Spatineo Monitor Log Upload Guide

Spatineo Monitor can analyze log files from your spatial web service and give you analysis and reporting based on that information. The analysis is based on what requests have been made to your spatial web service. To achieve this your log files must be transferred to Spatineo Monitor. This document describes which logs can be analyzed and how they are transferred to Spatineo.

The target audience for these instructions is IT professionals. These instructions and the upload mechanism is designed towards the automated transfer of log files.

Table of Contents

Quick start

To get started with uploading log files to Spatineo Monitor, you will need the following:

- Locate your access log files (see page 3 for details)
- An SFTP client software set up to send the files (see page 13 for examples)
- A key pair for authentication and authorization:
  - Private key for your safe-keeping on the server you will send the logs from
  - Public key for Spatineo to authenticate your server

Send the public key to Spatineo support. Also, if none of your log files have yet been analyzed, please send an example log file along with the public key.

We will register the key and create the appropriate log folders for you. We will contact you after everything is ready for you to start uploading the log files.

Access log

Typical web servers produce a file called an access log. This log type is typically a text which contains one row per request made to the server.

Spatineo Monitor can analyzer logs of multiple types of access logs and can work with log files with incomplete information.

Mandatory information
› When the request happened (timestamp)
› What resource was requested (URL path that matches the public URL of the service)
› What parameters were used (query string)
› HTTP method (GET, POST, HEAD, etc)

Recommended
› Requestor IP address
› Requestor User Agent (used browser and device)
› How long the request took (milliseconds or seconds)
› How much information the server response contained (bytes)
› HTTP status code

For more information see Appendix A - Supported log formats.

Logs from which server?

For the analysis to work, Spatineo Monitor needs to be able to match requests in the log files to the actual web services. Therefore the log should contain the full path to the service. The path should match the service endpoint defined in the service description. For example, for a WMS service, the URL path in GetMap requests in the log file must match the OnlineResource defined for GetMap requests.

In most typical deployments, the requests in the log files match the publicly known service endpoints. However in some cases, a frontend server or a reverse proxy is used to pass HTTP requests from the public network to internal backend servers. If the original request path is modified when passing the request to the backend servers, the log from the backend servers is typically not suitable for analysis. This is because the URL paths used in the backend servers do not match the service endpoints described in the public web service description. In these cases the appropriate logs the log from the intermediate server which logs the publicly used requests.

Which requests are analysed

The table below contains the request types that Spatineo Monitor analyses. The third column also gives an idea on what kind of information the correct log files for each service type will include.

Service Type	Request type	How to check if logs contain
Web Map Service (WMS)	GetMap (KVP)	Query parameter REQUEST=GetMap
Web Feature Service (WFS)	GetFeature (KVP)	Query parameter REQUEST=GetFeature
Web Map Tile Service (WMTS)	GetTile (KVP)	Query parameter REQUEST=GetTile
Web Map Tile Service (WMTS)	GetTile (RESTful)	Request path matches RESTful template
ArcGIS MapServer	export	Request path ends in /MapServer/export
ArcGIS MapServer	identify	Request path ends in /MapServer/identify
ArcGIS MapServer	find	Request path ends in /MapServer/find
ArcGIS MapServer	generateKml	Request path ends in /MapServer/generateKml
ArcGIS MapServer	tile	Request path ends in /MapServer/
ArcGIS MapServer	query	Request path ends in /MapServer/
ArcGIS MapServer	kml/mapImage.kmz	Request path ends in /MapServer/kml/mapImage.kmz
ArcGIS MapServer	layers	Request path ends in /MapServer/layers
ArcGIS MapServer	queryRelatedRecords	Request path ends in /MapServer/
ArcGIS MapServer	legend	Request path ends in /MapServer/legend
ArcGIS MapServer	exportTiles	Request path ends in /MapServer/exportTiles

Please note that all the above matching is case insensitive.

Transferring log files

Log files are transferred using the SSH File Transfer Protocol[1] (also known as Secure File Transfer Protocol, or SFTP). This protocol ensures a robust authentication scheme and encrypted network traffic.

