

Turn **Big Data** Into **Fast Data**



CLOUDERA

SUNLIGHT

EDGE - CLOUD - ON PREM



Scope

This solution paper is written for Big Data professionals interested in learning more about how to cut the costs of running big data workloads in on-prem and cloud environments. It looks at how the performance problems with legacy virtualisation and cloud technologies lead to slow and expensive implementations. It highlights how a new generation of hypervisor technology unlocks the performance that big data requires at low cost.

Big Data – a critical building block for future growth

Organisations rely on insights into their data to drive their core business processes. As organisations are increasingly able to instrument every aspect of their business – using Big Data technology, such as Hadoop, Cloudera, Spark and Databricks, to mine those insights has developed into a critical building block and competitive differentiator.

However, managing that firehose of data comes at a cost – collecting and processing the sheer volume of data requires ever more compute infrastructure, straining tight IT budgets further..

The most obvious way for organisations to respond to the need for compute is to turn to the cloud, where compute resources are plentiful, and available on-demand – perfect for the elastic nature of Big Data workloads. The problem is that to access the performance needed to process large datasets becomes extremely expensive extremely quickly.

SUNLIGHT

Benefits of virtualizing Big Data workloads

In order to run Big Data workloads flexibly, many organizations today run them virtualized on-prem or in the cloud, or are planning to do so in the near future. This is for good reason, as one of the key virtualization benefits is the ability to deploy and scale-out nodes on demand when required. This virtual flexibility is impossible to achieve by deploying directly on physical hardware. To get it to work properly, you basically have no choice but to overprovision hardware requirements which quickly can become prohibitively expensive and extremely difficult to manage. By going virtual or cloud, organizations gain flexibility compared to running Hadoop on physical servers.



To become and remain competitive, enterprises must seek to adopt advanced analytics, and adapt their business models, establish specialist data science teams and rethink their overall strategies to keep pace with the competition.

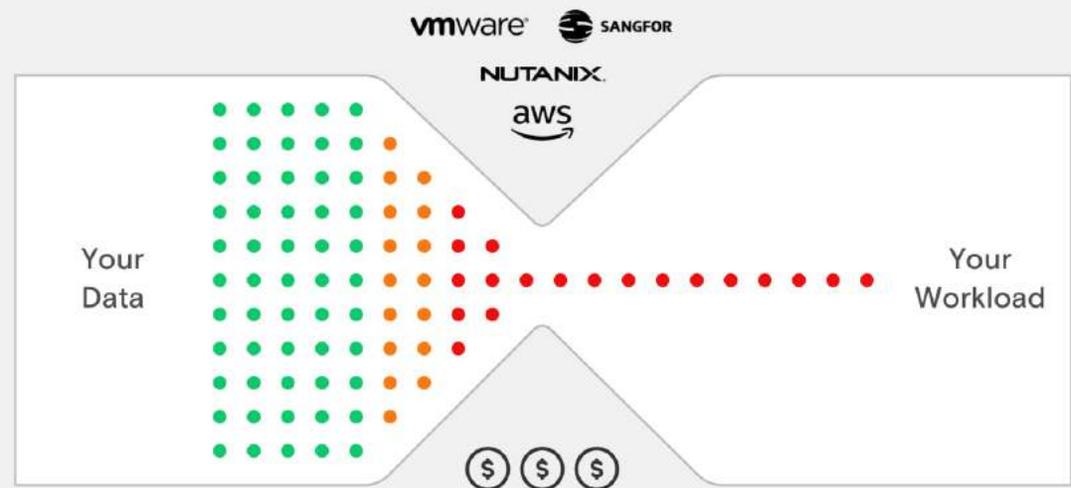
Gartner

Source: <https://www.gartner.com/en/doc/3746424-100-data-and-analytics-predictions-through-2021>

Legacy Virtualization wasn't built for Big Data

The catch is that in gaining the much-desired virtual flexibility you sacrifice performance as legacy hypervisors were simply never designed to support NVMe storage and fast networking. Some virtualization vendors add Big Data plug-ins or extensions to attempt to offset the lack of performance, but, in reality they just add extra layers of complexity and cost without delivering on the desired real performance improvements.

