# Sunlight – a new approach to edge architecture

WHITEPAPER

## A new frontier

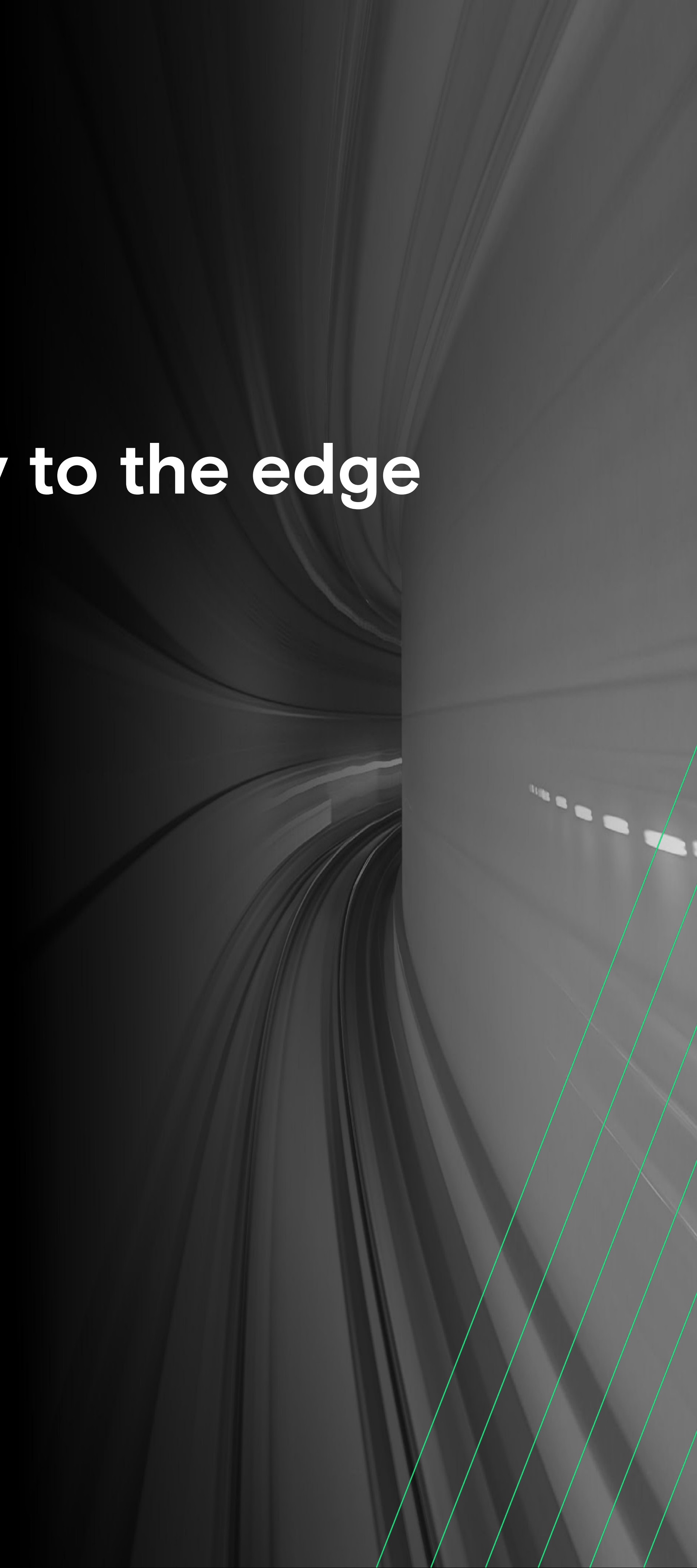
The edge computing market is exploding - largely driven by artificial intelligence, machine learning, analytics and IOT use cases. Huge volumes of data are now being processed in real-time, outside of the data center, and only a fraction of that data makes its way back into the data center.

The big problem is that the edge looks very different to the data center.

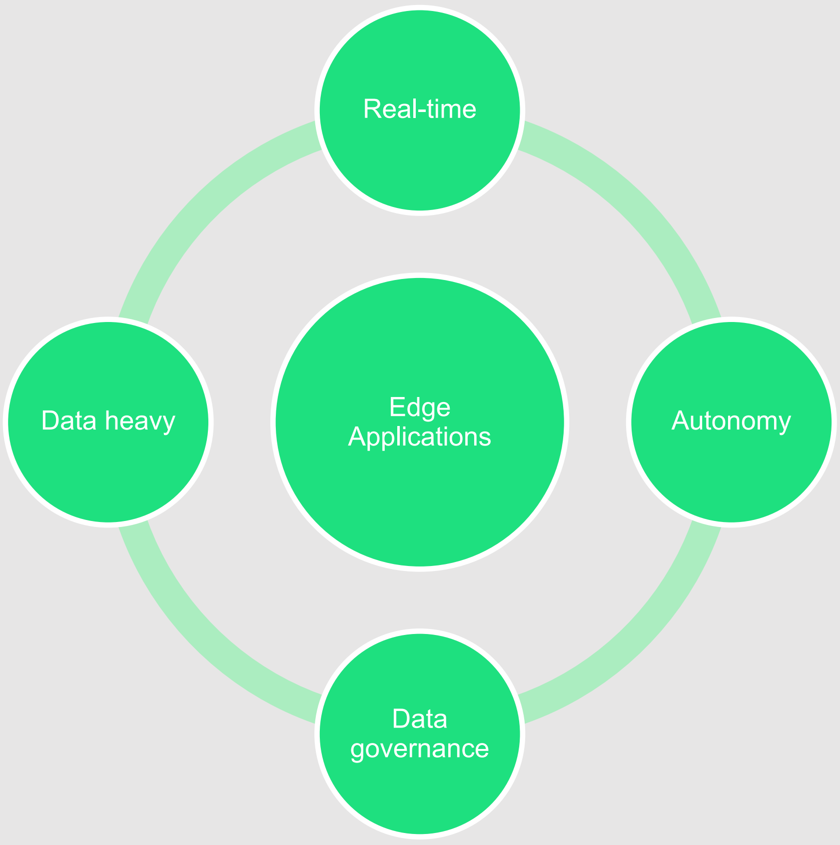
This poses several challenges – first – how to efficiently run these workloads at scale in these edge locations, and second, how to manage an infrastructure that consists of 1000s of geographically dispersed locations..