Server	upload.spatineo.com (54.247.70.234, static IP)
Port	22
Protocol	SSH File Transfer Protocol (SFTP)
Authentication	RSA key(s) + username (your main account name)

Spatineo recommends that you set up an automatic daily job which sends your newest completed log file. It may also be necessary to send older log files, so also that data can be used for reporting and ad-hoc analysis.

SFTP practicals

You should take care and inspect the following rules when designing your automated transfer.

- Folders appear as subdirectories in the remote SFTP file system root
- Folders (or subdirectories) may not be created using SFTP
- Files may be compressed with gzip
- Compressed archive formats (ZIP, TAR, RAR) are not supported
- Files may only be uploaded to folders or the special auto-folder
- Uploaded files are discarded if the file format cannot be auto-detected or the format does not match folder configuration
- Duplicate files are detected and discarded
- Only completed transfers will be saved for processing
- Files in a single folder must have a unique file name

> **NOTE!** Removing log files from the SFTP service will also remove any analyzed data derived from these files in Spatineo Monitor!

Key format

Spatineo Monitor accepts authorization via RSA keys. At least 2048 bit keys should be used. The key should be in either OpenSSH or RFC4716 format. See Appendix B for examples of these types of files. DSA keys are not supported.

You may configure multiple SSH keys for your account.

Apr 2015 At the time of writing this manual there is no way to upload SSH keys in Spatineo Monitor. Please send the public part of the SFTP key you wish to use to Spatineo by email: support@spatineo.com

1 http://en.wikipedia.org/wiki/SSH_File_Transfer_Protocol

Log folders

Log files sent to Spatineo Monitor are collected into folders. Each folder may contain log files for multiple different web services. It is generally good practice to send only one type of log file into one folder. This is especially important when folders contain configuration information regarding the type of log files it will contain.

There is no limit to how many log folders one may have.

Name	Visible name of the folder, also used when transferring files. Recommended name is the public DNS name of the server from which logs are uploaded to the folder.
Web service(s)	Which service or services are associated with the folder.
Timezone of log files	Required when using a log file format which lacks timezone information.
Log format	Required when using a log file format which cannot be auto-detected.

When uploading files to an existing folder, the upload service will detect if the log files contain requests to services not yet associated with the folder. This typically happens when you add new service endpoints to an existing server. When such services are found, the folder will be automatically extended, so log analysis can be performed for these services as well.

Auto-detection folder

In addition to normal folders, the log service provides a special folder called “auto”. When uploading files to this folder, the system will automatically determine to which folder it should upload the file to:

- The log file will be placed in the folder that is associated with services found in the log file. The folder will be modified if the file contained services not yet associated with the folder.
- If no folder is associated with the services in the log file, a new folder will be created. This folder will be named after the domain name common to all services found in the log file.

Only services monitored by Spatineo Monitor can be auto-detect. The monitoring leaves clues that the analysis can pick up when analysing the log files. This auto-detection works only for log files containing created after September 19th 2014.

Apr 2015 At the time of writing this manual you cannot configure folders within Spatineo Monitor. You can however use the auto folder as specified above. If you require more control of the log folder configuration, please contact Spatineo Support.

Appendix A - Supported log formats

Microsoft IIS - W3C Extended Log File Format

The default log file format, W3C Extended Log File Format[2], is supported for analysis. The log files are typically named u_exYYMMDD.log (where YY is the last two digits of the year, MM is month of year in two digits and DD is day of month in two digits).

This log format however does not include time zone information. This information must be provided to Spatineo Monitor via configuring the target folder.

