

GPU AS A SERVICE

# Service Description

GPU as a Service (GPUaaS) allows users to allocate high performance graphics processing resources to virtual machines hosted on the multi-tenant Infrastructure as a Service (IaaS) platform and managed through [MyCloudSpace](#).

Softsource Limited trading as Softsource vBridge

12 April 2024

---

# In this document

[Trial Offer Terms and Conditions →](#)

[GPUaaS Pricing model →](#)

[Licensing, Features and Environment →](#)

[Base Feature Set →](#)

[NVIDIA Product Feature Set →](#)

[General Features →](#)

[Scope of Service →](#)

[Responsibilities →](#)

[Exclusions →](#)

[Customer Responsibilities →](#)

[Service Availability and KPIs →](#)

[Service Level Performance Calculations →](#)

[Incident Request Service Levels →](#)

[Pricing Structure →](#)

# Trial Offer Terms and Conditions

## 1. TRIAL OFFER DURATION

The trial offer for GPU as a Service (GPUaaS) is available for two weeks. Users are granted access to high-performance virtual GPU processing resources during this trial period.

## 2. DATA MANAGEMENT

Upon the conclusion of the trial, the user's virtual machine will be completely wiped, and all stored data will be irretrievably deleted. Users are solely responsible for managing and securely removing their data and processes prior to the end of the trial period.

## 3. SOFTWARE AND DEVICE MAINTENANCE

Softsource vBridge will not be responsible for the maintenance, support, or any issues arising from software loaded into the trial virtual machine or any devices used in conjunction with the GPUaaS.

## 4. LEGAL COMPLIANCE

Users must comply with all applicable New Zealand laws and regulations during the trial period. This includes, but is not limited to, ensuring that all software and data used or stored on the GPUaaS do not infringe on intellectual property rights and are not used for unlawful purposes.

## 5. GOOD FAITH AGREEMENT

The trial offer must be agreed upon by both parties in good faith, with clear communication and understanding of the terms set forth herein.

## 6. NOTICE PERIOD

The terms of the trial offer are two weeks unless otherwise agreed. A one day notice period for termination of the trial will be advised in writing to the users supplied contact email.

## 7. LIMITATION OF LIABILITY

Softsource vBridge's liability in connection with the GPUaaS trial is limited to the extent permitted by New Zealand law. Users agree to not hold Softsource vBridge liable for any indirect, incidental, or consequential damages arising from the use of the service.

By participating in the GPUaaS trial, users acknowledge and agree to the terms and conditions outlined above. Failure to adhere to these terms may result in termination of the trial and potential legal action.

For further information on trial periods and legal requirements in New Zealand, users can refer to the resources provided by the Ministry of Justice and Employment New Zealand.

---

# GPUaaS Pricing model

*A single GPU profile comes in various sizes, standard size being 8gb however custom sizes can be allocated depending on available capacity. Profiles are presented to the virtual machine which is then visible to the operating system as a graphics card which can then be leveraged by customer applications.*

GPU profiles that are allocated to virtual machines on the IaaS environment can be billed in two ways.

- 1. Monthly-to-month**

The GPU profile will be allocated to a virtual machine that runs on the IaaS environment for the entire month. Useful for static workloads such as virtual desktop environments.

- 2. Hourly (pay as you go)**

The GPU profile can be consumed on an hourly basis but does attract a higher hourly cost when compared to the full month commitment.

The per month or per hour pricing depends on the size of the profile allocated. *See Pricing structure for more detail*

---

# Licensing, Features and Environment

The NVIDIA licensing models for this product determine what features the operating system can access on the graphics card.

- 1. Virtual Applications (vApps)**

For organisations deploying Citrix Virtual Apps and Desktops, RDSH or other app streaming or session-based solutions. Designed for PC level applications and server-based desktops. Essentially basic graphical acceleration.

