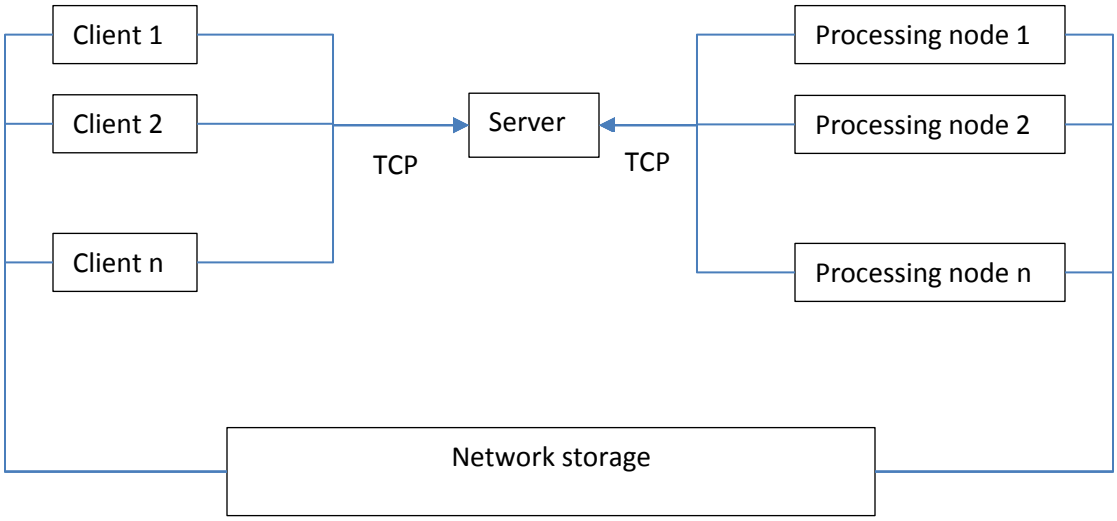
Agisoft PhotoScan Cluster Processing



Overview

Agisoft PhotoScan can be configured to run on a computer cluster where processing is distributed on several computer nodes connected to a local network. In such configuration multiple PhotoScan instances running on different processing nodes can work on the same task in parallel, effectively reducing the total processing time needed.

By default processing is divided between the nodes on a chunk by chunk, frame by frame basis (with the exception of camera alignment and optimization, where each chunk is processed as a whole on a single node). In addition a more fine grained distribution can be enabled for photo alignment and dense cloud generation, in which case processing of individual chunks/frames is further subdivided between several processing nodes.

Communication between processing nodes, server and clients is performed using TCP connections. In addition, shared network storage accessible for all processing nodes and clients is used for storing the source data and intermediate processing results.

Cluster components

Server

The server node coordinates operation of all processing nodes, maintaining a task queue for all projects scheduled for processing. A separate interface is provided for clients, which can connect to the server to start new processing tasks or monitor progress of existing ones.

The server itself doesn't perform any processing and can easily run on slow hardware. The server component is critical for operation of the cluster, and is not fault tolerant. Measures should be taken to ensure continuous operation of the server node.

The server node accepts TCP connections from processing nodes and clients on 2 separate interfaces, which can be attached to different physical network interfaces if necessary. The server node doesn't initiate any TCP connections itself.

Processing nodes

Processing nodes perform actual calculations, and thus need to be run on fast hardware. Each processing node connects to the server on startup and waits for a task to be assigned. Once the task is available, the node starts processing, informing the server about the progress. When processing is finished, the results are stored on shared network storage, and the server is informed about completion. Then a node switches to the next task when it becomes available.

Processing nodes can be added or removed from the cluster as needed. Abnormal shutdown of processing node doesn't cause cluster failure in most cases. Nevertheless, it is highly recommended to stop particular nodes using Agisoft Network Monitor before disconnecting them from the cluster.

Clients

Clients can connect to the server node to control cluster operation or to monitor its status. New tasks can be submitted for processing using Agisoft PhotoScan software configured as a network client, while cluster monitoring is performed using Agisoft Network Monitor. Multiple clients can be connected to the server node simultaneously.

Cluster setup

Before proceeding to the following steps please make sure that shared network storage is accessible from all processing and client nodes using the same absolute path. It should be either mounted to the same folder on all nodes (Linux), or should have the same UNC network path (Windows). In case such configuration is not possible (for example in mixed Windows/Linux cluster), a path prefix can be specified on each node to compensate for the differences.

Starting server

It is recommended to configure server with a static IP address, not dynamically assigned one. This IP address will be needed for each processing node and client.

