# Cirrascale CLOUD SERVICES

## Multi-GPU Cloud Services for Deep Learning and HPC

Cirrascale Cloud Services delivers cloud infrastruture solutions for Autonomous Vehicle, Medical Imaging, and Natural Language Processing workflows, as well as various Deep Learning HPC application workloads. Our cloud service offers cost-effective, high-speed solutions with the features you need in a cloud environment that's unmatched.

#### **Bare-Metal, Dedicated Servers and GPUs**

For dedicated development, virtualized or shared resources can underperform. CCS provides bare-metal servers ensuring you get every ounce of performance you paid for.

#### **Lightning Fast Storage Options**

Cirrascale has partnered with the industry's top storage vendors to supply our customers with the absolute fastest storage options available. Specialized NVMe hot-tier storage offerings such as WekalO Matrix or IBM Spectrum Scale ECE remove storage bottlenecks faced by customers. Options to saturate a GPU cluster and deliver more than 10GBytes/second per node across an InfiniBand network are something we specialize in implementing.

#### **Kubernetes and Cluster Management**

Cirrascale can help you with your various cluster management needs. We even have experts to help with Kubernetes -- a container scheduling, management, and orchestration platform. If you're running your containerized application in a single cloud or infrastructure, you're good with single Kubernetes pod/cluster. However, what if you want to build a hybrid application which will run in a multi-cloud environment? We can help answer all those questions, and get you on the right track.

#### Multi-GPU Provides True Scale-Up Capability

CCS uses servers that are designed to maximize PCIe bandwidth and minimize intercard latency. Our x86 dedicated cloud servers utilize PCIe switch riser technology enabling up to 8 GPUs to communicate directly with each other on the same PCI bus.

#### **Discounts for Long-Term Commitments**

CCS offers pricing packages that let you plan for the entire life cycle of your projects. Commit to three, six or twelve month usage of our cloud and save up to 20%.

#### Support for All Your Deep Learning Frameworks

Deep learning researchers and developers can feel comfortable knowing their favorite GPU-accelerated deep learning frameworks have been tested and benchmarked on cloud service to ensure the absolute best performance.

#### No Suprises When It Comes to Billing

With our service you'll always know what you're paying up front. We don't charge for the ingress or egress of data, so there are never any surprises with us.

#### **High Speed Networking**

Cirrascale provides options for multiple machine connectivity through 25Gb/50Gb/100Gb Ethernet or 100Gb EDR InfiniBand so that servers are connected together in a private, protected network. With InfiniBand, servers can move data without CPU involvement with Remote Direct Memory Access (RDMA) and Send/Receive offloads that are managed and performed by InfiniBand adapters providing distinct advantages over other network communications protocols, including higher bandwidth, lower latency, and enhanced scalability to perform complex AI and deep learning tasks.

#### **Cross Connections with Third-Party Providers**

You're just a cross connect away from a faster connection to cloud providers like Amazon AWS, Microsoft Azure, and Google Cloud through our Megaport connected data centers. It can also help to reduce data egress and port fees that are charged by hyperscale data centers.

#### Always the Latest GPU Technology Available

CCS is proud to use the latest NVIDIA Tesla Accelerators in our systems, including the NVIDIA Telsa V100 GPUs. You won't be using three-generation old cards with us.

#### **POWER9 and AMD Solutions for Deep Learning**

CCS has added POWER9 and AMD EPYC<sup>™</sup> systems to its cloud platform provide customers with even more options to test their favorite frameworks. With multiple configurations available, you can select the right choice for your needs.

#### **The Highest Level of Security**

CCS provides a secure cloud environment (SSAE-16 / ISAE SOC1 Type 2 and SOC2 Type 2). For agencies requiring the highest level of security, we can deploy a complete "shared nothing" environment where nothing is shared with our public cloud instances. The only connection out is a dark-fiber connection to your facility.

