In this work, a high-performance application for visualization of large-size grid models (about a million cells) with using of Vulkan technologies was developed. Vulkan is a new software interface (API) which controls the graphic processor (GPU). Vulkan became a low-level API, thanks to which the entire GPU capability was used, such as memory and synchronization control, error checking, creation of commands performed by graphic processor, etc. Thus it shows high performance with less load on the CPU. For the operation of the application without interruption, double buffering of vertex buffer and multi-threading of the processor was used. The results of each 100th iteration of the Jacobi method for solving the Poisson’s equation, namely data of each iteration, were taken for the 2D and 3D model visualization. Using of the above given methods, grid model examples are given. As a result of this work, a prototype of a visualizer was developed and presented, and it can be used for any results of numerical mathematical modeling on structured and unstructured 3D grids.

Key words: Vulkan, 2D, 3D, computer graphics, visualization, double buffering, grid model, multithreading.
1 Introduction

Vulkan is an API for graphics and computing devices. Like an OpenGL, Vulkan allows to display real-time 3D graphics with high performance, as well as a higher productivity and a lower load on CPU. Vulkan technology was based on AMD Mantle technology. Compared with OpenGL technology, Vulkan is a low-level API and generally has similar capabilities. This allows to use all the features of GPU for computing, to get low-level access to GPU and control its functioning.

Unlike from OpenGL, Vulkan shaders are represented by a binary intermediate representation of SPIR-V programs. SPIR-V is the only supported shader language for the Vulkan. It is adopted at the level of API and used for creation of conveyors designed to control the device and for the performance of work of application.

This paper proposes a real-time visualization of large grid models with using of Vulkan technologies on the typical personal computers equipped with a discrete graphic card. To store and update data, double buffering of vertex buffers and multithreading C ++ 11 was used. The proposed approach to the visualization of the grid model optimizes the speed of drawing. The paper also presents the visualization of a numerical algorithm for solving the Poisson algorithm 2D with a size of 1000 x 1000 (10^6 cells) and 3D with a size of 33 x 33 x 11.

Let’s note that the presented application can be used for any results of numerical mathematical modeling on structured and unstructured 3D grids.
2 Literature review

Currently, the OpenGL Shader Language (GLSL) became the main part of programming with the using of OpenGL library [1–3]. With help of GLSL, you can use the power of graphics processor for display and computing. Based on GLSL, the RGL library was created, which offers three-dimensional visualization in real time [4,5]. At work [6,7], the capabilities of the graphics processor and the shader language of the OpenGL library were used, and the results of numerical mathematical modeling were visualized.

Recently, Vulkan technology is gaining great popularity in the field of visualization [8]. Vulkan is a cross-platform graphical and computing API developed by the Kronos Group consortium [9]. Vulkan is an OpenGL recipient, but it differs very much. The main difference is that to make the OpenGL driver, is the responsibility of the programmer [10,11]. Another difference between Vulkan technology and OpenGL is the shader language. Vulkan supports the only shader language SPIR-V [12]. In the literatures [13–15], you can get acquainted with the specifications of the Vulkan technology.

At work [16], languages and libraries for multi-threaded programming are described, also, how to develop programming skills, and various testing and debugging methods developed for multi-threaded programs over the past 20 years are described and demonstrated.

3 Materials and methods

The purpose of this work was to develop a high-performance application for visualizing of results of numerical mathematical modeling, by the using of Vulkan technology.

In older APIs such as OpenGL, the driver controls synchronization and memory, checks for errors during the running of the application. This is convenient for programmers, but it takes CPU time. In the Vulkan, all responsibilities are transferred to the programmer, i.e. almost all status tracing, memory and synchronization management [13].

3.1 Coordinate systems

Vulkan works with segments and triangles, representing their vertexes as points in three-dimensional space. They are called vertices and are shown in Figure 1 [3].

For beginning 50 \times 50 sizes grid was created and filled with random colors. Each square of the grid consists of two triangles, which set by the coordinates [7]. In the functions of initVertices(), the vector was filled in, which contains the coordinates and color of each vertex of the triangle.

3.2 Memory and resources

Almost for all computing systems, including Vulkan, memory is extremely important. There are two basic types of memory in Vulkan: CPU memory (host memory) and GPU memory (device memory). When creating a new object in Vulkan, you will need memory to store data. The regular CPU memory is used for it.

Vulkan works with data, and data is stored in resources. There are two basic types of resources in Vulkan: buffers and images [10]. A buffer is a simple linear piece of memory that
can be used for various purposes. They are used to store linear structured and unstructured data, which can be in format or simply be bytes in memory.

3.3 Device memory management

When Vulkan works with data, this data should be stored in device memory (device memory). In Figure 2, the GPU and CPU memory circuit, each with its own memory [10].

3.4 Device memory allocation

Device memory allocation is represented as a VkDeviceMemory object created with the help of vkAllocateMemory() functions, the prototype of which is shown below [10, 11, 14]:

```
VkResult vkAllocateMemory (VkDevice device,
const VkMemoryAllocateInfo* pAllocateInfo,
```
const VkAllocationCallbacks* pAllocator,
VkDeviceMemory* pMemory
);

After allocating the memory of the device, it can be used to store data. Thus, vertexBuffer called buffer was created and used to store the vertex data. After all works we get the following Figure 3:

![Figure 3: Grid with size of 50 × 50](image)

When visualizing in real time, new data will come in, and this is the color data of each cell. Since the coordinates and color of the cells are recorded in one buffer, the coordinates will be overwritten, and this is the waste of a precious time of CPU. To solve this problem, a separate buffer was created for these colors. As a result, we get two buffers: posBuffer - for storing coordinate data and colBuffer - for storing color data. The program will work in real time, which means data will be flowing continuously. Let’s create a new colBuffer1 buffer to store the new data.

