TM Research & Development

Cloud Native Containerisation Meets the Edge: A New Era of Smarter Services

White paper

Introduction to the world of eXtended realities

Video traffic will account for 69 per cent of all mobile data traffic by the end of 2025, and is further forecasted to increase to 79 per cent by 2027 [1]. This is a clear indication that the latest trends today in multimedia applications are moving towards eXtended Reality (XR), an umbrella term that encompasses three realities, namely virtual reality (VR), augmented reality (AR) and mixed reality (MR). This technology explains how people's everyday lives can be transformed with the use of groundbreaking scientific knowledge to create captivating interactions, environments and experiences.

Let's look at how each of the realities are used in today's world:

- **Virtual reality** – this application uses a headset to fully enter and be engrossed in a computer-simulated reality. A good headset will generate sounds and images while combining the use of all five senses for a full experience.

- **Augmented reality** – in contrast to immersing users, this reality uses a device, often a camera on your phone to superimpose digital graphics and sounds in a real-world situation.

- **Mixed reality** – this reality is a mix of both of the above, whereby it blends real and virtual worlds that result in an intricate environment with physical and digital elements that can communicate in real life.

Besides XR there are many other trends that are moving fast enough to drive change in different sectors and industries today. These include high-definition video, real-time multimedia interaction and real-time uploading of self-produced content. Simply put all real-and-virtual combined environments and human-machine interactions generated by computer technology and wearables are on the rise [2].

Evaluating XR challenges

Like every other application, there are challenges that are bound to arise, and this includes the XR application.

One of the main challenges is easing the visualisation of digital objects and handling action commands from the XR terminal. This has driven the system integrator to move the processing functions to the cloud, while the video streams are delivered to XR terminals wirelessly. Utilising the powerful processing power of the cloud, the weight and size of the XR terminals were significantly reduced, which subsequently improved user experience.

Reducing the complexity of the XR terminals greatly helps to promote the massive development of affordable XR terminals. However, the cloud XR application scenarios require a large user-experienced data rate of around 100 Mbit/s to ensure a high-definition video experience (>2K), and a network delay of 5~10 ms due Motion-To-Photon (MTP) delay of 20 ms to eliminate user motion sickness effects.

Robotics for edge containerisation

Robotics has also become a vital part of our modern life. It is used in manufacturing to replace repetitive tasks in hazardous places or to collaborate with humans during the production of goods leading to increased productivity. To simplify the hardware of industrial robots and a better implemented artificial intelligence (AI) software control-driven approach, moving control functionality into an edge [3] located for example, in a control room of a factory has revealed several benefits. One of them being robots that become cheaper and occupy less space on the shop floor. There is also a potential to deploy more robots using coordinated methods to improve efficiency in utilising them for specific purposes.

In order for the above scenarios to be realised, it is critical to have functions or services close to the proximity of users that can enable low latency, real-time responses and low energy consumption. In addition to that, these emerging technologies should also help to avoid high bandwidth demands on communication networks that go back and forth in data centres. Protecting the privacy of users and sensitive application data is also considered an important capability for these emerging technologies. We have termed these emerging technologies as edge containerization, a powerful combination of edge computing and containerized software applications.

What is edge computing?

Edge computing is a distributed computing paradigm bringing computing and storage resources closer to the user. This takes place when virtualisation technology over the edge is adopted. As illustrated in Figure 1.0, virtual machine spawning using a hypervisor is an efficient virtualisation method for resource utilisation that has been around for decades. Containerisation is indeed a step ahead of the widely adopted virtual machine where an application's software files, including code, dependencies, and configuration files, are bundled into a package and executed on a host by a container runtime engine. Although the concept of containerization is not new and has been introduced or existed since Unix's chroot system in 1979, it gained popularity due to the introduction of Docker [4] and was followed by cloud-native initiatives [5]. Containers are now used in all areas of software and are critical in many projects that require integration pipelines.

To manage the workloads or containerised applications, a container orchestration platform is required and thus, Kubernetes [6], an open-source container orchestration platform, has become the de-facto industry standard. The container orchestration platform consists of master nodes and worker nodes. Master nodes act as control planes of the container platform to assign computing, storage and networking resources to the applications running over the worker node. As illustrated in Figure 2.0, the worker nodes can reside in the cloud, like in data centres, or sit at the edge, for instance in the control room of a factory or theme park that offers an immersive gaming experience. If the worker node resides at the edge, the user's experience is better in terms of offered services since the latency is lower. In our example, the machine learning (ML) application can perform inference over the edge by having multiple cameras connected to it. In addition to that benefit, users may also receive an immediate notification should an intruder go into the premises.