This paper looks at these challenges, and shows how Sunlight has addressed the problem.

## Chapter 1
# Journey to the edge



# Realtime interaction with physical world drives edge applications




**Why are workloads moving to the edge?**

New technologies that instrument and interact with the physical world in real time are exploding in importance. Analysis of and reaction to the huge volumes of time-series data they generate is creating new business possibilities that were simply impossible before - from new process efficiencies in a quick service restaurant to fully autonomous operations on a manufacturing production line.

**New application characteristics at the edge**

Most edge applications exhibit at least one of the following characteristics:

1. **Real-time** – need for instant decision making and action means latency must be as low as possible, for example picking defective items off a conveyor belt.
2. **Data heavy** – so much data is being generated, for example from video streams, that there isn't enough bandwidth, or bandwidth is too costly, to push data up to the cloud.
3. **Data governance** – commercially sensitive, or personally identifiable data cannot leave the location - for example in facial recognition use cases.
4. **Autonomy** – intermittent network connections cannot interrupt local operations, or systems need to be air-gapped – for example in a military forward operating base use case.

SUNLIGHT

# Building a scalable edge

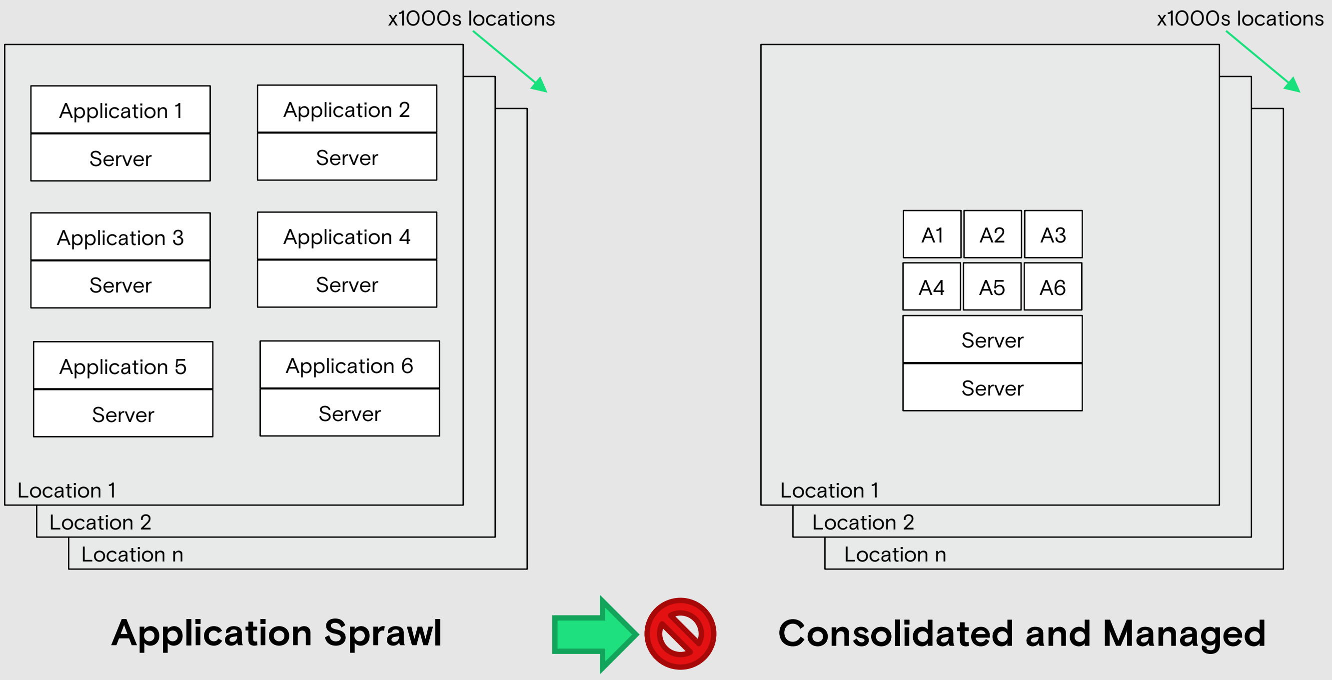
It is all too easy for an enterprise to find itself with several disparate edge projects on the go. Each project uses its own infrastructure and management tools. This can be fine for a pilot, but scaling becomes prohibitive as application sprawl leads to sky-high infrastructure and management costs when multiplied across thousands of locations. In addition, future agility is massively constrained because there is no flexibility in infrastructure.

### The barriers to consolidation

Consolidation of applications onto a common 'cloud-like' infrastructure is the clear answer to the above issue, but is difficult to achieve in practice, because of the very different nature of edge. The same platforms that allow us to do this in a data center simply don't work at the edge as we will explore in the next chapter.

