

Towards Hardware OpenGL Support in Cloud Computing

Eugene Denisov¹ and Irina Kalugina¹

¹ Keldysh Institute of Applied Mathematics, Miusskaya sq., 4 Moscow, 125047, Russia

Abstract

Modern realistic computer graphics applications, such as physically accurate lighting simulation systems, require a lot of computer power for images generation. Usage of the resources of cloud computing platforms for such calculations allows to avoid additional expenses for purchase and maintenance of own computer farms. However often such simulation systems use OpenGL for 3D images display, for example during scene preparation and modification. Since cloud-based virtual machines had only software (that is, slow) OpenGL display support, it was not convenient for the users to work with their habitual computer graphics applications in such environments, and typical workflow was to prepare all data on local computer and then execute simulation in the cloud service (usually using distributed processing). Recently several cloud computer service providers started to suggest users the option of hardware (GPU-based) OpenGL support in their cloud virtual machines. This article is devoted to the investigation of hardware OpenGL display options, suggested by various providers of cloud computing services, and their comparison. Available types of hardware GPU were checked and compared, along with the conditions for their usage.

Keywords

Cloud computing, computer graphics, hardware OpenGL.

1. Introduction

Modern realistic computer graphics is based on physically accurate lighting simulation in different media. The simulation requires a lot of computing power, and time of image generation for large scenes having millions of objects with complex optical properties, as well as many light sources, can be as large as tens hours. Naturally, users are looking for additional computational resources to be used for their goals. The ideas to provide such resources via Internet were elaborated in the past [1]. Now a convenient cloud computing services appeared, and they are considered as computational resources along with the other possibilities [2]. Cloud computing is applied for various resource-consuming calculations, such as overlay analysis [3] or motion capture [4].

However, users of lighting simulation and realistic visualization software would prefer to not use separate services for each time-consuming operation, but to stay in the workbench they have accustomed to, and use its habitual thus convenient user interface. This becomes possible, if to arrange Virtual Machine in cloud service with all necessary software, and perform all operations in that machine. But most of modern visualization systems intensively use OpenGL for 3D scene preview, interactive navigation and scene modelling, and sometimes even for displaying of user interface objects. For smooth interactive operations with such systems, a GPU with hardware OpenGL support is necessary.

Recently several providers of cloud computing services started to suggest their clients the virtual computing platforms with the support of hardware OpenGL. The goal of our study was to investigate what services can be really obtained from them and for what cost (under cost we consider not only money but the complexity of usage as well). This article compares hardware OpenGL features, suggested by several famous cloud computing platforms. Such comparison may help users to select proper platform, which maximally fits their needs.

GraphiCon 2021: 31st International Conference on Computer Graphics and Vision, September 27-30, 2021, Nizhny Novgorod, Russia

EMAIL: eed@spp.keldysh.ru (E. Denisov); qik@gin.keldysh.ru (I. Kalugina)

ORCID: 0000-0002-0614-9100 (E. Denisov); 0000-0001-5558-3045 (I. Kalugina)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

The following cloud computer service platforms are compared below:

- Google Cloud Platform;
- Microsoft Azure;
- Amazon Web Services.

In spite various visualization services are already provided in the cloud computing platforms, for example on Amazon [5], this does not allow users to stay completely in the frames of their customary lighting simulation software.

All three above services provide option to add various GPUs (NVIDIA and/or AMD) to their Virtual Machines (VM). Depending on user needs, there are cheap (for accelerated 3D rendering) or expensive (for GPU calculations) options. It is necessary to note, that only visualization by OpenGL is considered below; other visualization techniques (such as DirectX), as well as GPU calculations, are not considered.

In the cloud computing the speed of remote display naturally depends on network speed, which in turn depends on location of (distance to) the VM. Those services have different datacenters throughout the world (with different prices depending on location). To achieve the best visualization performance, it is necessary to select location of VM to be nearest to the user.