- ✓ Throttles new fast storage and network hardware performance by **70%+**
- ✓ Too bloated to run in constrained edge computing environments
- ✓ Necessitates massive overinvestment to achieve required performance

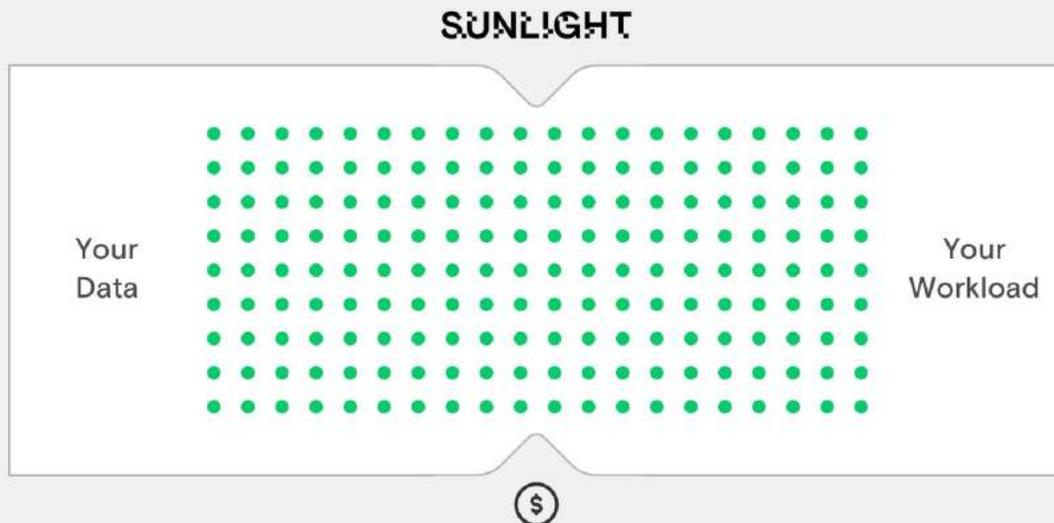


Sunlight NexVisor fixes the flexibility vs performance gap

Sunlight NexVisor is a next-gen HCI stack, which has been specifically designed from the ground up to fix the flexibility versus performance gap that exists in the market. Sunlight has architected a hypervisor that can efficiently handle Big Data workloads superfast with bare metal performance without sacrificing on virtualization flexibility.

Sunlight NexVisor takes full advantage of NVMe flash storage and fast networking giving you the ability to get 3x the performance for Hadoop, Spark, Databricks or Cloudera workloads compared to running on legacy virtualisation. This extends to your AWS workloads – which can continue to run in AWS on NexVisor with no changes. Workloads can be scaled elastically and with linear performance – giving you maximum flexibility.

Sunlight NexVisor supports hybrid deployments and can run on-prem, in the cloud (AWS) or at the edge.



- ✓ Takes advantage of near **100%** of the hardware capability
- ✓ Runs efficiently in constrained edge computing environments
- ✓ Slashes IT costs and maximizes workload performance

AT-A-GLANCE

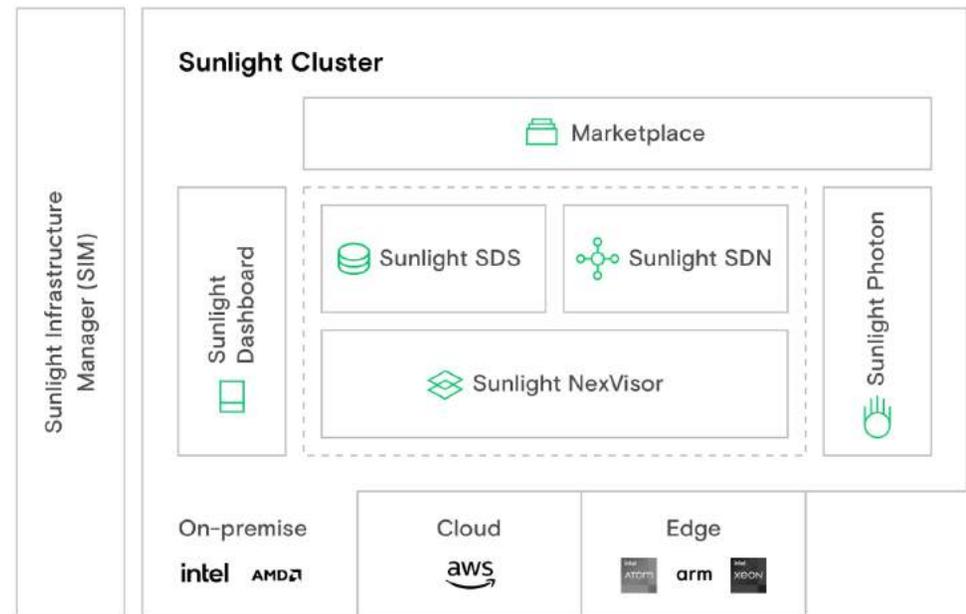
Sunlight Converged Infrastructure Platform