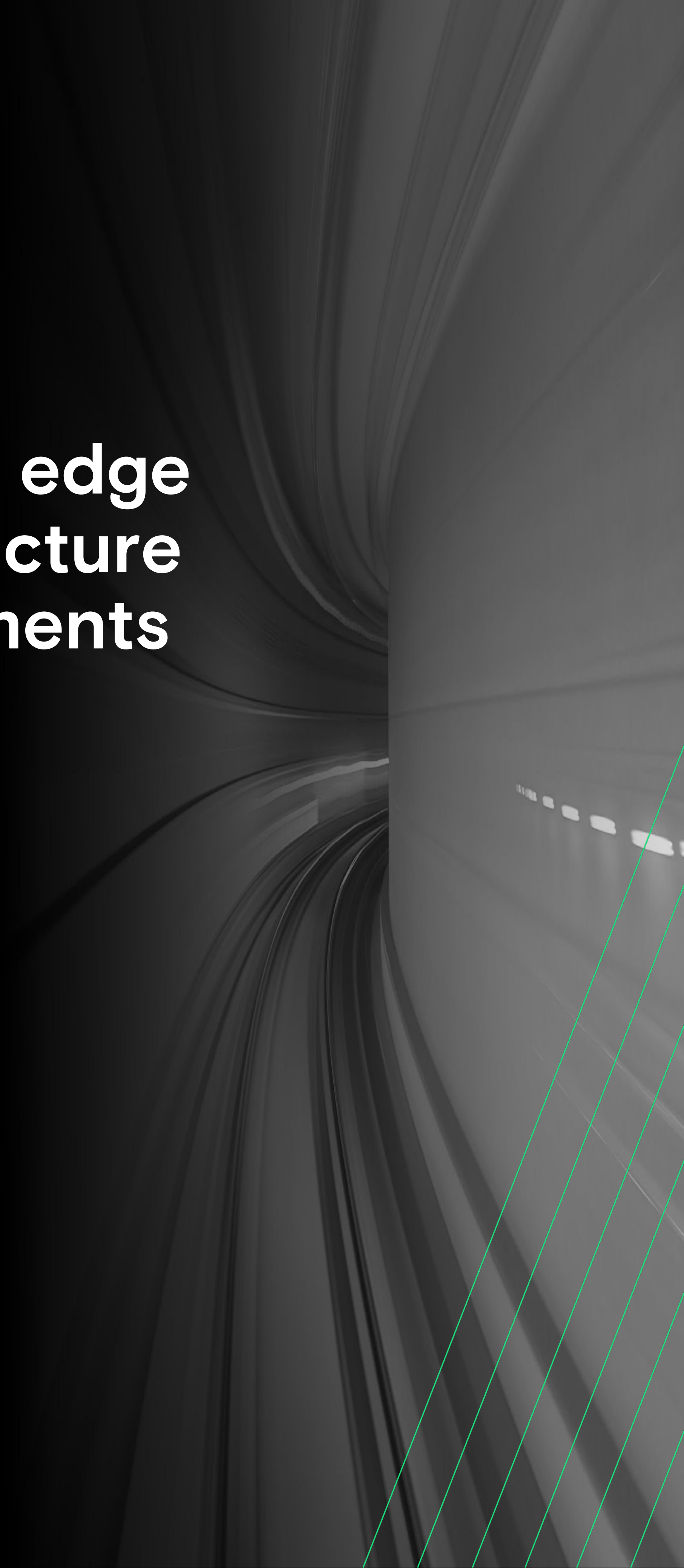
Application architectures such as Kubernetes are only part of the answer, as separate container infrastructures are still required for Linux and Windows, and legacy applications must be re-written to support them.

Being able to deploy applications into a shared infrastructure in each location is essential to minimising infrastructure and management costs and maximising business agility.



## Chapter 2

# The new edge infrastructure requirements



# Edge applications require a new type of infrastructure

> **by 2025 the volume of data will grow to 175 zetabytes, and 30% of that data will be processed directly at its source, at the edge, in realtime**

Hewlett Packard Labs

Kirk Bresniker
Chief Architect

### Data center applications

Today, the majority of applications tend to live in the cloud – in big data centers, ultimately sitting on powerful servers in a very controlled environment. Power and cooling are kept absolutely stable. The servers all sit close to the management plane, and with a constant, high bandwidth connection, and 'smart hands' are always a few steps away if something goes wrong. If the application needs more capacity - the infrastructure can be scaled out with ease. Everything is 'software defined' - making it fast and efficient to deploy and manage the precise virtual infrastructure each application needs.

### Edge applications live in the real world

In order to meet the exacting requirements of edge applications, they must be placed right next to where the data is generated and action is to be taken, as the latency and capacity of network connections makes it impossible to pump the data to a central cloud and wait for a response. In mission-critical or revenue-impacting use cases – it can be too risky to rely upon an intermittent network connection. In fact Hewlett Packard Labs' Chief Architect, Kirk Bresniker, predicts that of the 175 zetabytes of data being generated by 2025, "75% of those zetabytes will never be in the data center"

### Edge locations look nothing like a data center

Edge locations are nothing like a data center. They can be extremely harsh environments - such a hot, steamy fast-food kitchen, or a freezing wind turbine. Network connectivity can be intermittent and slow. Power and cooling can be uncontrolled. Space can be at a huge premium. Data center hardware doesn't stand a chance, which has led to a new breed of server, such as Hivecell, Lenovo's SE70 and SE350, HPE's Edgeline range and Antillion's military grade servers. These share a few characteristics. they are small, often ruggedised, have constrained compute and memory, and often use power-efficient Arm CPUs, unlike the ubiquity of x86 in the data center.

An additional complication is that there are often thousands of geographically dispersed edge locations to manage, each with a small cluster of these tiny servers. This is the opposite to a data center where there are very few locations, each with a large numbers of servers. If systems can't be maintained and problems can't be fixed remotely, an expensive and time-consuming truck-roll is required.

What this ultimately means is that traditional data center software defined infrastructure and management tools don't work at the edge.