- 2. Virtual RTX Workstation (vWS)**

For users who want to be able to use remote professional graphics applications with full performance or solutions that require machine learning and AI.

Licensing is allocated on per-user model and not available on a per-hour billing basis even under the pay as you go model. User licenses are billed in monthly increments.

These graphics processing resources are situated on a highly performant, highly available and resilient set of infrastructure technologies located in TIA certified Data Centres in New Zealand. We provide this infrastructure using industry supported best practises.

To leverage the graphics processing resources the customer is required to install and use an Operating System Environments (OSE) within the virtual environment hosted on the GPUaaS platform and ensure that appropriate NVIDIA drivers are installed and functioning. The customer is fully responsible for the support and administration of the OSE and OSE data installed and running within the virtual environment.

As with virtual machines on the IaaS platform full management capability through MyCloudSpace, backup and other services are offered as standard, refer to the Infrastructure as a Service description for detailed information.

---

## Base Feature Set

### BASE DETAILS

---

#### NVIDIA Products

Support for the following NVIDIA products on a per concurrent user basis.

- NVIDIA Virtual Applications (**vApps**)
- NVIDIA RTX Virtual Workstation (**vWS**)

---

#### Virtual Machines (VM's)

To consume a GPUaaS profile an Infrastructure as a Service virtual machine must be deployed.

Any configuration of CPU, Memory\*, Storage, Networks, and Backup are allowed.

---

#### Operating Systems

Support for any OS that runs on the x86 architecture and has available NVIDIA drivers.

# NVIDIA Product Feature Set

FEATURE	VAPPS	VWS
RDSH App Hosting	✓	✓ (With vApps License)
RDSH Desktop Hosting	✓	✓ (With vApps License)
Compute Virtualisation		✓
Windows OS Support	✓	✓
Linux OS Support	✓	✓
Maximum Displays	1	4
Maximum Resolution	1280x1024	7680 × 4320 (8K)
RTX Enterprise Software Features		✓
CUDA and OpenCL Support		✓
OpenGL, DirectX and Vulkan	✓	✓
ECC and Page Retirement		✓
Profile Sizes Supported*	1gb / 2gb / 3gb / 4gb 6gb / 8gb / 12gb / 16gb / 24gb / 32gb / 48gb	1gb / 2gb / 3gb / 4gb 6gb / 8gb / 12gb / 16gb / 24gb / 32gb / 48gb

\* Support for GPU profile sizes subject to availability and capacity

---

# General Features

SCALE UP/DOWN

24x7 via MyCloudSpace

---

RESILIENCE

All components designed with N+1

---

DATA CENTRES

NZ DIA Approved Facilities.

Fully Redundant Services.

---

INFORMATION SECURITY  
MANAGEMENT

ISO 27001 Certified

# Scope of Service

This service is provided on the following basis.

1. GPUaaS platform Availability 24 x 7
2. MyCloudSpace Web Portal Availability 24 x 7

Technical support assistance between 8:00am and 5:00pm, weekdays, excluding NZ public holidays.

SERVICE CATEGORY	SERVICE DESCRIPTION	HOURS OF OPERATION
GPUaaS	Full-service availability	<b>24 hours x 7 days/week</b>  <b>Note:</b> System Maintenance is performed on a weekly basis and published in advance to all customers whether the maintenance is service impacting or not.
	Technical Support (via email, web portal or Service Desk phone number)	<b>24 hours x 7 days/week</b>  <b>Note:</b> Calls/Event alerts are logged any time, only P1 & P2 alerts will be actioned after-hours. All other Calls/Event alerts will be actioned next working day between the hours of 8:00am to 5:00pm Monday to Friday Excludes NZ public holidays.



# Responsibilities

The GPU as a Service responsibilities are focused on the provision, housing, proactive monitoring, and management of the base shared infrastructure only.

The following table outlines those elements covered in the GPU as a Service - GPUaaS offering.