Sunlight is the fastest converged infrastructure platform for commodity hardware consisting of:

-  **Sunlight NexVisor**
The NexVisor hypervisor – built from the ground-up to support today’s high performance hardware technologies with almost zero overhead
-  **Sunlight SDS**
Distributed Software Defined Storage for easy scaling and low latency
-  **Sunlight SDN**
Software Defined Networking enabling creation of ethernet-based virtual private networks
-  **Sunlight Dashboard**
The easy to use Sunlight Dashboard and API to configure, monitor and manage your virtual datacenter

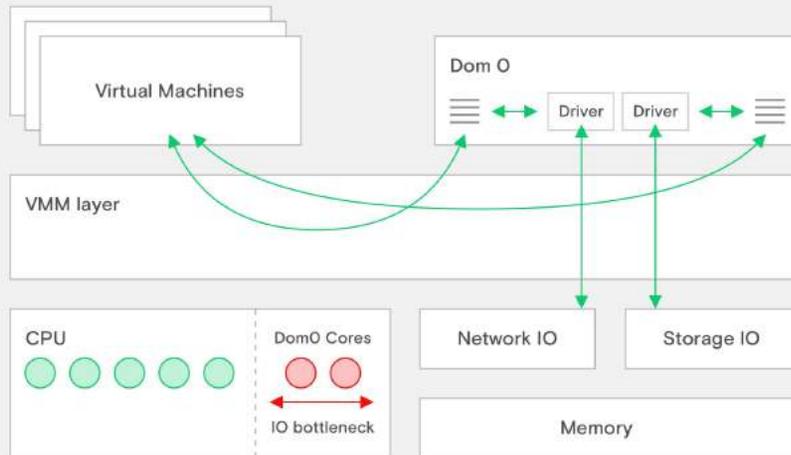
Performance increase on exactly the same hardware

3x

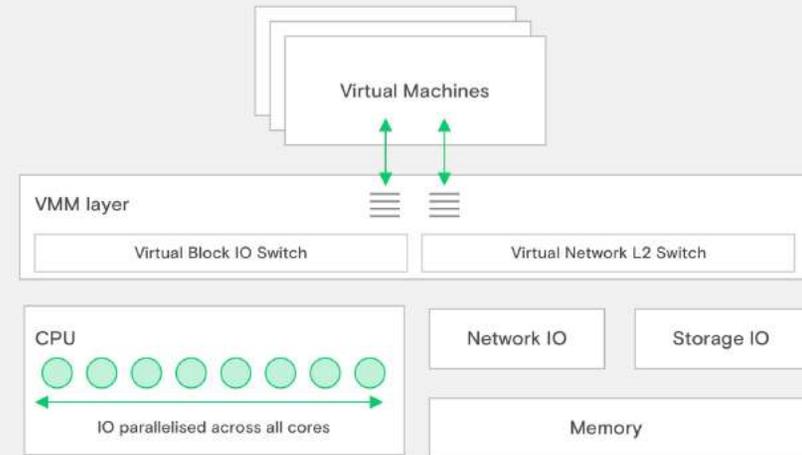


- ✓ Runs on-prem with easy migration for VMware workloads
- ✓ Runs on-demand in AWS with binary compatibility for AWS machine images (AMI)
- ✓ Sold as a simple per-server hourly, monthly or annual subscription
- ✓ Big Data products can be deployed with a single click from the Sunlight Marketplace

Legacy Hypervisor



Sunlight



Designed for Performance

Sunlight NexVisor has been purposely designed for high performance and incorporates a brand new virtualization architecture to enable customers to handle the most demanding workloads without loss of performance. Be it on premise, in the cloud or at the edge using low-energy processors.

Ultra-thin Hypervisor

NexVisor has the smallest footprint of any HCI solution – meaning you can run more VMs on your cluster for the best possible density. With legacy platforms – you’re wasting a third of your cluster’s precious memory just running the HCI stack – meaning fewer VMs, lower density and higher costs.