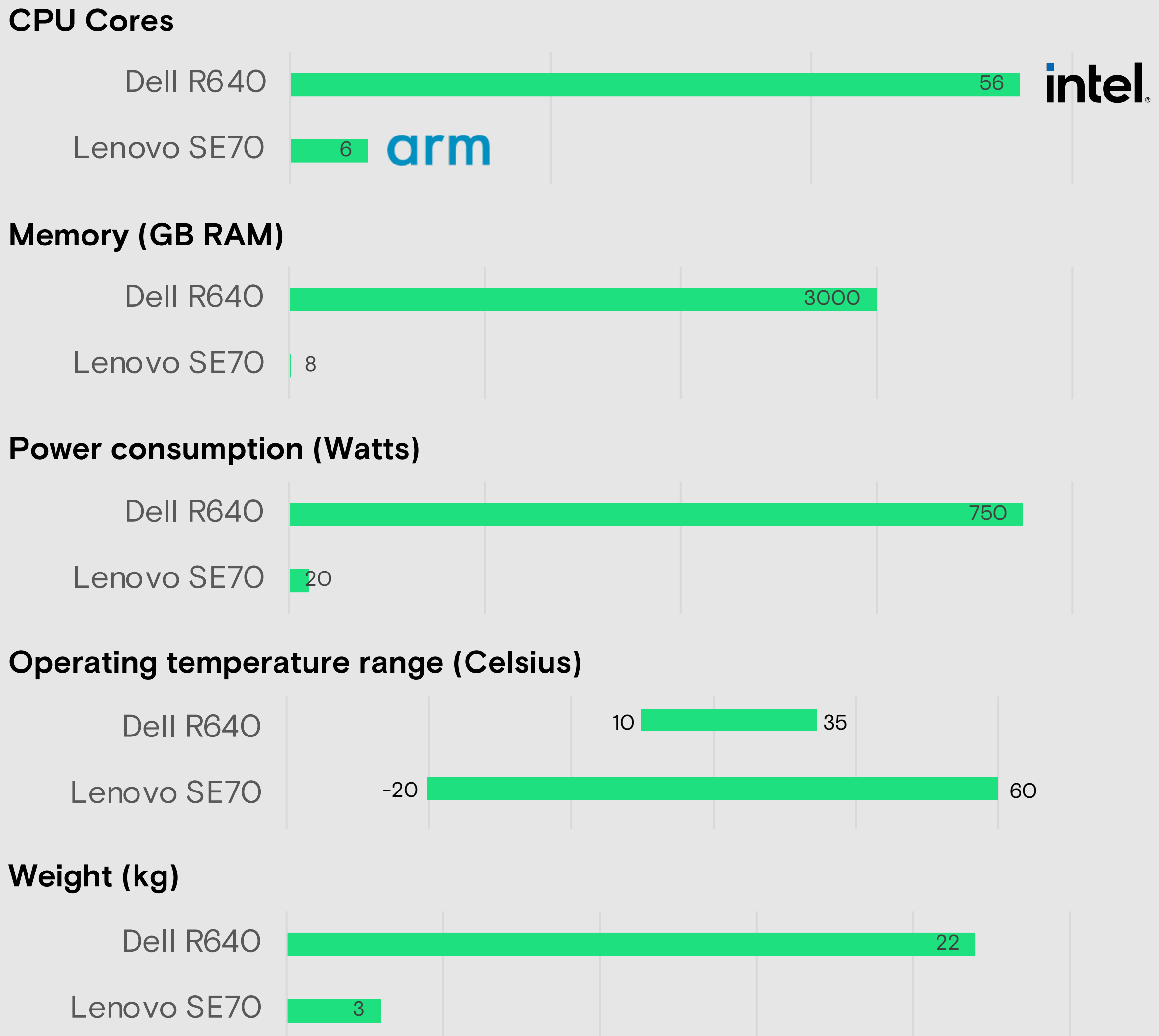
## Edge servers demand efficiency

Far edge architectures tend to rely on compact, ruggedised servers, running small scale Intel Atom or Arm processors. These systems have a significantly smaller footprint than a data center server, and as they operate outside of the data center, they must have low power requirements, and operate in very wide temperature ranges. They house fewer cores running at a lower frequency for power efficiency, and now standardise on modern NVMe storage for high performance.

## Comparing typical edge and data center servers

Edge Server - Lenovo SE70

Data center server - Dell R640



## The new edge infrastructure requirements

1. **Support a mix of application architectures** - some legacy, some cloud native on a single common infrastructure

2. **Provide high performance** – support the performance needs of AI and video analytics applications

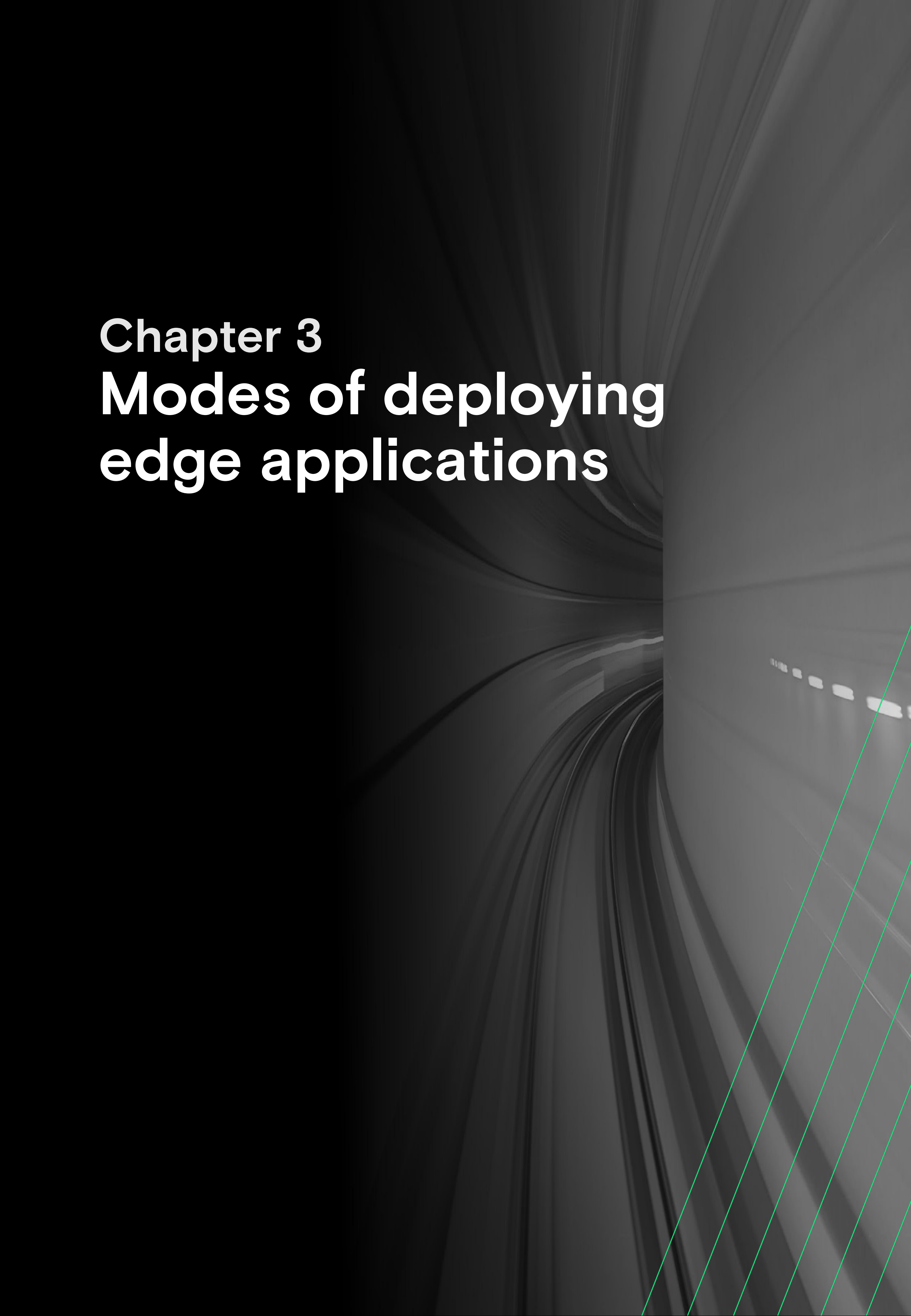
3. **Centrally and remotely manageable** – keep up-to-date and fix problem quickly without requiring a truck roll