ELEMENTS	SERVICES INCLUDED
<i>GPUaaS</i>	<p>Softsource vBridge are responsible for providing a compute platform that supports the customers virtual machines in a highly available, secure, and performant manner. This includes compute, storage, network, hypervisor, and backup operations. Softsource vBridge also ensures that all related components are kept up to date and that security updates are applied in a timely fashion.</p> <p>The customer is responsible for the management of WAN connections to the GPUaaS platform and all OSE installed onto virtual machines.</p> <p>Responsibilities of the Service Provider (Softsource vBridge):</p> <ul style="list-style-type: none"> <li>• Ensure that the GPUaaS platform is maintained to supported levels as stipulated by the associated vendors.</li> <li>• Monitor resources and provide tuning and advisement for GPUaaS Platform service provider issues.</li> <li>• Engineer all layers of the GPUaaS technology stack for high availability and compliance to Softsource vBridge policy.</li> <li>• Provide access to tools and capabilities that ensure the customer can operate their virtual machines through the MyCloudSpace (MCS) portal.</li> <li>• Apply patches, updates, and upgrades as recommended for the infrastructure.</li> <li>• Ensure that the platform meets performance service levels.</li> <li>• Where required provide licensing through the Microsoft SPLA programme.</li> </ul>

	<ul style="list-style-type: none"><li>• Respond to MAC requests for actions that end users cannot complete inside the portal.</li><li>• Manage capacity over the medium to long-term ensuring that customers can scale up and down to meet their needs.</li><li>• Manage changes and uptime ensuring that target service levels are achieved.</li><li>• Where required ensure that backup services are delivered to specification.</li></ul>
<i>Backups</i>	<p>Backups will be performed for the Customer Information and operating system in accordance with the following provisions:</p> <ul style="list-style-type: none"><li>• Where backup to disk is selected in the MyCloudSpace portal the Customer Information is duplicated on a disk backup and such duplicate is stored on disk separately from the original to ensure it is recoverable in the event of the original becoming deleted, corrupted, or infected by malicious software.</li><li>• Where backup to tape is selected in the MyCloudSpace portal the Customer Information is duplicated and stored on tape media separately from the original to ensure it is recoverable in the event of the original becoming deleted, corrupted, or infected by malicious software and is never lost completely.</li><li>• Backups will be run daily to an offsite location.</li><li>• Backup to Tape are performed Weekly, Monthly and Yearly as per the retention policy.</li></ul>
<i>Backup Retention</i>	<ul style="list-style-type: none"><li>• <b>Daily Incremental and Full</b> - backup to be retained for no less than 10 Business Days.</li><li>• <b>Weekly Tape Backups</b> - backup to be retained for no less than 4 Weeks.</li><li>• <b>Monthly Tape Backups</b> - backup to be retained for no less than 12 Months.</li><li>• <b>Year End Tape Backups</b> - backup to be retained for no less than 10 Years.</li></ul>

# Exclusions

The GPU as a Service excludes the following activities, tasks, and items, these being outside of the entrada support agreement.

SERVICE ELEMENTS	SERVICES EXCLUDED
GPUaaS	<ul style="list-style-type: none"> <li>• All diagnostics and remediation activities relating to the customer provisioned guest environment (processing, storage, networks security, OS, applications, etc.). As these are configured by the customer, any support is on a T&amp;M basis unless the client has engaged Softsource vBridge Managed Services.</li> <li>• Any availability / performance monitoring toolsets installed on the guest environment.</li> <li>• Technical support for the guest OS environment, including patch management, upgrades, etc. As these are the customer's responsibilities, any entrada support is on a T&amp;M basis unless the client has engaged Softsource vBridge Managed Services.</li> <li>• The provision of any technical support for all non-approved OS installed on the guest server environment.</li> <li>• The removal or retirement of any self-provisioned services, including final archive to disk - As these are the customer's responsibilities, any support is on a T&amp;M basis.</li> <li>• Management and compliance with any vendors software license used by the customer.</li> <li>• The provisioning of any anti-virus service – such products and management services can be acquired from Softsource vBridge</li> </ul>
Backups	<ul style="list-style-type: none"> <li>• Restoration of data/VM required due to customer changes. Any support is on a T&amp;M basis.</li> </ul>

The GPUaaS solution excluded activities, tasks and items also apply to.

