AEROSPIKE

Multi-access Edge Computing (MEC)

USE CASES

Multi-access Edge Computing (MEC) market explodes

According to new estimates from IDC, worldwide edge spending is expected to hit almost $274 billion in 2025, growing at an 18.7% compound annual growth rate (CAGR) during that period. That growth is being driven by AI and analytics apps that have choked off some data passages to and from the cloud, IDC notes.

On an even more vigorous note, Grand View Research reports that the CAGR for the global multi-access edge computing (MEC) market size is expected to be more than 42% from 2021 to 2028. This robust growth is attributed to the increasing use of IoT devices across various industries, and the need for faster decision-making in industries such as IT and telecom, manufacturing, transportation and logistics and datacenters. The research group also reports that the adoption of MEC offers a "safer environment for data transfer and reduces latency."

STL Partners echoes similar optimism around the MEC market – in their estimates, the total addressable market (TAM) for edge was worth only $10 billion in 2020, but is going to explode to $534 billion in 2030.

MEC is primarily about low latency

The core concept of MEC is very straightforward: lowering the latency to deliver faster service to the end-users. More often than not, MEC-related discussions circle around low latency of the network (by bringing network gears closer to the end-users) and low latency of computation (by introducing accelerators such as GPU, FPGA, or special ASICs).

Still, there is a third dimension that usually doesn't get as much attention: lowering the latency of data processing at the edge.

End-users really don't care about the source of latency. Their main concern is how much time it takes for a service to be delivered to the requesting endpoint i.e. end-to-end latency. It's immaterial to them whether the latency is contributed by access network distance, computation, or by the processing of data.

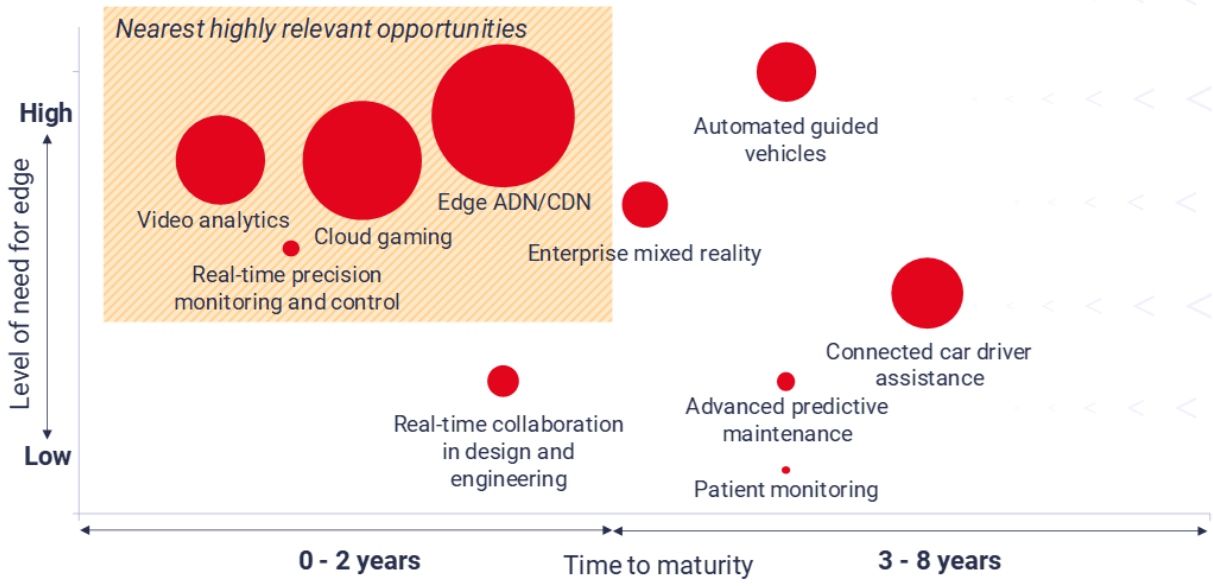
Low and consistent latency is important for the supplier side as well. A key value proposition for Verizon adopting AWS Wavelength is ensuring consistently low latency to the application developers as they serve new use cases for end-users from the edge of the telco network. And, low latency data processing becomes extremely relevant for edge-use cases that are "data-heavy."

That's why we consider MEC a "sweet spot" for Aerospike. We offer not only real-time at scale performance but also the smallest footprint (and subsequently the lowest TCO) at the edge, thanks to our patent-winning Hybrid Memory Architecture (HMA).

MEC use cases

At this point, there is a universal consensus that telco edge/MEC is an opportunity for telcos. They can move up the value chain beyond connectivity and capture revenues from edge platforms and more importantly, applications i.e. use cases.