4. **Support several applications** on a single hardware stack to keep costs low and maintainability and manageability high

5. **Fit anywhere** – cope with limited space and power with no cooling, and harsh environments

6. **Deploy-and-forget** - be easy to on-board new devices and not require special skillsets on-site to keep things moving

7. **Low cost per location** – keep hardware and software infrastructure minimal and contained for lowest TCO



Chapter 3

# Modes of deploying edge applications

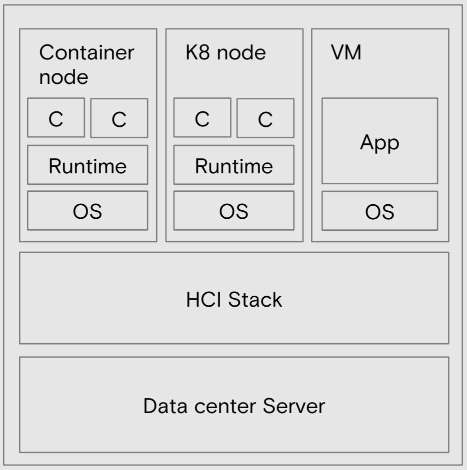
# Traditional hyperconverged

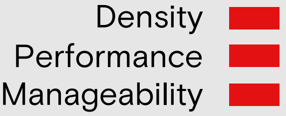
Use traditional hyperconverged solutions (e.g. Nutanix or VMware) with data center hardware.

**Pros**
- Run VM and container-based apps
- Full storage and networking services
- High availability
- Very strong resource isolation

**Cons**
- Poor remote manageability
- Very large overhead
- Requires datacenter-style hardware
- High on-site skillset requirement
- High licensing costs

**Bottom line:**
Not possible to deploy in most far-edge use cases due to footprint, performance and physical power and space constraints.




Overall TCO $$$$$

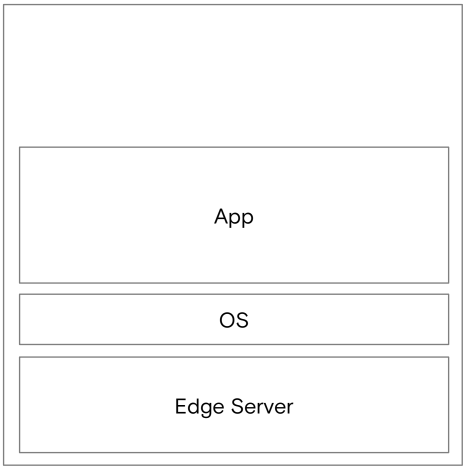
# Bare Metal

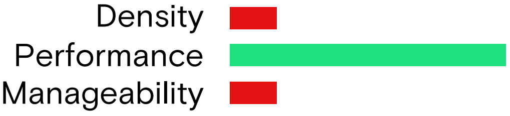
Use dedicated edge hardware for each application.

**Pros**
- High performance
- Ultimate resource isolation

**Cons**
- No hardware abstraction
- Very poor remote manageability
- High cost/low efficiency
- Each app needs its own hardware
- No high availability
- Requires storage and other services to be layered on

**Bottom line:**
Suitable for limited pilot deployments, but hardware sprawl and poor manageability make it too costly for production use



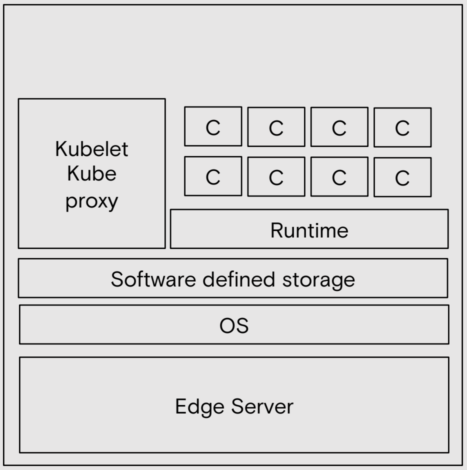




Overall TCO $$$$$

SUNLIGHT

# Kubernetes on bare metal

Run a Kubernetes distro on edge hardware.




**Pros**
- High performance
- Run container workloads

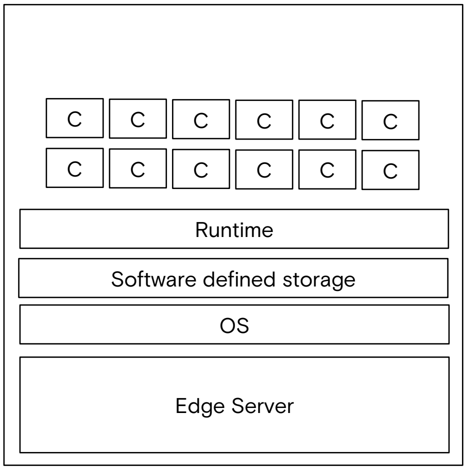
**Cons**
- Limited isolation
- Can't run VM-based apps
- Separate clusters for Linux and Windows
- Need to integrate software defined storage
- No high availability out-of-the-box
- May not work well disconnected from master running in the cloud


Overall TCO $$$

**Bottom line:**
Good for container-only deployments, but doesn't support non-containerized applications. Still requires additional products to be integrated for storage and remote management.