2. Google Cloud Platform

Google suggests various types of VM in frames of their «Compute Engine» service [6]. New users get \$300 in credit to spend for Google Cloud over the next 90 days. There are some free VM types available, but GPU there is not free and cannot be attached to free (poor) VM types.

All set of available resources in Google Cloud Platform [7] is split into «regions» and then into «zones» inside «regions». Not all zones have GPU, and those which have, sometimes have no available ones (i.e. all are busy). There is no regions in Russia. In Germany, there is one «Frankfurt» region, which has three zones, and only one zone has GPUs. In Japan, there are two regions: «Tokyo» and «Osaka». «Tokyo» region has three zones, two of them has GPUs. But at the moment of VM creation there was no available GPU in the one of these zones, and only in the other one the VM of necessary type could be created. But even after the VM was created, sometimes it could not be started because «zone does not have enough resources, try again later». It is not explained which resources are not enough, but perhaps it was GPU. One remedy is to recreate VM in «less busy» region (without warranty that VM in another region will eventually start when necessary). Another remedy is to use «reserved» VM - all equipment for VM is allocated for whole life of this VM (and is paid round the clock regardless of whether VM is turned on or not).

Only VMs of «N1» family support the visualization from GPU (i.e. hardware OpenGL). They are based on Skylake, Broadwell, Haswell, Sandy Bridge, and Ivy Bridge Intel CPU platforms; specific platform can be selected during VM creation. They are available as both predefined VMs and custom VMs. Custom VMs can be created within a wide range of memory-to-core ratio, ranging from 0.95 GB per core to 6.5 GB per core. During creation of predefined VM user can select among three possible variants: «standard» (balanced), «high memory» (more RAM), «high cpu» (less RAM). Table 1 shows all possible predefined variants for «N1» family of VMs.

Cost of these VMs includes 50 GB standard HDD, Windows Server 2019 (with license) and GPU NVIDIA Tesla T4 (with NVIDIA GRID license). Instances with one GPU can have at most 24 cores; with two GPUs – at most 48 cores; for more cores four GPUs are required. All prices are available in the price list [8].

The following configuration of «standard N1» VM was created for testing in the «Tokyo» region: 4 cores (Intel Xeon 2.20 GHz), 15 GB RAM, NVIDIA Tesla T4 attached to it, and Windows Server 2019 installed. For Windows VMs only Windows Servers are readily available, though user can provide and install own OS of any version; then its licensing/registration is up to user as well.