The server process can be started by executing PhotoScan with the following command line arguments:

```
photoscan --server --control <ip>[:port] --dispatch <ip>[:port] [--root prefix]
```

--server parameter specifies that PhotoScan should be started in a server mode.

--control parameter specifies the network interface to be used for communication with clients. In case port value is omitted, the default port 5840 is used.

--dispatch parameter specifies the network interface to be used for communication with processing nodes. In case port value is omitted, the default port 5841 is used.

--root parameter can be used to specify network storage mount point or prefix path in case it is not the same across network.

Example:

```
photoscan --server --control 10.0.1.1 --dispatch 10.0.1.1
```

In this case PhotoScan will use the same interface for both clients and processing nodes with default port values.

Starting network nodes

The processing node can be started by executing PhotoScan with the following command line arguments:

```
photoscan --node --dispatch <ip>[:port] [--root prefix]
```

--node parameter specifies that PhotoScan should be started as a processing node.

--dispatch parameter specifies IP of the server to connect. In case port value is omitted, the default port 5841 is used.

--root parameter can be used to specify network storage mount point or prefix path in case it is not the same across network.

Example:

```
photoscan --node --dispatch 10.0.1.1
```

This command will start processing node using 10.0.1.1 as a server IP with default port 5841.

Checking cluster status

Start Agisoft Network Monitor application. In the host name field enter server IP used for client connections (10.0.1.1 in our example). Modify port values in case non-default port was specified. Click Connect button when done.

A list of available network nodes currently connected to the server will be displayed in the bottom part of the window. Please make sure that all started processing nodes are listed.

The top part lists tasks currently being processed (finished tasks are removed from the list as soon as they are completed). The task list will be empty if no processing was started yet.

Starting network processing

1. Configure Agisoft PhotoScan for network processing.

Start Agisoft PhotoScan on any computer connected to the cluster network.

Open Preferences dialog using Preferences… command from the Tools menu. On the Network tab make sure Enable network processing option is turned on, and specify the server IP used for client connection in the Host name field. In case non-default port was configured on a server, modify the port value accordingly.

In case you are going to process a few single-frame chunks with large number of photos, we recommend to enable fine-level task distribution for all supported operations (Match Photos, Align Cameras, Build Dense Cloud). In case you are going to process a large number of small chunks, or chunks with large number of frames, fine-level task distribution can be left in disabled state.

Press OK button when done.

2. Prepare project for network processing.

Open the project file to be processed. Click Save As… command from the File menu and resave project to a folder on a shared network storage in PhotoScan Unpacked Project (*.psx) format. Alternatively you can save only particular chunks using Save Chunks… command from the Chunk context menu in the Workspace pane.

Important! Please make sure that source images are located on shared network storage and not on a local hard drive. Otherwise processing nodes will be unable to load them.

3. Start processing.

Start processing using corresponding command from the Workflow menu, or using Batch Process command to execute a command sequence. A network progress dialog should appear displaying current progress.

4. Wait for processing to complete.

You can disconnect from the server at any time using Disconnect button in the Network Progress dialog in case you need to work on other projects. Processing will continue in the background.

To see processing status after you have disconnected from the server simply open corresponding .psx project on the network storage. Alternatively you can use Agisoft Network Monitor to see processing status for all projects being processed.

5. Save processing results as normal .psz project.

Once processing is finished, you can resave the project in normal .psz format. In case you have disconnected from the server, you need to open .psx project first using Open… command from the File menu. Then save it to a desired location using Save As… command from the File menu, and selecting PhotoScan Project Files (*.psz) file type.

Cluster administration

Adding processing nodes

New processing nodes can be added by starting Agisoft PhotoScan on additional computers in the network node mode, as outlined in the Starting network nodes section above.

Removing processing nodes

Aborting or disconnecting processing node during cluster operation is not absolutely safe, and can lead to project corruption if performed during final project update at the end of processing. Although the chances of this situation are relatively low, we strongly recommend against such approach. To safely disconnect processing node from the cluster you need to stop processing on this node using Agisoft Network Monitor first.

1. Start Agisoft Network Monitor. Make sure the server IP is properly configured in the host name field and click Connect button.
2. In the bottom list of network nodes identify the node to be removed. From the node menu select Pause command to stop the node after it finishes current job, or Stop command to abort processing immediately.
3. Wait until the Batch # and Progress items for a selected node become empty. This indicates that the node has finished processing. The status of the node should be Paused.
4. Now you can safely disconnect the node by aborting PhotoScan process.