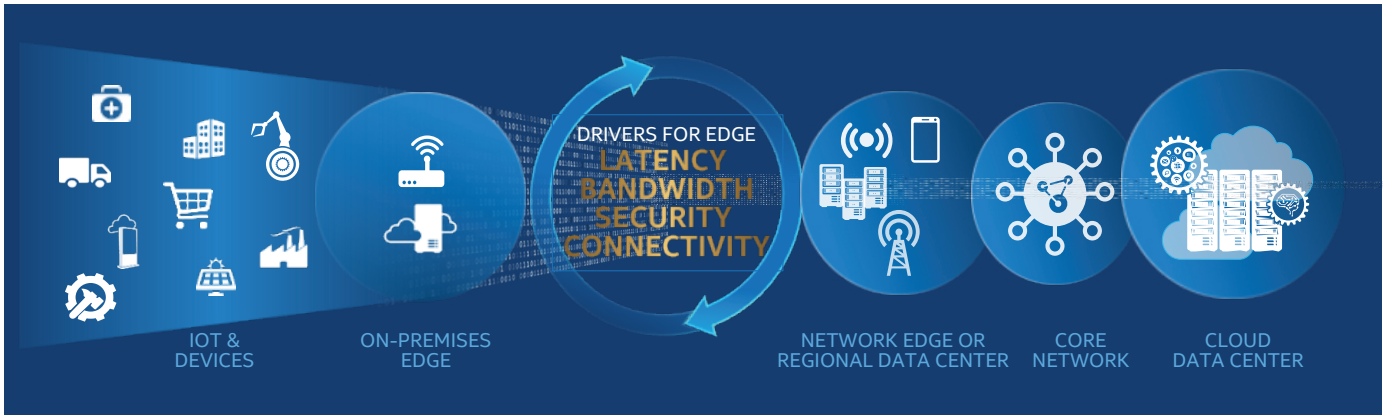

Figure 4. Edge network ecosystem

Converged Edge Architectural Approach

Utilizing a converged edge architectural approach enables deployment of edge solutions in a flexible, modular manner to various edge locations—on premises or as part of the network edge (see Figure 4). By being able to add platform infrastructure in a modular manner, specific ingredients are brought together to achieve convergence of IT, OT, and CT workloads. The approach also results in driving a commonality across infrastructures that might be located in various edge locations, thereby helping service providers to leverage common portfolio investments as well as orchestration systems.

Companies like Lenovo can use elements of the architecture to develop unique edge servers. Key considerations for converged edge architecture include the following:

- **Edge server location:** A converged edge architecture recognizes that different edge locations have different requirements based on availability of power, need for cooling, indoor/outdoor cabinet, and the different applications that might be served from that location. The architecture is defined modularly so that OEMs can optimize their systems for a variety of different locations.
- **Power efficiency:** Energy utilization is a key tenant of edge network design because it has the highest cost, over time, of all aspects of the edge network solution. Energy consumption needs to be considered in all network architecture design decisions, including the backup compute nodes that are in place for redundancy and failover protection.
- **Technology mapping:** A converged edge architecture specifies a network platform, microservices through a toolkit like OpenNESS, ecosystem components, and orchestration layers that can be configured to provide the best architecture definition for a location and for an application.
- **MANO:** The designs factor in how edge servers are to be managed by management and orchestration (MANO) stacks.
- **Interfaces:** A converged edge architectural approach considers all of the open interfaces that could be exposed to the applications to have full access to all the features provided by the hardware. This analysis is fundamental for maximizing value at the edge through operator manageability and accessibility/usability by the end services.

Lenovo Edge Solutions Are Optimized for 5G Use Cases

5G is expected to be a big driver of the move to edge computing. Consider 5G vRAN, an access edge deployment that requires a larger number of small cell base stations to offer the same level of connectivity as 4G networks. In addition, the low latency and high throughput of 5G is an enabler for new applications such as V2X that utilize edge computing. Even if these applications do not need low latency, they are expected to create a significant amount of data, which will require costly expansion of backbone transport networks if not processed on the edge.

Figure 5 shows a sample 5G implementation that demonstrates Lenovo's modular system. The network was set up across three cities and offered V2V, V2X, video analytics, CDN and safety applications via the ThinkSystem SE350 servers located in enterprise edge points of presence (PoPs). The modularity of the systems allowed data planes to be customized for each site as well as the option to offer different compute accelerators to match the needs of each site, although in this instance, each of these locations utilized Intel® Movidius™ Myriad™ X Vision Processing Unit (VPU) to provide a dedicated neural compute engine for vision processing and other application acceleration.