### 3.5 Buffer update

To update the data inside the buffer, the vkCmdUpdateBuffer function was used [15]. vkCmdUpdateBuffer() copies data directly from CPU memory to buffer memory. Data is collected from the CPU memory when calling, and upon returning from vkCmdUpdateBuffer(), one can free up this memory or write new data to it.

### 3.6 Double buffering of vertex buffer

Double buffering is a data preparation method that provides the ability of out-turn of the finished result without interruption [6]. This method is commonly used when working with frame buffer. This paper describes how to use a similar approach for buffering of color data on a model of sheet. This method is as follows. A second buffer is created, and the data is filled only on it. As soon as the reading process is completed, the buffers change places, and data output will start from the second buffer, and new data will be filled into the first buffer.
This is similar to double frame buffering and is used for vertex buffers. Figures 4 and 5 show examples of drawing each buffer.

Figure 4: Drawing of the colBuffer

Figure 5: Drawing of the colBuffer1

When reading and writing new data, the drawFrame() drawing function stops and the program will not react until the data is initialized. Despite the fact that Vulkan processes on the CPU and GPU occur asynchronously, interactive control of a model as the rotation and transfer the model using the mouse and keyboard should be done on the CPU. If all actions on the central processor will occur in the same thread, we will not be able to interact with the three-dimensional model while writing to the buffer. For example, if we call the read and write function every ten seconds, and this function will be executed in six seconds, as a result we will get a program that does not respond at the time of initialization and only works for the remaining time. A visual example is shown below in Figure 6.

Figure 6: Green - Drawing, gray - Copy (A-host to host, B - host to device)

In order the input data recorded without stopping the work of drawing, it must be performed in parallel. For this we will use C++11 multithreading [16]. Modern computers have multi-core processors, in which multithreading is performed by the fact that several processes are executed on different cores. Using the multithreading feature, we pass the command to read and fill the buffer into another thread, and it will be executed in parallel without affecting into the drawing. That is, when the input data from the first buffer is displayed on the screen, the second buffer will be updated in parallel. As soon as the second buffer is ready, the buffers change places. A visual example is shown in Figure 7.

As you can see in Figure 6 and 7, initialization of data takes place in two stages: reading and writing data into the system memory (Gray A), and copying from system memory to GPU memory (Gray B). It takes longer time to initialize the reading and writing of new data into the system memory, i.e. host to host.
4 Results and discussion

To get the result, a personal computer (Core i7 3770 3.40 GHz, 8Gb DDR3) equipped with a discrete graphics card (nVidia GeForce GTX650Ti, 1Gb GDDR5) was used. As an input data, the results of each 100th iteration of the Jacobi method for solving the Poisson equation \[17,18\] with a size of \(1000 \times 1000\) for 2D and \(33 \times 33 \times 11\) for 3D, the quantity of iterations of 20,000 were taken. The Dirichlet boundary conditions for the problem under consideration are: \(u(0, y) = 0, u(1, y) = 0, u(x, 0) = 0, u(x, 1) = 0\). To visualize the three-dimensional model was used format GRDECL [19]. On Figure 8 the prototype of the application that visualizes the result of the Poisson equation grid model is shown.

Measurements were made and the results presented in Table 1 were obtained.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>2D, (1000 \times 1000)</th>
<th>3D, (33 \times 33 \times 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPS (frames per second)</td>
<td>1750</td>
<td>4415</td>
</tr>
<tr>
<td>copy host to host, ms</td>
<td>8438</td>
<td>97</td>
</tr>
<tr>
<td>copy host to device, ms</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>

When running the application for a grid model with the above mentioned dimensions, the more time takes copying host to host, i.e. reading and writing data into the system memory, which depends on computer resources: reading and writing speed of the hard disk, data transfer speed via the Internet. Copying data from system memory to device memory depends on the following GPU characteristics: memory bus, memory interface, memory transfer speed, memory bandwidth, type of dedicated video memory. Characteristics of the graphics processor on which the work was performed are shown in Table 2.

<table>
<thead>
<tr>
<th>Memory bus</th>
<th>PCI Express x16 Gen3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory interface</td>
<td>128 bit</td>
</tr>
<tr>
<td>Memory data transfer rate</td>
<td>5400 M Ghz</td>
</tr>
<tr>
<td>Memory bandwidth</td>
<td>86.40 GB/s</td>
</tr>
<tr>
<td>Type of dedicated video memory</td>
<td>1024 MB GDDR5</td>
</tr>
</tbody>
</table>

5 Conclusion

By using the Vulkan technology, a high-performance application was developed which visualizes real-time grid models of numerical mathematical modeling. In order the application
Figure 8: Examples of visualizations of the grid model of the results of the Jacobi iterative method for solving the Poisson equation: a) 2D model of the first iteration, b) 2D model of the latest iteration, c) 3D model of the first iteration, d) 3D model of the latest iteration, e) 3D model with inactive areas of the first iteration, f) 3D model with inactive areas of the latest iteration

worked without interruption, double buffering of vertex buffers and multithreading of C ++ 11 was used. The double buffering method of vertex buffers used for color data buffering on the model of oil reservoir. Such an approach to the visualization of the grid model optimized the speed of the application. An application that visualizes 3D graphics in real time, with high performance and lower CPU load, was introduced. For example, the results of the Poisson equation 2D and 3D were taken. The ready application will be used for numerical mathematical modeling results’ visualization.
6 Acknowledgements

This work was performed as part of the grant funding Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan on the topic "Development of intellectual high-performance information system for analysis of oil production technologies «iFields-II»".

References