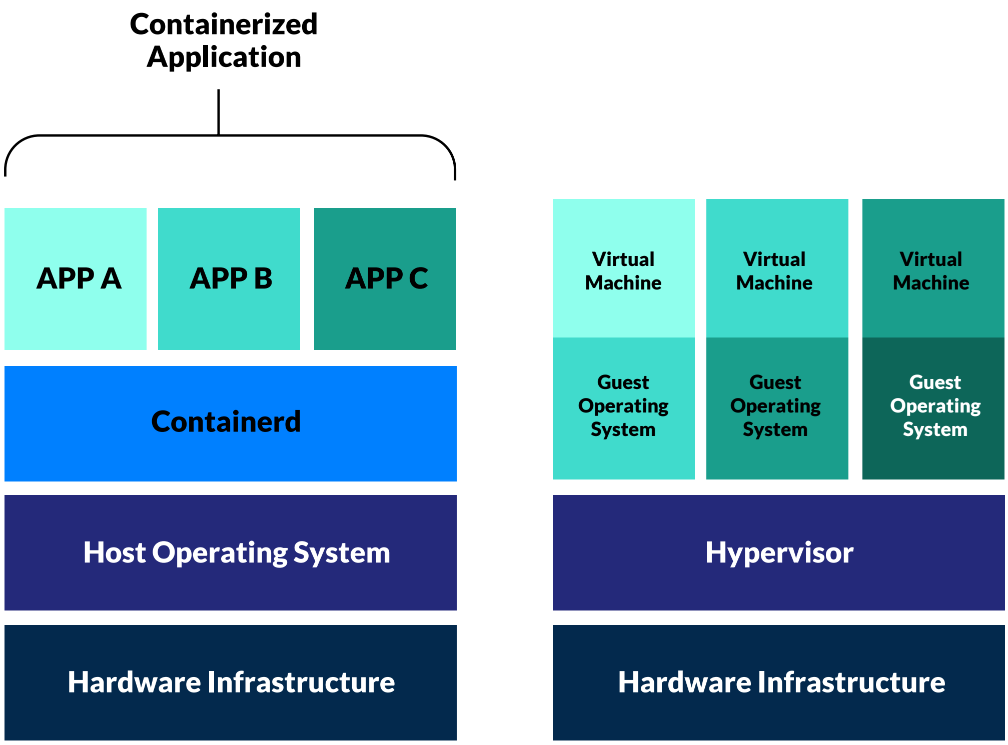

Figure 1: Current State-of-the-Art Visualization

Furthermore, the scaling of containerised machine learning applications is made easier by using a control plane that has full visibility of the overall cluster resources. Based on the demands at the edge and the service level agreement (SLA), the master node can decide whether to assign the desired workload to the worker node located at the edge, or at the cloud.