Logging should be configured so that the following fields are present:

```
date time cs-method cs-uri-stem cs-uri-query c-ip cs(User-Agent) sc-status time-taken sc-bytes
```

Example content of suitable log file (an empty row depicts a new line in the log):

```
#Software: Microsoft Internet Information Services 7.5
#Version: 1.0
#Date: 2012-04-08 00:00:51
#Fields: date time s-ip cs-method cs-uri-stem cs-uri-query s-port cs-username c-ip cs(User-Agent) sc-status sc-substatus sc-win32-status time-taken


2012-04-08 00:00:51 193.166.21.134 GET
/ArcGis/services/INSPIRE/SYKE_Hydrografia/MapServer/WMSServer.agsx
VERSION=1.3.0&SERVICE=WMS&REQUEST=GetMap&LAYERS=Reporting.WFDCoastalWater&STYLES
=&CRS=CRS
%3a84&BBOX=23.658734797013413%2c63.42763513451248%2c24.987023158488526%2c64.7559
2349598759&WIDTH=256&HEIGHT=256&FORMAT=image%2fpng&EXCEPTIONS=XML 80 -
46.137.73.20 Spatineo+Serval/0.2 200 0 0 1953


2012-04-08 00:00:52 193.166.21.134 GET
/ArcGis/Services/INSPIRE/SYKE_Maanpeite/MapServer/WMSServer.agsx
VERSION=1.3.0&SERVICE=WMS&REQUEST=GetMap&LAYERS=CorineLandCover2000_25m&STYLES=&
CRS=CRS
%3a84&BBOX=25.569043403499037%2c62.90239286671174%2c28.866385808635236%2c66.1997
3527184794&WIDTH=256&HEIGHT=256&FORMAT=image%2fpng&EXCEPTIONS=XML 80 -
46.137.73.20 Spatineo+Serval/0.2 200 0 0 562
```

2 http://www.microsoft.com/technet/prodtechnol/WindowsServer2003/Library/IIS/be22e074-72f8-46da-bb7e-e27877c85bca.mspx?mfr=true

NCSA Common / Combined

The NCSA Common[3] log format is a very common access log format. The combined log format is a variation where two additional fields (referrer and user agent) are appended to the end of each row. This format is supported by Spatineo.

This log format however does not include information about how long requests take.

Example content of suitable log file (an empty row depicts a new line in the log):

```
88.113.90.70 - - [05/Sep/2012:07:42:36 +0300] "GET /wms?LAYERS=bar
%3Afoo_kaavat_tm35&TRANSPARENT=true&ID=95&STYLES=&FORMAT=image
%2Fpng&SERVICE=WMS&VERSION=1.1.1&REQUEST=GetMap&EXCEPTIONS=application
%2Fvnd.ogc.se_inimage&SRS=EPSG
%3A3067&BBOX=319488,6823936,320000,6824448&WIDTH=256&HEIGHT=256 HTTP/1.1" 200
30255
"http://www.paikkatietoikkuna.fi/widget/web/fi/julkaisijankartta/-/MapPublished_
WAR_mapportlet?id=490" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:15.0)
Gecko/20100101 Firefox/15.0"

193.111.93.44 - - [05/Sep/2012:07:43:16 +0300] "GET
/wms/b670654d446fe018e28e491e30a310a0?LAYERS=bar
%3Afoo_vkartta_gk24&SERVICE=WMS&VERSION=1.1.1&REQUEST=GetMap&STYLES=&FORMAT=imag
e%2Fjpeg&SRS=EPSG
%3A3878&BBOX=24496101,6813108,24516581,6833588&WIDTH=256&HEIGHT=256 HTTP/1.0"
200 4614
"http://trew38.tac.fi/facta/VAADIN/widgetsets/com.logica.facta.vaadin.component.
gwt.FactaWidgetSet/kartta.html" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
5.1; Trident/4.0; .NET CLR 1.1.4322; .NET CLR 2.0.50727; .NET CLR 3.0.4506.2152;
.NET CLR 3.5.30729; InfoPath.3)"

88.113.90.70 - - [05/Sep/2012:07:44:27 +0300] "GET /wms?LAYERS=bar
%3Afoo_kaavat_tm35&TRANSPARENT=TRUE&ID=95&STYLES=&FORMAT=image
%2Fpng&SERVICE=WMS&VERSION=1.1.1&REQUEST=GetMap&SRS=EPSG
%3A3067&BBOX=320256,6823168,320512,6823424&WIDTH=256&HEIGHT=256 HTTP/1.1" 200
13450 "http://www.paikkatietoikkuna.fi/web/fi/kartta?
zoomLevel=7&coord=326802_6822672&mapLayers=base_35+100+!default!
&showMarker=false&forceCache=true" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:15.0)
Gecko/20100101 Firefox/15.0"
```