1. All service and support calls activities that are made to Service Desk that do not relate to the in-place GPU as a Service agreement and those resources provisioned by the customer.
2. All changes to the customer provisioned infrastructure (MACs) that are deemed projects. Such changes may include OS version upgrades, resource expansions, the addition of new resources, decommissioning of self-provisioned resources no longer required by the customer.
3. Any customer-provisioned guest environment that requires support for Incident investigation and resolution. All such activities will be performed on a T&M basis.
4. All setup and support of software applications currently, or to be installed, on the guest servers within the GPUaaS environment.

---

## Customer Responsibilities

The customer is responsible for:

5. Manage the OSE including patching, updates, and configuration.
6. Ensuring that virtual machines are protected by firewall and/or other appropriate services.
7. Provision and management of WAN services and connections to GPUaaS.
8. Testing backups and disaster recovery restoration testing.
9. Notifying immediately, via the Service Desk, of any service outage or degradation in performance to any of the customer provisioned GPUaaS components or services.
10. Ensuring that only customer approved in-house IT support staff can log GPUaaS support calls with the Service Desk and that all Incidents or Service Requests are logged via the Service Desk rather than directly with individual Operations Centre staff.
11. Ensuring that all operating software, including hypervisors, software applications, installed on the GPUaaS platform are fully licenced in accordance with the software vendor licencing policies.
12. All Software used with the GPUaaS.

---

# Service Availability and KPIs

The GPU as a Service - GPUaaS availability parameters are:

SERVICE AVAILABILITY – 24/7.

PERFORMANCE MONITORING – 24/7.

The KPI target for all the above service availability parameters is 99.97%

---

# Service Level Performance Calculations

Softsource apply the following formula to determine all Service Level calculations when reporting on performance against stated targets.

Availability % = ((Monthly Standard Hours of Operation + Planned Downtime) - Unplanned Downtime)

# Incident Request

## Service Levels

- The table below states defines the Service Levels and Response Time targets of the above Service Categories, as applied to Softsource Standard Hours of Operation.
- Please note – the target times below relate to hours/minutes within the defined Softsource Standard Hours of Operation.

SERVICE   ACTIVITY	Service Level To Achieve	P1	P2	P3	P4	P5
Time to log call and acknowledge (create the Service Ticket)	100%	15 mins.	15 mins.	15 mins.	15 mins.	15 mins.
Response Time (Support Engineer commences activity)	95%	1 hour	2 hours	4 hours	8 hours	16 hours
Target Resolution Time (Ticket in Resolved Status)	95%	4 hours	6 hours	18 hours	24 hours	36 hours

---

# Pricing Structure

The charging structure for the provision of GPUaaS self-provisioning services is a monthly usage/consumption fee for resources, and monthly basis for licensing of concurrent users.

All costs are based on customer selectable parameters/quantities and may be adjusted at any time by adding /removing components from the environment.

## LICENSING

Billed monthly, per-concurrent user.

Available in nvidia vws or vapp license types.  
Feature set defined in this document.

---

## VIRTUAL RESOURCES

GPU profiles are available as

1. Monthly commit (full calendar month)
2. Pay as your go (calculated hourly)

Available for all profiles, price breaks for monthly commitment.

Standard profile size

8gb Profile

Custom sizes supported by the license type available on request and subject to availability and capacity.

Virtual Machine consumption, including backup type, operating system (if required) and other related product sku's over and above the allocated GPU profile and licensing type. See the Infrastructure as Service (iaas) product description for more information.