Table 1

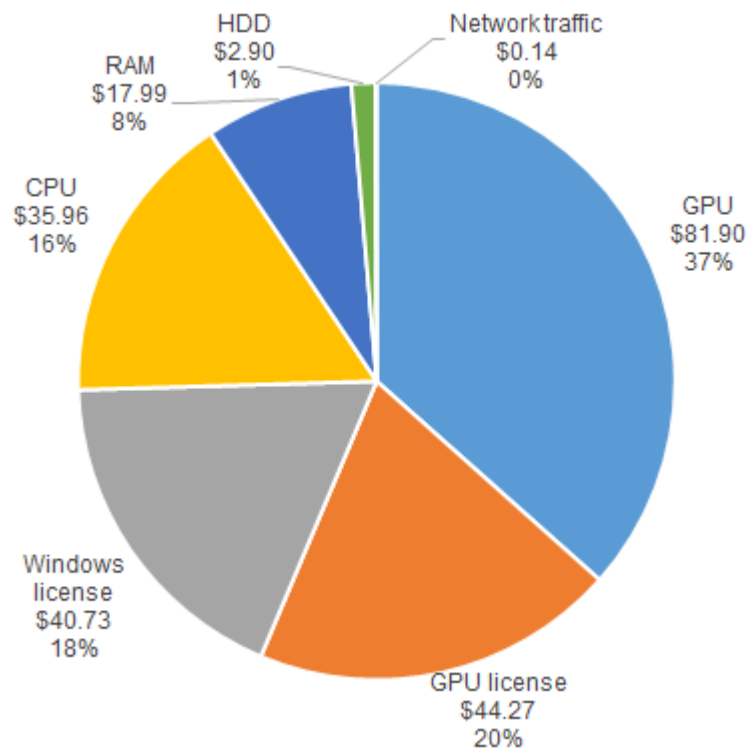
Available RAM sizes for N1 type of VM

Number of cores	Number of GPUs	Standard	High RAM	High CPU	Cost per hour for Standard	
					Germany	Japan
1	1	3.75 GB	N/A	N/A	\$0.579	\$0.551
2	1	7.5 GB	13 GB	1.8 GB	\$0.668	\$0.64
4	1	15 GB	26 GB	3.6 GB	\$0.846	\$0.817
8	1	30 GB	52 GB	7.2 GB	\$1.201	\$1.172
16	1	60 GB	104 GB	14.4 GB	\$1.912	\$1.882
32	2	120 GB	208 GB	28.8 GB	\$3.82	\$3.76
64	4	240 GB	416 GB	57.6 GB	\$7.637	\$7.516
96	4	360 GB	624 GB	86.4 GB	\$10.48	\$10.355

Only NVIDIA GPUs are available in Google Cloud Platform. After VM creation GPU must be configured - proper drivers must be installed into the guest OS. Whole procedure is well described in user manual, including link to drivers themselves. There was no any problems with it, and after drivers installation applications could successfully recognize the presence of hardware OpenGL accelerator and could use it for 3D display. Speed of GUI operations and OpenGL drawing looks satisfying taking into account remote nature of operation with VM.

Initially connection is possible via standard Windows Remote Desktop only. After first connection user can install other tools for remote access, such as Chrome Remote Desktop, Team Viewer, etc.

Below is a quote from actual billing for VM in Google Cloud. VM runtime was 221 hours (9.2 days), total sum \$224 (about \$1 per hour). This includes cost of CPU cores, RAM, GPU, Windows license, NVIDIA GRID license (necessary for usage of GPU for visualization in VM), disk space, network traffic, etc. Cost breakdown is presented on the Figure 1.

**Figure 1:** Cost breakdown

«GPU license» here is «NVIDIA GRID license» necessary for display of hardware OpenGL from GPU. This «display from GPU» option is activated and paid separately from GPU itself. If one does not need display from GPU (e.g. GPU is used for GPU calculations only), then he does not activate «display from GPU», hardware OpenGL visualization is not used and «GPU license» is not charged. It is remarkable that cost of hardware OpenGL visualization support is more than half of total sum.

Here are several comments on the way of resources charging:

- GPU, GPU license, CPU, RAM are charged for VM working hours.
- Windows license is charged for working hours * cores. I.e. for 4-cores VM one has to pay 4 times more for Windows license than for 1-core VM.
- HDD is charged round the clock (even when VM is turned off).
- Network traffic is charged for actual data transferred to/from VM.

3. Microsoft Azure

Microsoft suggests various types of VM in frames of their «Azure» service [9]. New users get \$200 for 30 days to be spent for Microsoft Azure services. After registration Microsoft Azure suggests a free VM instance - during first 12 months new user can use free of charge a VM of «B1S» size (1 core, 1 GB RAM, 4 GB HDD). GPU in Microsoft Azure is not free and cannot be attached to free VM instance.

All set of available resources in Microsoft Azure is split into «regions». Not all regions have GPUs. There is no regions in Russia. In Germany there are four regions, but none of them has GPUs. In Japan there are two regions, «Japan East» and «Japan West». «Japan West» has no GPUs, «Japan East» has it, and so VM was created there.

The complete list of available VM configurations (they call it «VM sizes») is available in [10]. All possible VM sizes with support of visualization from GPU (i.e. hardware OpenGL) are listed in the Table 2.