According to STL Partners' estimate, by 2030, 70% of the addressable revenue for edge should come from the applications with 21% from the edge infrastructure/platform/integration/managed services. Only 9% will go to pure connectivity. Their categorization of MEC use cases is as follows:



Source: STL Partners - Edge computing market sizing forecast

We have a similar view in the sense that in the near term, data-heavy edge use cases such as cloud gaming, video analytics and edge CDNs will dominate the conversation. There also should be an array of connected machine/sensor/device applications under the umbrella of IoT use cases.

Let's delve into those use cases below with an emphasis on what Aerospike's real-time data platform brings to the table:

Multiplayer cloud gaming

As gaming is emerging as the dominant non-streaming killer app for consumer broadband, the need for dynamic, interactive, multiplayer gaming is on an exponential rise. Providing a true real-time experience to the gamers requires processing lots of data. It also must extract insights out of that data and act on that data in a short amount of time. This needs to happen while the gamers are still engaged with whatever end devices they are using. Hence comes MEC to the rescue. As highlighted by the Cox Edge gaming use case, by default the states of all entities related to a particular game need to be stored at edge locations closest to the end-users. This is to enable extreme performance and dynamic gaming engagement in virtual worlds that are very large in size (up to tens of thousands of players/game) and have a huge number of complex AI and gameplay behaviors.

The diagram below summarizes the flow of events for this signature MEC use case and how Aerospike's real-time data platform adds value for everyone involved. For gamers, it's a more engaging and fulfilling experience. For the providers, it's an opportunity to keep the platform stickier with potential revenue upsides through cross-selling and upsell.

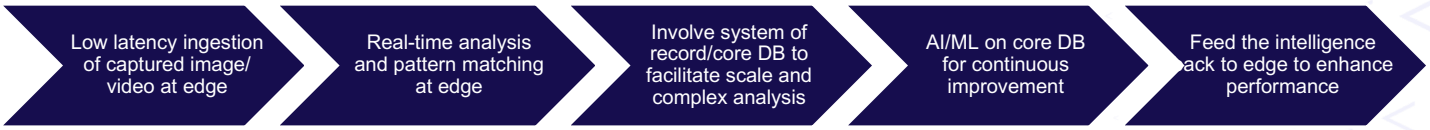
Benefits with Aerospike

- ✓ **Avoids rigidity of a SQL data schema –** Allows a flexible data model where specific parameters of a user are not statically provisioned to a specific server/database instance
- ✓ **Consistency across edge and core –** Cross Datacenter Replication (XDR) helps core DB to get copy of edge data in real time and aggregate from multiple edge systems for further processing. XDR's bi-directional filter allows to choose data attributes to be shared between edge and core.
- ✓ **Integration w/ AI/ML Engines –** Best-in-class Spark connector to get data in and out of the AI/ML engine in real-time. Push the intelligence back to the edge to facilitate personalized experience and cross-sell/upsell offers (by the providers)

Image/video analytics

Image/video analytics is a near-term edge use case that is already happening. The key value proposition here is very simple. First, the huge amount of image/video data that's captured 24/7 by a range of edge devices/processes, do not need to travel in its entirety to the core network. It is much more efficient to process that data near to where it's captured and only forward relevant filtered data to the core, saving a ton of bandwidth cost along the way. Second, a subset of this analytics use case – namely facial/object recognition – maintains its relevance/efficiency only if the decision can be rendered in real-time. Here comes MEC again with its low latency value proposition. Earlier in a blog post, we wrote about a Silicon Valley-based AI company that uses the Aerospike data platform to power their facial recognition solution for a banking industry client. Anyone walking into a branch of that bank undergoes facial recognition at the entrance to verify identity and subscription tier. Our data platform plays a key role not only by storing the metadata of the contour of human faces but also by supplying that data to the AI module in real-time to facilitate a fast and seamless decision. As MEC becomes more mainstream, we expect to see an explosion of this kind of use case.

The diagram below highlights some key aspects of this use case:

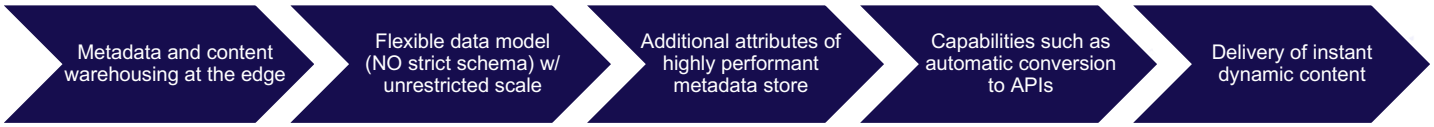


Benefits with Aerospike

- ✓ **Predictable low latency –** performing equally good with both read and write workloads without deteriorating under load. No spike in latency due to garbage collection as seen in Java Virtual Machine (JVM) environment
- ✓ **Supporting scale with small footprint at Edge –** there is no "Server Sprawl" at edge even if the data volume gets massive, thanks to our Hybrid Memory Architecture (HMA) that leverages performance optimized SSDs to store data
- ✓ **Handling of bursty workloads –** allows to maintain stable performance even under peak load during sudden spikes (e.g., sports/concert events, busy train stations)