3 http://en.wikipedia.org/wiki/Common_Log_Format

Tomcat AccessLogValve

Tomcat application servers produce access log via the AccessLogValve[4] mechanism. Spatineo supports the log file produced with this mechanism. The default log pattern produced by Tomcat is NCSA Common and thus covered by the previous log format. NCSA Combined is also a typical setting for Tomcat.

When a more advanced pattern is used, the pattern must be configured to the target folder as there is however no way to auto-detect this format.

To support logs with time taken in microseconds, the Tomcat AccessLogValve supports a special field %{TimeTakenMicroseconds}.

The recommended pattern for logs is NCSA Combined with time taken appended to the end:

```
%h %l %u %t "%r" %s %b "%{Referer}i" "%{User-Agent}i" %D
```

If your logs contain further non standard information you would like to include in the log analysis, the Tomcat log format allows you to specify special user defined fields. These fields will be visible in Spatineo Monitor as custom dimensions in the log analysis. Such custom fields might indicate whether the request is coming from an internal network or a public website, or it could be used to group users into categories (“free”, “customer”). You may include up to 4 such custom fields.

To make full use of this feature, contact Spatineo support.

4 http://tomcat.apache.org/tomcat-5.5-doc/config/valve.html

HAProxy access log

The popular HAProxy server software produces a well known and supported access log format[5]. Default format used by the server has most information required by the analysis. However servers may be configured to store request header fields. This is strongly recommended as without headers the referer and user agent information is missing. These are required for analysis based on the used technology or source of traffic.

To take advantage of the headers, the folder must be configured with the relevant header configuration.

Example content of a suitable log file that contains both referer and user agent information in the headers (an empty row depicts a new line in the log):

```
May 14 10:02:09 localhost.localdomain haproxy[27102]: 10.30.54.19:48093
[14/May/2014:10:02:09.077] main gsnlvector/gsnlvector3 0/0/0/552/553 200 2752 -
- ---- 12/12/4/1/0 0/0 {acceptatie.geodata.nationaalgeoregister.nl|Jakarta
Commons-HttpClient/3.1|||} "GET /global/brtachtergrondkaart/wms?
BBOX=375200.96%2C22598.08%2C595401.9199999999%2C683200.96&TRANSPARENT=TRUE&EXCEP
TIONS=application%2Fvnd.ogc.se_xml&VERSION=1.1.1&FORMAT=image
%2Fpng&SERVICE=WMS&HEIGHT=768&LAYERS=brtachtergrondkaarttijdelijk&REQUEST=GetMap
&STYLES=&WIDTH=256&SRS=EPSG%3A28992 HTTP/1.1"


May 14 10:02:09 localhost.localdomain haproxy[27102]: 10.30.54.19:52727
[14/May/2014:10:02:09.155] main gsnlvector/gsnlvector4 0/0/0/614/615 200 14249 -
- ---- 11/11/3/1/0 0/0 {acceptatie.geodata.nationaalgeoregister.nl|Jakarta
Commons-HttpClient/3.1|||} "GET /global/brtachtergrondkaart/wms?
BBOX=375200.96%2C683200.96%2C595401.9199999999%2C903401.9199999999&TRANSPARENT=T
RUE&EXCEPTIONS=application%2Fvnd.ogc.se_xml&VERSION=1.1.1&FORMAT=image
%2Fpng&SERVICE=WMS&HEIGHT=256&LAYERS=brtachtergrondkaarttijdelijk&REQUEST=GetMap
&STYLES=&WIDTH=256&SRS=EPSG%3A28992 HTTP/1.1"


May 14 10:02:09 localhost.localdomain haproxy[27102]: 10.30.54.19:45399
[14/May/2014:10:02:08.977] main gsnlvector/gsnlvector1 0/0/0/822/823 200 3383 -
- ---- 10/10/2/0/0 0/0 {acceptatie.geodata.nationaalgeoregister.nl|Jakarta
Commons-HttpClient/3.1|||} "GET /global/brtachtergrondkaart/wms?BBOX=-
285401.92%2C683200.96%2C375200.96%2C903401.9199999999&TRANSPARENT=TRUE&EXCEPTION
S=application%2Fvnd.ogc.se_xml&VERSION=1.1.1&FORMAT=image
%2Fpng&SERVICE=WMS&HEIGHT=256&LAYERS=brtachtergrondkaarttijdelijk&REQUEST=GetMap
&STYLES=&WIDTH=768&SRS=EPSG%3A28992 HTTP/1.1"
```