The ThinkSystem SE650 prototype was used in a telco edge site that allowed it to connect the other locations and utilized its processing performance to provide additional compute power for the far edge servers and also to provision additional applications such as AR/VR.

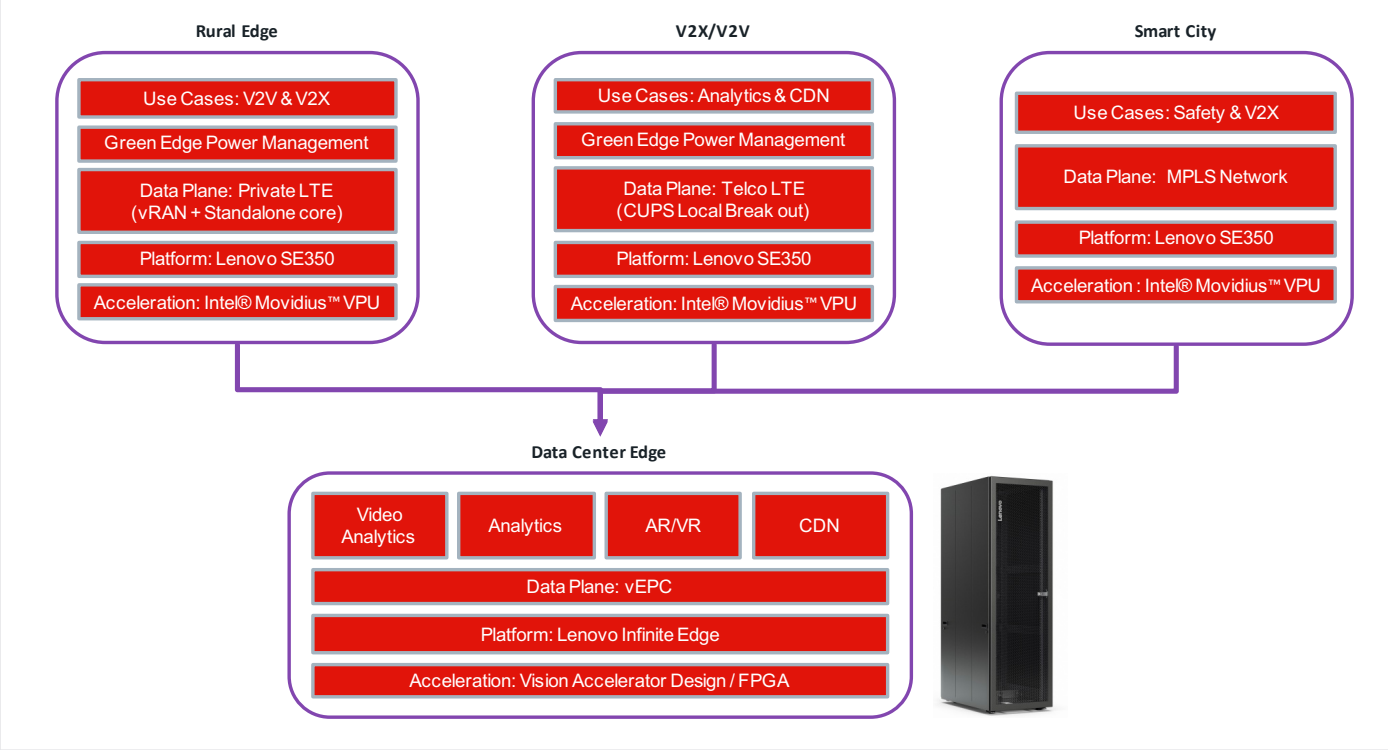

Figure 5. Example of Lenovo modular edge servers in 5G application

Conclusion

Converged solutions are accelerating service innovation at the edge and proliferating new use cases. By taking a converged edge architecture approach and utilizing its Intel-based ThinkSystem SE350 and ThinkSystem SE650, Lenovo has a modular edge computing solution that is optimized for 5G and can be utilized for a wide range of service deployments at different edge locations.

Learn More

Lenovo: https://www.lenovo.com

Lenovo ThinkSystem SE350: https://www.lenovo.com/us/en/data-center/servers/edge/ThinkSystem-SE350/p/77XX6DSSE35

Lenovo is a member of Intel® Network Builders ecosystem: http://networkbuilders.intel.com

Notices & Disclaimers

[1] Figures 1, 2, 3, and 5 provided courtesy of Lenovo

Intel technologies may require enabled hardware, software or service activation.
No product or component can be absolutely secure.
Your costs and results may vary.
Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

0720/DO/H09/PDF ♻ Please Recycle 343955-001US