Content management at the edge

As claimed in the Cox Edge dynamic content delivery use case, by moving data and queries to the edge, complex application functions such as authentication, content personalization, search, customization, and ordering can be done 10X-50X faster than with conventional application backends or content systems. The ability to work with any data model and turn that into an API that can then be integrated with web pages/applications/microservices is a key attribute of dynamic content delivery. When the content delivery happens from a backend centralized database/storage – even with CDN caching on – dynamic data requests need to travel long distances, adding latency in the process. This invariably deteriorates user experience. Another aspect of content management is a high-performance metadata store. The high performance can come from a flexible schema, the ability to support low latency read/write operations at massive scale, high uptime (>99.999%), and capabilities around replication, elastic search, secondary index, and expression filters, etc., all supported by Aerospike.



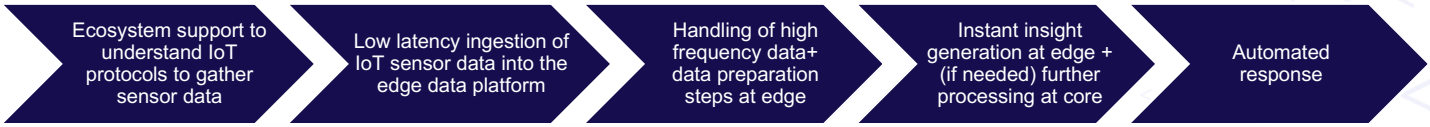
Benefits with Aerospike

- ✓ **Small footprint facilitating edge deployments –** Our small footprint (via denser nodes) alleviates space constraints at edge facilities
- ✓ **High performance datastore and ecosystem support –** Along with supporting low latency, massive scale, and high uptime, we support (by moving to an elastic search) full text search type access patterns via our ESP connector
- ✓ **Avoids stale content –** No backhauling of dynamic data requests over long distances means NO dreaded spinner. Also, our NO caching mechanism means NO cache misses

Internet of Things (IoT) at any scale

IoT use cases generally require time series capabilities. Although Aerospike originated as a key-value store database, it now supports time-series data models through Aerospike Time Series API and a combination of buffered writes and efficient map operations that allows us to optimize for both the reading and writing of time series data.

To process IoT sensor data at the edge, the requirements include cleansing/filtering, formatting, ingesting, combining and transforming massive volumes of data. Then there is the requirement of real-time analysis to make sense of the data gathered. Finally, it needs to facilitate the generation of actionable insights leading into use cases such as precision control or predictive maintenance that benefit by being served from the edge (closer to the sensors/endpoints). The key here is to take advantage of MEC's low latency value proposition and not require the data to be backhauled all the way to the cloud. However, along with processing at the edge, the ability to replicate and gather that data into a core system of record (SoR) can facilitate additional use cases (e.g. complex AI/ML-related tasks) as discussed under the multi-player cloud gaming use case above.



Benefits with Aerospike

- ✓ **Ingests any data format at the Edge –** Not Schema-dependent. Supports time series data models. Works with industry leading data capture and messaging platforms (e.g. Ably) that speak IoT protocols such as MQTT or CoAP
- ✓ **Handles massive scale linearly –** Through our performance-optimized SSDs, parallel architecture, smart partition algorithm and smart client
- ✓ **Cloud-like analytical capability at the Edge –** Allows sensor data analysis and pattern/anomaly detection and insight generation within milliseconds away from IoT sensors/devices. Through XDR can replicate/gather data at core for further analysis

MEC enables many new services that were not possible until now because of the need for guaranteed low latency connectivity and processing of data. The four use cases that we highlighted above are just a snapshot of what sort of newer use cases have been enabled by the emergence of and strong momentum behind MEC. But, there can be many more. If you are exploring/considering ways to deploy MEC use cases in a future-proof yet cost-efficient manner, we would like to hear from you. We can work with you to generate both the architectural blueprint and business case.

The Aerospike Real-time Data Platform enables organizations to act instantly across billions of transactions with predictable performance from terabytes to petabytes of data, while reducing server footprint by up to 80 percent.

2525 E Charleston Road, Mountain View, CA, 94043 | (408) 462-2376 | aerospike.com