5 http://www.haproxy.org/download/1.6/doc/configuration.txt

Overview of log formats

Format	Pro	Con
Microsoft IIS - W3C Extended Log File Format	Auto-detected by Spatineo Monitor Log format fields are included within the log file: › No configuration required › Spatineo Monitor can adapt dynamically to format field changes	Timezone must be manually configured for the target folder May lack crucial log fields depending on server configuration.
NCSA Common	Auto-detected by Spatineo Monitor Very commonly used stable log format	No user agent No time taken
NCSA Combined	Auto-detected by Spatineo Monitor Very commonly used stable log format	No time taken
Tomcat AccessLogValve (defaults)	See NCSA Common	See NCSA Common
HAProxy (defaults)	When correctly configured, will contain all required information. Basic information auto-detected by Spatineo Monitor.	Cannot auto-detect user agent or referer. Needs manual configuration.
Customized Tomcat AccessLogValve (custom pattern) or HAProxy (configured headers)	When correctly configured, will contain all required information. You may specify custom log analysis fields.	No auto-detection: › Server and target folder configuration must be matched › All files in folder must share the same pattern May lack crucial log fields if configured wrong.

Appendix B - Supported key formats

OpenSSH

```
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAABAQCi4CjzNHFmAwaxtg2t3jrZCkM2pllcJsOHcEj+oKa8PsmXZUff
+sOgAae4mjsgKbmi67EYb1PFTfgdhh8+bomKpfv0OqOxGgI8IyZ6Wz8d8KBHSy3+4drMHcO7iyQgLt14
rIztfxP6lNCtxCPgI/W6JWhNb9Gt6gxhY4/eSNpSpiQCRwz9FP8RyFKXDuz60G2L1lFMKnrYqUmXqgqk
cNrBNz8sgFtoOmvI34+cP7xb81OeF1cQM5wyVfR3+fNLx4/E26+S09xpvxd2Qmu/7LIw9nfWZGPLTZoW
q0MkTEcSxG1hKeSpa5wTc/EIUbNsOFLGeto8kzoJLW5isFiniTsd test@spatineo.com
```

(single line without newlines)