Table 2
Available VM sizes with GPU visualization support

CPU, GPU	Number of cores	RAM size	Cost per hour, Japan
Intel Xeon E5-2690 v3 (Haswell)	6	56 GB	\$1.778
+	12	112 GB	\$3.555
NVIDIA Tesla M60	24	224 GB	\$7.11
Intel Xeon E5-2690 v4 (Broadwell)	12	112 GB	\$2.132
+	24	224 GB	\$4.264
NVIDIA Tesla M60	48	448 GB	\$8.528
AMD EPYC 7V12 (Rome)	4	14 GB	\$0.522
+	8	28 GB	\$1.044
AMD Radeon Instinct MI25	16	56 GB	\$2.087
	32	112 GB	\$4.175

There is no possibility to get «more RAM» or «less RAM» (as in Google Cloud Platform); size of RAM is fixed for every VM size.

For testing purpose the AMD-based VM of «NV4as_v4» size (4 cores, 14 GB RAM) was selected, for several reasons, among them are: to compare cost of VM with performance similar to one used in Google Cloud Platform, and to test visualization by AMD GPU this time.

Again after VM creation GPU must be configured properly. Procedure is well described in documentations. Proper drivers can be installed into the VM using Microsoft Azure Portal. Here also no any problems were found, and OpenGL visualization worked as expected. Everything looks the same as it was in Google Cloud Platform.

Cost of that VM is ~\$0.6 per hour. It is necessary to note that the cheapest NVIDIA-based VM (6-cores one) costs 3 times more, ~\$1.83 per hour; the same NVIDIA GRID technology is used (as in Google Cloud Platform). All prices are available in online pricing calculator [11].

Like in Google Cloud Platform, connection is provided via standard Windows Remote Desktop protocol, but other remote access programs can be installed into the VM.

4. Amazon Web Services

Amazon suggests Virtual Machines in their «Elastic Computer Cloud» service (EC2, [12]). There are 21 geographical regions where VMs can be created. There are no regions in Russia. There is one region in Germany, «Frankfurt», it has GPUs. There are 2 regions in Japan - «Asia Pacific (Osaka)» and «Asia Pacific (Tokyo)». Only «Asia Pacific (Tokyo)» has GPUs. Among Windows OS only Windows Server is available, versions 2003 and newer.

All VM instance types available in EC2 are listed in [13]. Complete list of VM instance types which support hardware OpenGL visualization using GPU is presented below in Table 3. NVIDIA-based instances use the same NVIDIA GRID technology as in other cloud services.

Table 3

Available VM sizes with GPU visualization support

CPU, GPU	Num of cores	RAM size	Cost per hour	
			Germany	Japan
2-nd Generation Intel Xeon Scalable (Cascade Lake)	4	16 GB	\$0.842	\$0.894
+	8	32 GB	\$1.308	\$1.383
NVIDIA T4 Tensor Core	16	64 GB	\$2.241	\$2.361
	32	128 GB	\$4.192	\$4.41
	48	192 GB	\$7.098	\$7.489
	64	256 GB	\$8.384	\$8.819
	96	384 GB	\$14.196	\$14.978
Intel Xeon E5-2686 v4 (Broadwell)	4	30.5 GB	\$1.122	\$1.224
+	16	122 GB	\$2.161	\$2.316
NVIDIA Tesla M60	32	244 GB	\$4.322	\$4.632
	64	488 GB	\$8.644	\$9.264
2-nd generation AMD EPYC processors	16	64 GB	\$1.82	\$1.906
+	32	128 GB	\$3.64	\$3.812
AMD Radeon Pro V520	64	256 GB	\$7.28	\$7.624

However, besides classic GPUs Amazon Web Services suggests also another approach for hardware OpenGL visualization: «Amazon Elastic Graphics». Amazon Elastic Graphics provides graphics acceleration for Windows instances. Elastic Graphics accelerators are available in multiple sizes and are a low-cost alternative to using GPU graphics instance types. A user can choose an instance type that meets the compute, memory, and storage needs of his / her application. Then, add the accelerator to the instance that meets the graphics requirements. An Elastic Graphics accelerator does not form part of the hardware of the cloud service instance. Instead, it is network-attached through a network interface. An Amazon-optimized OpenGL library detects the attached accelerator. It directs OpenGL API calls from user's instance to the accelerator, which then processes the requests and returns the results.