Out with Dom0

NexVisor makes bare metal performance for VMs a new reality by removing Dom0, or Controller VMs (CVM) and parallelizing IO across all CPU cores – thus removing the IO bottleneck. Now you can take full advantage of your hardware without paying the virtualisation tax.

Fully integrated distributed storage

Because the Sunlight HCI stack includes fully integrated and optimised software-defined storage, the performance benefits apply even when logical storage volumes are distributed across multiple NVMe drives on different physical servers – meaning high-availability comes as standard.

Performance for your Big Data workload without the headaches

Customer benefits of deploying Big Data workloads on Sunlight NexVisor include:



Maximum workload performance

Sunlight's IO performance enables Big Data workloads to run 2-3 times faster than other virtualisation and cloud solutions.



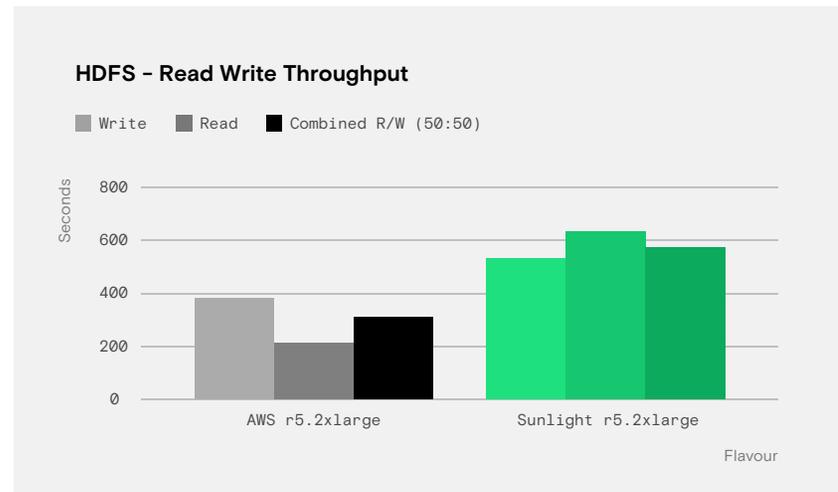
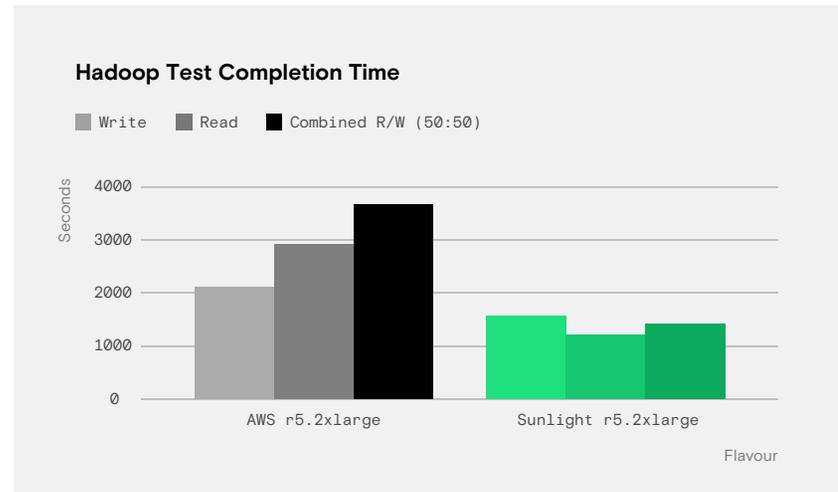
Minimum infrastructure spend

Because Sunlight delivers the full performance of your underlying hardware, and frees up CPU and RAM normally consumed by virtualization stacks - you can increase density and slash infrastructure costs.



Easy management

Sunlight's management portal makes it easy to manage your entire hybrid infrastructure from a single pane of glass



Get started Try Sunlight

If you would like to see how Sunlight can solve your Big Data performance problems without sacrificing flexibility and saving you 70% on infrastructure costs, then get in touch for a free trial

www.sunlight.io/free-trial ↗



sales@sunlight.io



www.sunlight.io



Castle Park, Cambridge, United Kingdom

SUNLIGHT

V1.0 - 9 FEB 2021

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-premise on standard data center hardware, in AWS and on resource-constrained far-edge devices.

in



www.sunlight.io

Copyright © 2021 Sunlight.io and respective copyright owners. All rights reserved.