RFC4716

```
--- BEGIN SSH2 PUBLIC KEY ----
Comment: This is my public key for use on \
servers which I don't like.
AAAAB3NzaC1kc3MAAACBAPY8ZOHY2yFSJA6XYC9HRwNHxaehvx5wOJ0rzZdzoSOXxbET
W6ToHv8D1UJ/z+zHo9Fiko5XybZnDIaBDHtblQ+Yp7StxyltHnXF1YLfKD1G4T6JYrdH
YI14Om1eg9e4NnCRleaqoZPF3UGfZia6bXrGTQf3gJq2e7Yisk/gF+1VAAAAFQDb8D5c
vwHWTZDPfX0D2s9Rd7NBvQAAAIEAlN92+Bb7D4KLYk3IwRbXblwXdkPggA4pfdtW9vGf
J0/RHd+NjB4eo1D+0dix6tXwYGN7PKS5R/FXPNwxHPapcj9uL1Jn2AWQ2dsknf+i/FAA
vioUPkmdMc0zuWoSOEsSNhVDtX3WdvVcGcBq9cetzrtOKWOocJmJ80qadxTRHtUAAACB
AN7CY+KKv1gHpRzFwdQm7HK9bb1LAo2KwaoXnadFgeptNBQeSXG1vO+JsvphVMBJc9HS
n24VYtYtsMu74qXviYjziVucWKjjKEb11juqnF0GDlB3VVmxHLmxnAz643WK42Z7dLM5
sY29ouezv4Xz2PuMch5VGPP+CDqzCM4loWgV
---- END SSH2 PUBLIC KEY ----
```

Example from http://www.ietf.org/rfc/rfc4716.txt

Appendix C – Common SSH clients and generating keys

OpenSSH

The private and public keys are created using the command line tool ssh-keygen. The following command will create a 2048 bit RSA key which will be written to files spatineo-monitor.rsa (private key) and spatineo-monitor.rsa.pub (public key)

```
ssh-keygen -b 2048 -t rsa -f /path/to/keys/spatineo-monitor.rsa
```

The file spatineo-monitor.rsa.pub contains the public key and it should be sent to Spatineo support to setup log transfer. Never expose the private key to anyone, not even Spatineo!

To use the key when connecting to the log upload service, use the -i flag to refer to the private key file.

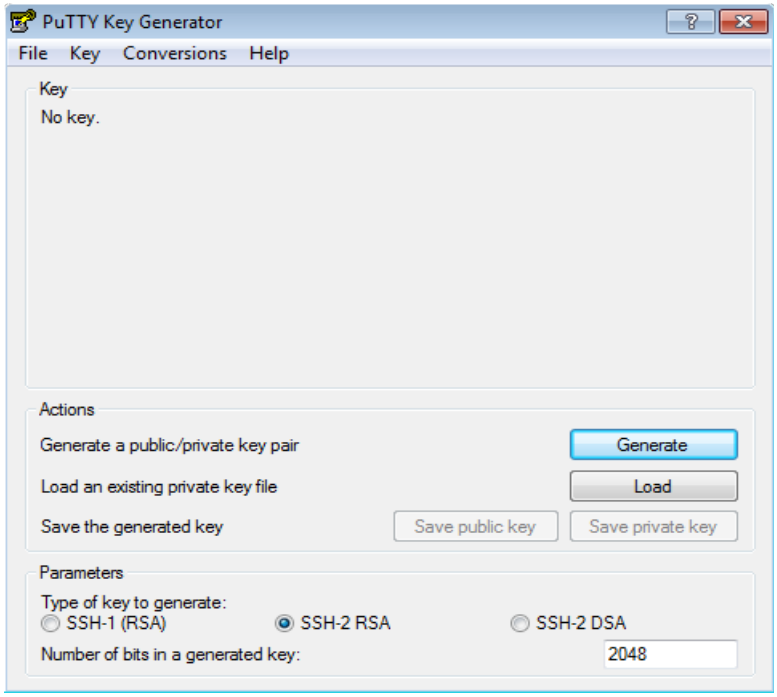
```
sftp -i /path/to/keys/spatineo-monitor.rsa username@upload.spatineo.com
```