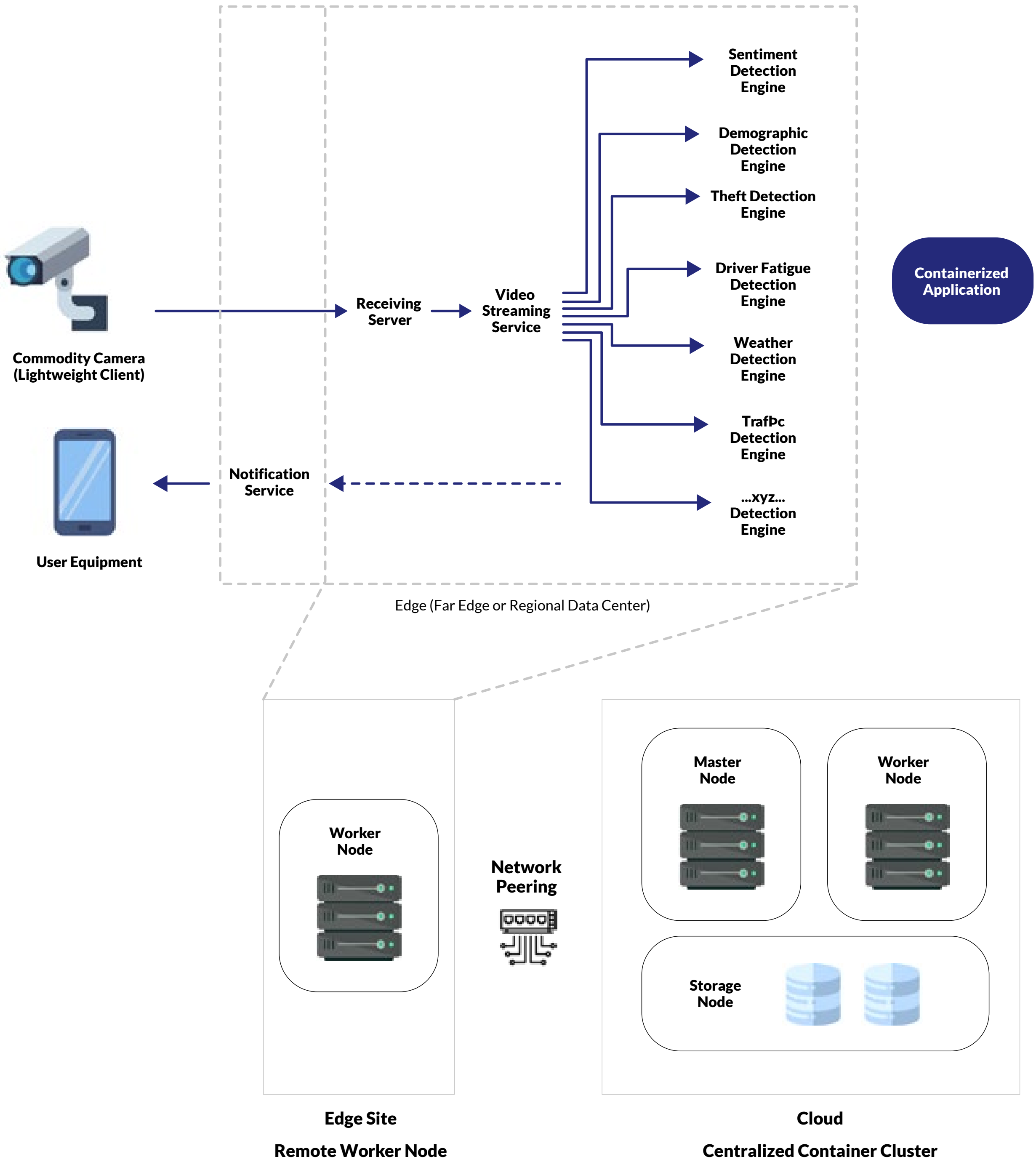

Figure 2.0: Container Platform Deployment with Edge Application Workload

Benefits of edge containerisation

Reliability

The key aspect of its reliability is it avoids fault scenarios. Containers are isolated, meaning all the application dependencies are packaged within the container and cannot conflict with other software in other containers. It is also considered reliable because it can recover and continue operations under fault conditions.

Scalability

Different software systems must be scalable to be able to support connecting and coordinating high numbers of computing nodes. Being scalable also allows container-based systems to adapt as mission priorities shift.

Maturity

Container technologies like Docker are considered mature and trail-tested. In addition, no retraining is required, so developers or researchers using these tools can continue using the same ones.

Reduced bandwidth

centralised applications have the tendency to have high network charges because all the traffic is concentrated to the cloud data centre. Because of the edge container's closeness to the user, it can provide pre-processing and caching.

Conclusion

Edge containerisation brings many benefits to future 5G usage scenarios and is always an important capability to realise 5G network slicing. However, the challenges and areas of improvement remain to be tackled. Some of these challenges include container build and deployment processes with built-in container minimisation. Another key challenge is runtime engines for microcontrollers and Internet of Things (IoT) devices as well as the integration of container runtimes with real-time operating systems.

Nevertheless, it is widely recognised that containers can bring vast benefits to not only developers but infrastructure and operations teams, as well as small and large organisations. Being one of the latest developments in the evolution of cloud computing, TM R&D is currently working on a solution to tackle the containerised application in disconnected, intermittent, and limited networks that can help to fulfil the resource requirements over the edge. In line with this, we encourage those interested to reach out to us.

Reference

1. Ericsson Mobility Report, June 2022.
   https://www.ericsson.com/49d3a0/assets/local/reports-papers/mobility-report/documents/2022/ericsson-mobility-report-june-2022.pdf.
2. Mobile XR Experience, March 2023.
   https://www.qualcomm.com/research/extended-reality
3. 5gsmart.eu, March 2023.
   https://5gsmart.eu/wp-content/uploads/5G-SMART-D1.1.pdf
4. Docker, March 2023.
   https://www.docker.com/
5. Docker, March 2023.
   https://www.docker.com/
6. Kubernetes, March 2023.
   https://kubernetes.io/

Creator of Smarter Ecosystems for a better Malaysia

Established in 2000, TM R&D is the innovation arm for TM Group focusing on creating smarter ecosystems to make business and life easier for a better Malaysia. TM R&D's solutions are clustered around four (4) pillars namely Intelligent Platforms, Data Brokerage, Connectivity/Tools and IR4.0/Digital Solutions.

Growing from strength to strength since 2016, TM R&D has won multiple global awards and generated more than 2,800 Intellectual Property Rights (IPRs) and 1,400 digital assets to-date.

TM R&D's innovations are all developed in-house and cut across multiple verticals such as Utilities, Retail, Agriculture, Healthcare and Education with safety and productivity as the top priority.

As TM R&D continues to expand beyond connectivity and into smarter digital ecosystems, its role in TM has become more prominent and exciting.

We are looking for remarkable people to join us. People who are courageous enough to push boundaries, curious enough to experiment with new technologies, and who have the determination to drive new ideas forward. A new opportunity awaits you here in TM R&D.

Be a part of our family at https://www.tmrnd.com.my/jobs/ or email recruit@tmrnd.com.my

For more information about TM R&D and its products and services, visit **www.tmrnd.com.my**