# Containers on Bare Metal

Run static containers on edge hardware.




**Pros**
- High performance
- Run container workloads
- Lightweight

**Cons**
- Limited isolation
- Poor manageability
- Can't run VM-based apps
- Separate clusters for Linux and Windows
- Need to integrate software defined storage
- No high availability out-of-the-box

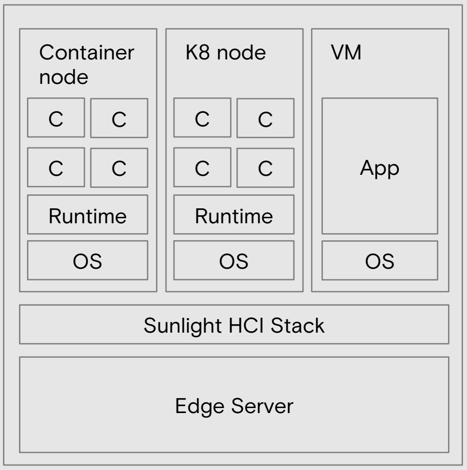

Overall TCO $$$

**Bottom line:**
Good for container-only deployments, but doesn't support non-containerized applications. Manageability a big issue.

SUNLIGHT

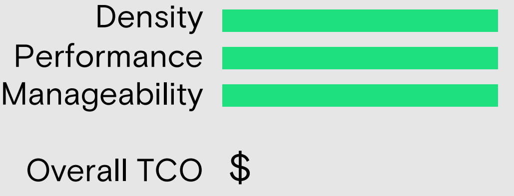
# Sunlight NexVisor edge platform

Use Sunlight running on modern edge servers with any mix of VM and container applications.




**Pros**

- Run VM and container-based apps
- Full storage and networking services
- High availability
- Very strong resource isolation
- Centralised remote management of locations
- Tiny overhead for maximum app density
- Runs on tiny x86 and Arm edge servers
- Low on-site skillset requirements
- Transparent, low licensing costs




**Bottom line:**
Perfect for deploying easily manageable edge infrastructure across 1000s of locations

**Sunlight is approximately 81% more CPU efficient and 98% more RAM efficient than industry leading hyperconverged infrastructure (HCI) solutions."**

Alex Arcilla
Senior Validation Analyst
Enterprise Strategy Group

SUNLIGHT

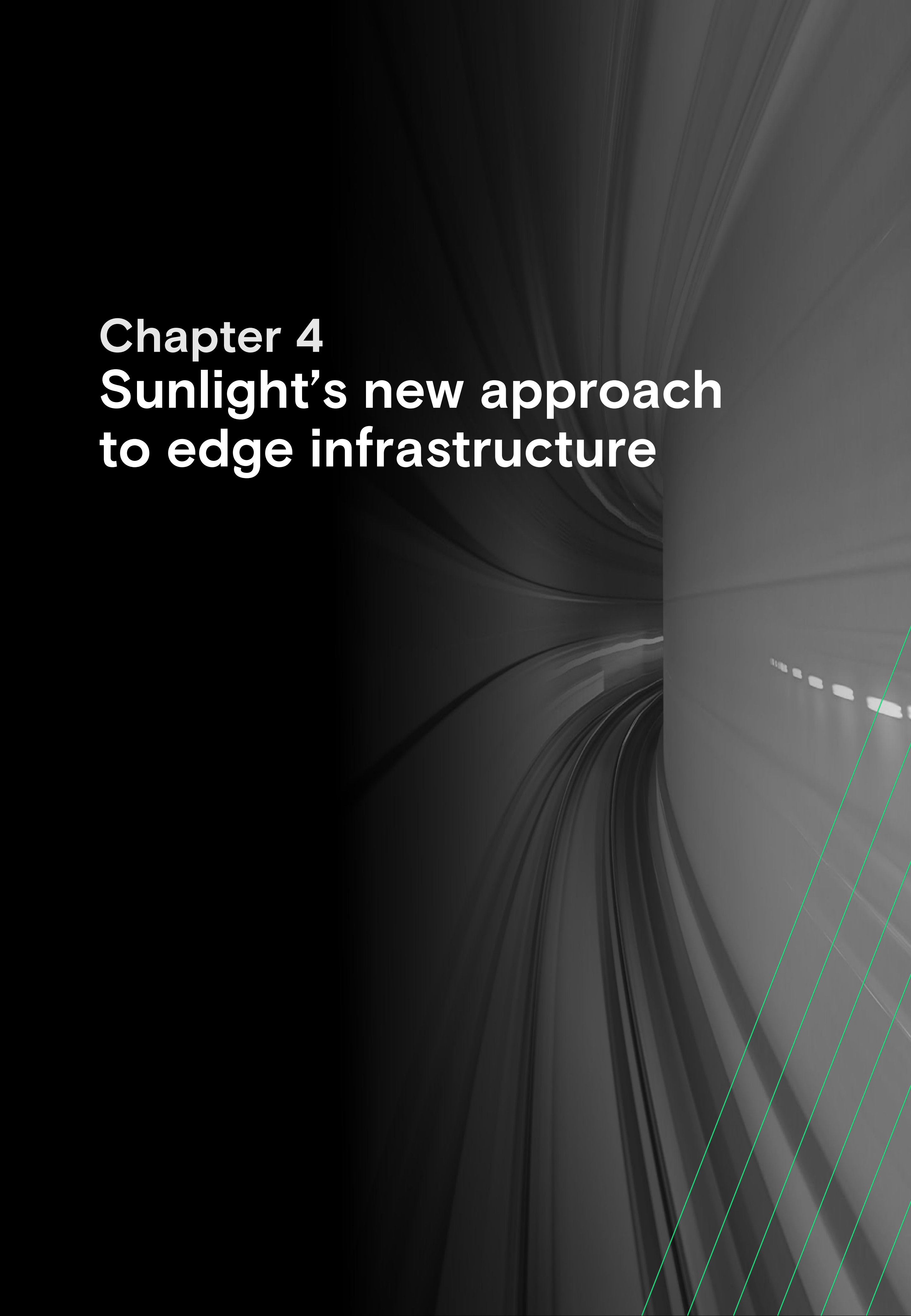
# Approaches to running applications at the edge