WinSCP

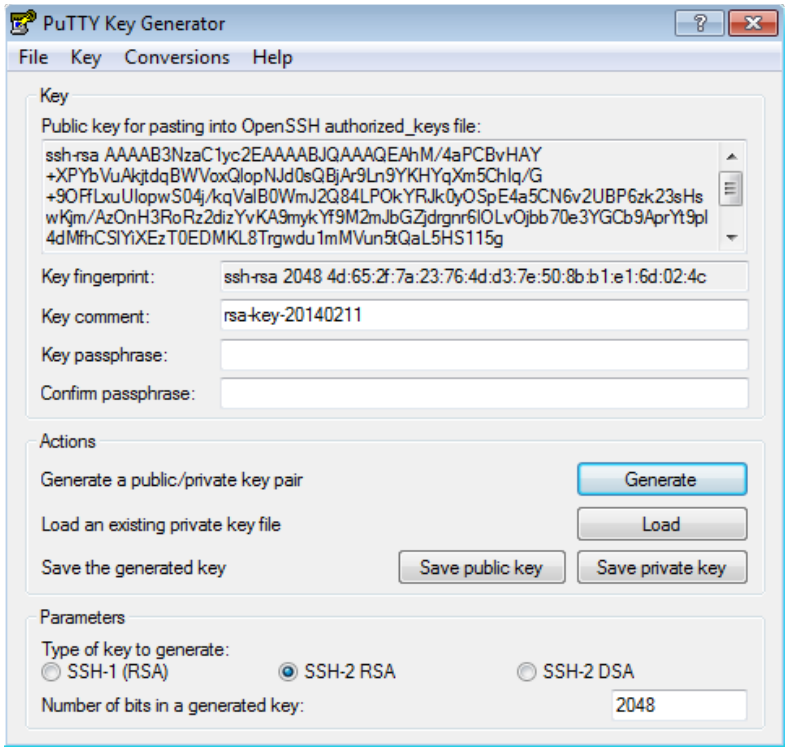
WinSCP is a popular tool for SFTP file transfers on Windows. It's open source (licensed under GPLv3 or later) and freely available at http://winscp.net/

To generate keys with WinSCP you may use the PuTTYgen tool that is bundled with WinSCP. You will find puttygen.exe inside the WinSCP installation, typically in the directory C:\Program Files\WinSCP\PuTTY