### **NVIDIA Tesla GPU Instance Pricing**

GPU	Processor	RAM	Local Storage	Hourly Equiv. Pricing*	Monthly Pricing	6-Month Pricing	Annual Pricing
8X Tesla V100 (DGX-1)	Dual Intel <sup>®</sup> Xeon <sup>®</sup> E5-2698 v4	512GB	4X 1.92TB NVMe	\$20.55	\$14,999	\$13,499	\$11,999
8X Tesla V100	Dual Intel <sup>®</sup> Xeon <sup>®</sup> E5-2630 v4	256GB	1TB NVMe	\$17.81	\$12,999	\$11,699	\$10,399
8X Tesla V100	AMD EPYC 7402P	256GB	1TB NVMe	\$15.34	\$11,199	\$10,079	\$8,959
4X Tesla V100	Dual IBM POWER 9	256GB	1.6TB NVMe / 2TB SATA	\$12.78	\$9,329	\$8,396	\$7,463
4X Tesla V100	Intel Xeon E5-2630 v4	256GB	1TB NVMe	\$10.27	\$7,499	\$6,749	\$5,999
2X Tesla V100	Dual IBM POWER 9	128GB	1.6TB NVMe / 2TB SATA	\$7.94	\$5,799	\$5,219	\$4,639
8X Tesla P100	Dual Intel Xeon E5-2630 v4	256GB	1TB NVMe	\$9.59	\$6,999	\$6,299	\$5,599
4X Tesla P100	Intel Xeon E5-2630 v4	256GB	1TB NVMe	\$5.48	\$3,999	\$3,599	\$3,199
4X Tesla P100	Dual IBM POWER 8	256GB	1TB NVMe / 1TB SATA	\$7.53	\$5,499	\$4,949	\$4,399

### **NVIDIA Quadro GPU Instance Pricing**

GPU	Processor	RAM	Local Storage	Hourly Equiv. Pricing*	Monthly Pricing	6-Month Pricing	Annual Pricing
8X Quadro RTX 8000	Dual Intel <sup>®</sup> Xeon <sup>®</sup> E5-2630 v4	256GB	1TB NVMe	\$11.77	\$8,599	\$7,739	\$6,879
8X Quadro RTX 6000	Dual Intel <sup>®</sup> Xeon <sup>®</sup> E5-2630 v4	256GB	1TB NVMe	\$9.45	\$6,899	\$6,209	\$5,519

## Workstation-as-a-Service Instance Pricing

GPU	Processor	RAM	Local Storage	Hourly Equiv. Pricing*	Monthly Pricing	6-Month Pricing	Annual Pricing
2X GeForce RTX 2080	AMD Ryzen Thread- ripper 2990WX (3GHz)	64GB	512GB NVMe / 2TB SATA	\$1.75	\$1,279	\$1,151	\$1,023
2X GeForce RTX 8000	Intel <sup>®</sup> Xeon <sup>®</sup> W-2145 (3.7GHz)	128GB	1TB NVMe	\$3.83	\$2,799	\$2,519	\$2,239

## **Cloud Storage Solutions Pricing**

Solution	Capacity	Price / GB / Month		
NVMe Hot-Tier Storage	50TB or greater	\$0.20		
Object Storage	Less than 2PB	\$0.02		
Object Storage	2PB or greater	\$0.01		

\* Below pricing is based on Cirrascale's No Surprises billing model with no hidden fees. Pricing shown for servers are per server per month. Hourly equivilent pricing is shown as a courtesy to customers for comparison against other clouds. Cirrascale does not offer hourly service.

## Looking for a Different Configuration?

We have one of the largest varieties of configurations available to cloud customers in order to help meet their specific requirements. We work closely with each one to identify and develop solutions that meet their needs. Contact us today to discuss how we can help your company.

Cirrascale Cloud Services 5775 Kearny Villa Road, San Diego, CA 92123 USA Phone 888-942-3800 Web www.cirrascale.com

©2019, Cirrascale Cloud Services. All Rights Reserved. Cirrascale and the Cirrascale logo are registered trademarks of Cirrascale Cloud Services. NVIDIA and Tesla are trademarks or registered trademarks of NVIDIA Corporation or its subsidiaries in the United States and other countries. All other names or marks are property of their respective owners. No part of this document may be reproduced without consent from Cirrascale Cloud Services. Technical specifications and pricing subject to change without notice.