| Edge Requirements | Traditional HCI | Kubernetes on bare metal | Containers on bare metal | Dedicated bare metal | Sunlight |
|---|---|---|---|---|---|
| Support a mix of application architectures | ✓ | ✗ | ✗ | ✗ | ✓ |
| Provide high performance | ✗ | ✓ | ✓ | ✓ | ✓ |
| Centrally and remotely manageable | – | – | ✗ | ✗ | ✓ |
| Consolidate several applications | ✓ | ✓ | ✓ | ✗ | ✓ |
| Fit anywhere | ✗ | ✓ | ✓ | – | ✓ |
| Deploy-and-forget | ✗ | ✗ | ✗ | ✗ | ✓ |
| Low cost per location | ✗ | – | – | ✗ | ✓ |

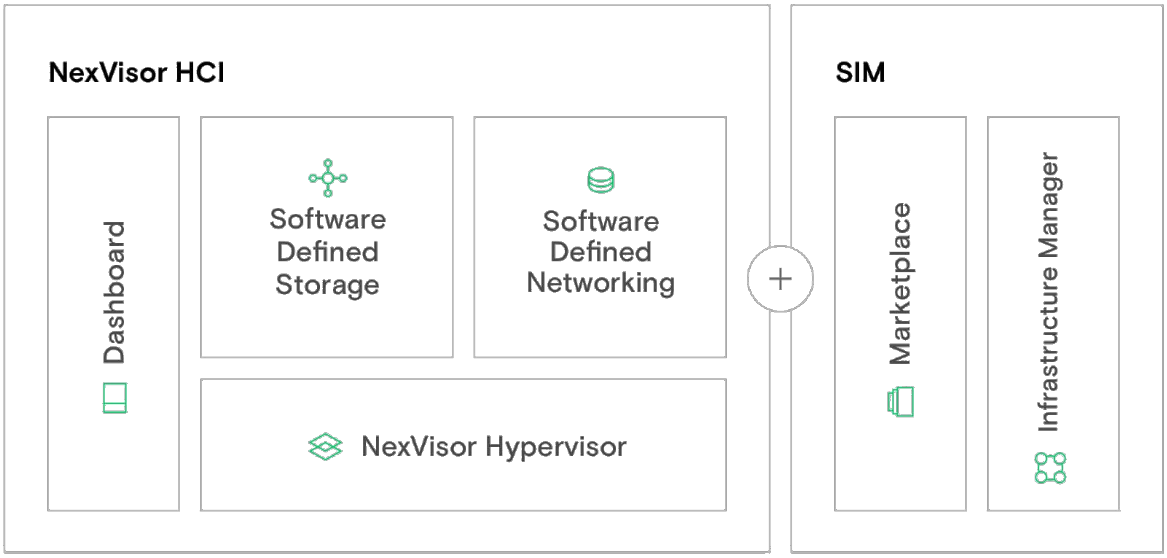
SUNLIGHT

Chapter 4

# Sunlight's new approach to edge infrastructure



# Sunlight – built for edge




### The Sunlight NexVisor Stack

Sunlight provides a complete software-defined edge infrastructure platform that makes deploying and managing edge applications easy. It enables you to run all your VM-based and cloud-native applications on a common hyperconverged platform with full high-availability and isolation – even on constrained hardware.

It's simple – all-in-one compute, storage and networking, and centralised application deployment and management. No more having to licence and integrate separate virtualization, software defined storage and management tools.

### Leverage new Arm servers

Far edge hardware is all about efficiency, which is why Arm hardware is becoming increasingly popular. Sunlight is the only Edge platform that runs on both x86 and Arm based hardware so you can choose the best tools for the job.

### Manage your edge locations

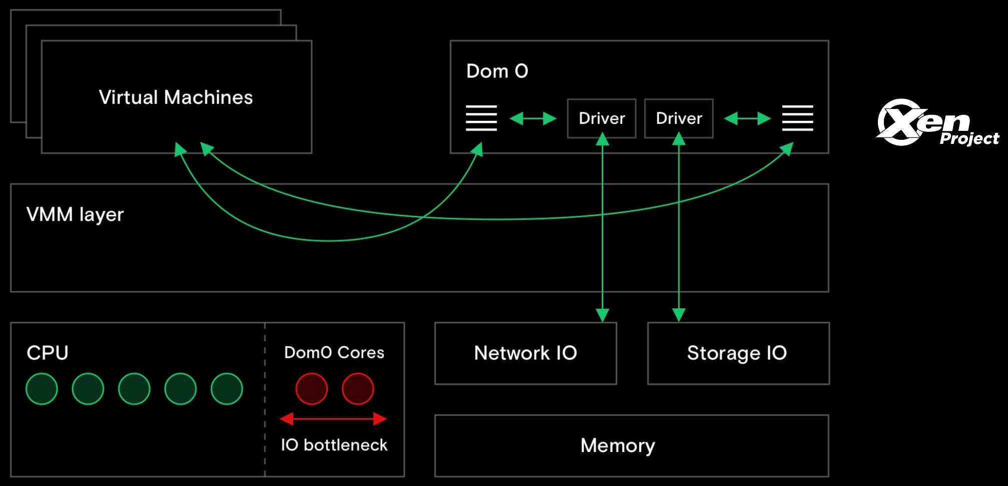
Sunlight Infrastructure Manager (SIM) centrally manages and monitors 1000's of locations edge-to-cloud from a single pane-of-glass

Together with the Sunlight Marketplace, this allows you to fully automate the end-to-end deployment of applications - all powered by the Sunlight platform and our network of certified partners.

### True edge-as-a-service

Pulling it all together – Sunlight makes it possible to deliver a true edge-as-a-service capability for your business, letting you focus on delivering business outcomes.

# Traditional HCI isn't an option at the edge




## Traditional hypervisor architectures

Traditional HCI stacks from the likes of VMware and Nutanix were born in the data center, designed to scale out across large scale, power hungry x86 hardware infrastructure. They were never designed with performance and small footprint in mind as you could always increase CPU and RAM capacity by throwing more hardware at the problem.