By other words, «Amazon Elastic Graphics» can be attached to ordinary VM instances (without GPU). Cost of «Amazon Elastic Graphics» is presented in Table 4. All prices for all VM types are listed in [14].

Thus «Amazon Elastic Graphics» can be attached e.g. to instance of general purpose «t3.xlarge» type (4 cores, 16 GB RAM) which is a cheapest among 4-cores instances with cost \$0.2912 per hour, and give total cost \$0.3842 per hour, which is 2.3 times less than cheapest 4-cores instance with GPU.

«Amazon Elastic Graphics» can be used only for OpenGL acceleration; no DirectX support, no GPU-based calculations. This option might be interesting as a mean to minimize cost of VM with hardware OpenGL visualization support.

Table 4

Cost of «Amazon Elastic Graphics»

Graphics Memory size	Cost per hour, Japan
1 GB	\$0.093
2 GB	\$0.186
4 GB	\$0.372
8 GB	\$0.744

5. Conclusion

It seems selection of appropriate cloud computing service with GPU supporting hardware OpenGL is not an easy task, and typically, such feature seriously increases cost of service. Therefore perhaps such approach is justified for short-term projects requiring a lot of computing power (besides visualization), or when it is unprofitable to make serious initial investments (purchase of powerful computers). The convenience of cloud computing is that for every particular project it is possible to select most optimal configuration in cost-performance terms. Meanwhile when purchasing own PC, one cannot be sure its capabilities (e.g. RAM size), which seems to be enough for current project, will be enough for future projects as well.

6. References

- [1] B. Kh. Barladyan, A. G. Voloboi, N. I. V'yukova, V. A. Galaktionov, N. B. Deryabin, Illumination Modeling and Generation of Realistic Images Using Internet Technologies, *Programming and Computer Software* 31(5) (2005) 282-291. doi:10.1007/s11086-005-0037-1.
- [2] V. S. Vavilala, Combining high-performance hardware, cloud computing, and deep learning frameworks to accelerate physical simulations: probing the Hopfield network, *European Journal of Physics* 41(3) (2020) 035802. doi:10.1088/1361-6404/ab7027.
- [3] K. Zhao, B. Jin, H. Fan, W. Song, S. Zhou, Y. Jiang, High-Performance Overlay Analysis of Massive Geographic Polygons That Considers Shape Complexity in a Cloud Environment, *ISPRS International Journal of Geo-Information* 8(7) (2019) 290. doi:10.3390/ijgi8070290.
- [4] Y. Xie, H. Su, Research on the application of cloud computing mode in athletes' technical movement capture, *Concurrency and Computation – Practice & Experience* 31(10) (2019) e4747. doi:10.1002/cpe.4747.
- [5] M. Raji, A. Hota, T. Hobson, J. Huang, Scientific Visualization as a Microservice, *IEEE Transactions on Visualization and Computer Graphics* 26(4) (2020) 1760-1774. doi:10.1109/TVCG.2018.2879672.
- [6] Google Cloud – Compute Engine. URL: <https://cloud.google.com/compute>.
- [7] Google Cloud – Machine families. URL: <https://cloud.google.com/compute/docs/machine-types>.
- [8] Google Cloud – Compute Engine pricing. URL: <https://cloud.google.com/compute/all-pricing>.
- [9] Microsoft Azure – Virtual Machines. URL: <https://azure.microsoft.com/en-us/services/virtual-machines/>.
- [10] Microsoft Azure – Sizes for virtual machines. URL: <https://docs.microsoft.com/en-us/azure/virtual-machines/sizes>.
- [11] Microsoft Azure – Pricing calculator. URL: <https://azure.microsoft.com/en-us/pricing/calculator/?service=virtual-machines>.
- [12] AWS – Amazon EC2. URL: <https://aws.amazon.com/ec2/>.
- [13] AWS – Amazon EC2 Instance Types. URL: <https://aws.amazon.com/ec2/instance-types/>.
- [14] AWS – Amazon EC2 pricing. URL: <https://aws.amazon.com/ec2/pricing/>.