They were based upon existing virtualisation stacks – Xen, ESXi and KVM. When these stacks were first architected, disks were all mechanical, with typical performance from hundreds to a few tens of thousand IOPS. This meant that the IO bottleneck was always the storage interface itself, not the virtualization layer. IO would be managed from a specialised VM (in Xen architecture – Dom0) which would utilise a number of dedicated CPU cores. In fact several of these slow storage devices could be driven with a fraction of a CPU core, so this wasn't a big compromise. It didn't matter what the virtualization overhead from context switching and data copying was because it was lost in the noise.

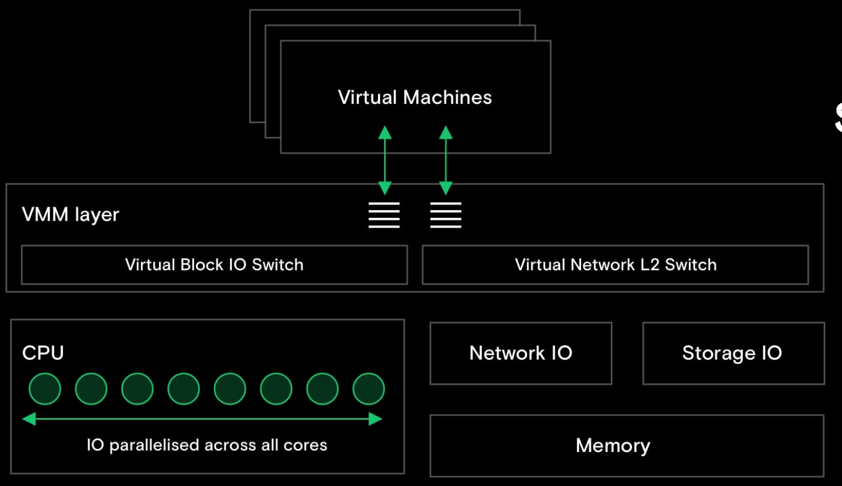
With the advent of NVMe flash storage, raw IOPs performance has rocketed – the latest Intel Optane drives can deliver 1.5 million IOPs from a single device – orders of magnitude better performance than a spinning disk. Networking devices now operate at 100Gbps. This breaks the traditional hypervisor architecture, as suddenly many CPU cores need to be dedicated to managing IO, which means fewer cores available for actually running a workload.

Hyperconvergence requires a number of components that needed to be layered on top such as software defined storage, software defined networking, management tools and so on. These were generally discrete products (e.g. vSAN, NSX and vCenter in the case of VMware), which led to complexity and bloat for the complete solution.

In addition, due to the poor IO performance of the underlying storage subsystem, software defined storage products have to resort to heavy caching – both in memory and using 2-tier storage architectures. This again consumes large amounts of valuable RAM, and increases complexity and physical footprint of the solution.

On an edge device with a handful of cores and small amount of RAM – these factors make traditional HCI stacks unusable.

# Sunlight – architected for the edge




SUNLIGHT

### The genesis of Sunlight

Sunlight was born out of a collaboration with Arm technologies in Cambridge, UK. Sunlight worked with Arm to develop a next generation virtualisation stack that could run on top of their new data center reference architecture. It was a novel scale-out architecture using embedded mobile CPUs, capable of creating dense multi-core systems. Sunlight built a virtualisation stack that could run very efficiently on top of these resource constrained platforms. This evolved into the Sunlight NexVisor HCI stack.

### Sunlight's new approach

Sunlight's new architectural approach involved ripping out all the unnecessary parts of the hypervisor stack, optimising the IO architecture considerably and building in software defined storage, software defined networking and the management plane. Some of the major innovations included

- Doing away with the 'Dom0' approach, and instead building the storage subsystem into the VMM layer itself. This makes it possible to map all of the virtual IO device queues much more efficiently onto the hardware DMA queues themselves using a zero-copy architecture.
- Parallelising IO across all of the CPU cores, rather than requiring dedicated cores. This allows the full IO capability of the underlying hardware to be presented to the fully virtualized guests running on top of NexVisor – so they appear to be running with bare metal performance.
- Incorporating all of the additional functionality required to deliver true hyperconvergence into the stack. This significantly reduced the CPU and memory footprint of the complete hyperconverged stack – making 81% more CPU efficient and 98% more RAM efficient than traditional hyperconvergence.
- Creating a management plane designed to centrally manage infrastructure and applications across thousands of intermittently connected edge clusters.

Finally – as edge infrastructure is a hybrid of x86 (for compatibility) and Arm (for efficiency). Sunlight is available on both architectures.

This complete re-architecture of the traditional hyperconverged stack makes Sunlight uniquely suited to far-edge devices

## Conclusion

The biggest challenge of edge computing is deploying at scale. As we have explored in this paper, there are many ways to deploy applications on constrained edge hardware– each with their own pros and cons. Using an edge platform is almost certainly going to be the right choice, but applying the right selection criteria is critical. These criteria revolve around central manageability, efficiency in the edge location, breath of application support and basic economics.

Sunlight's edge platform is uniquely designed to meet all these edge criteria – whilst providing cloud-like simplicity. If you'd like to give Sunlight a go – you can organise a free trial at the link below.

**“Sunlight.io's compelling message of lightweight SDI should resonate with customers looking to build, expand or modernize their edge environments.”**

Christian Perry,
Senior Analyst
451 Group

451 Research
S&P Global
Market Intelligence

Get your free trial at
sunlight.io/free-trial

sales@sunlight.io www.sunlight.io Castle Park, Cambridge, United Kingdom

SUNLIGHT

3 MAR 2022

Sunlight makes performance possible anywhere – from the cloud to the edge. Demanding applications like AI, Big Data, Analytics and Rendering run 3x faster on Sunlight compared to legacy virtualisation, and because Sunlight has a tiny footprint – it's perfect for the edge.

Enterprises and MSPs use Sunlight to cut the costs of delivering high performance IT by 70%. Sunlight is a complete HCI stack that can be deployed on-premises, in AWS, and on resource-constrained far-edge devices.

www.sunlight.io