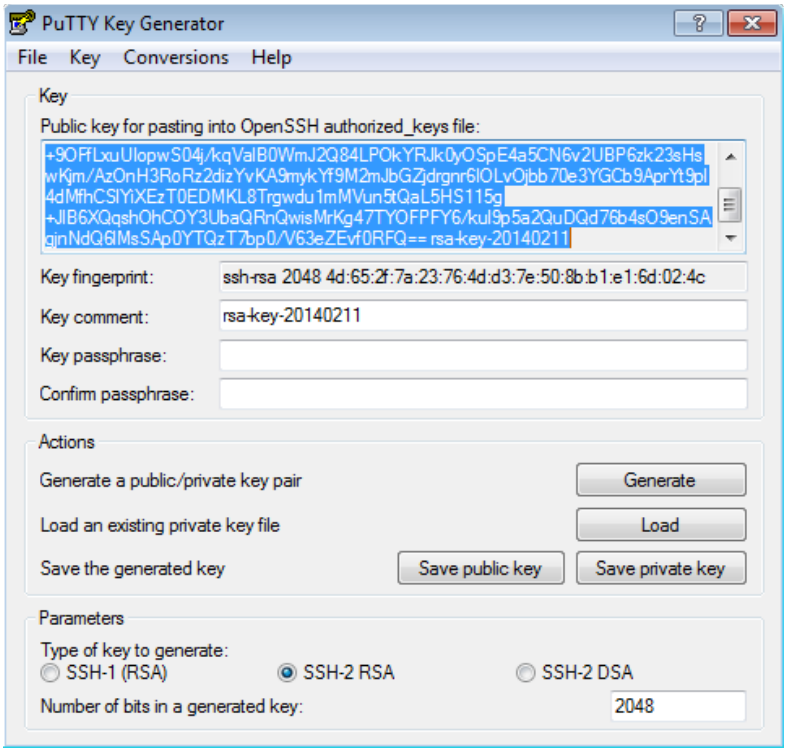
1. Start puttygen.exe



2. Click “generate”

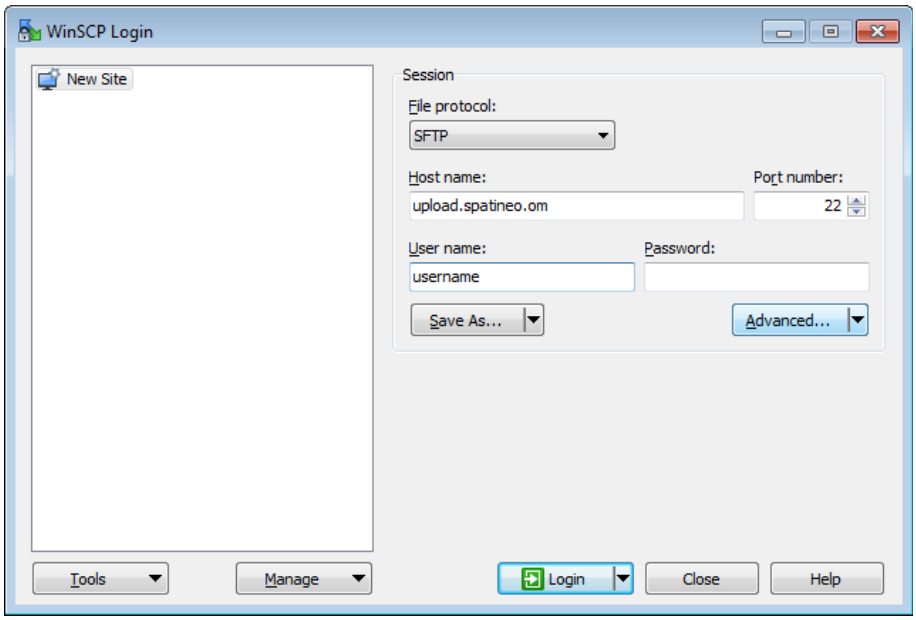


3. Copy public key text and save private key. Make sure you copy the entire key – there is typically more text that what fit in view at once.

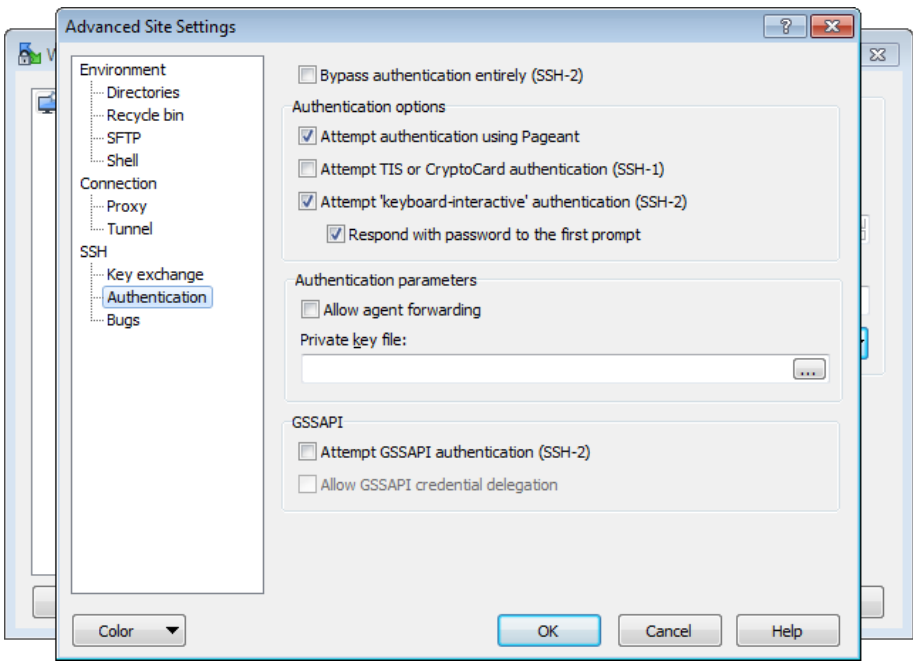


Send the public key copied in step 3 to Spatineo support. Do NOT send the private key. After you have received confirmation that they key is registered, you can start using WinSCP:

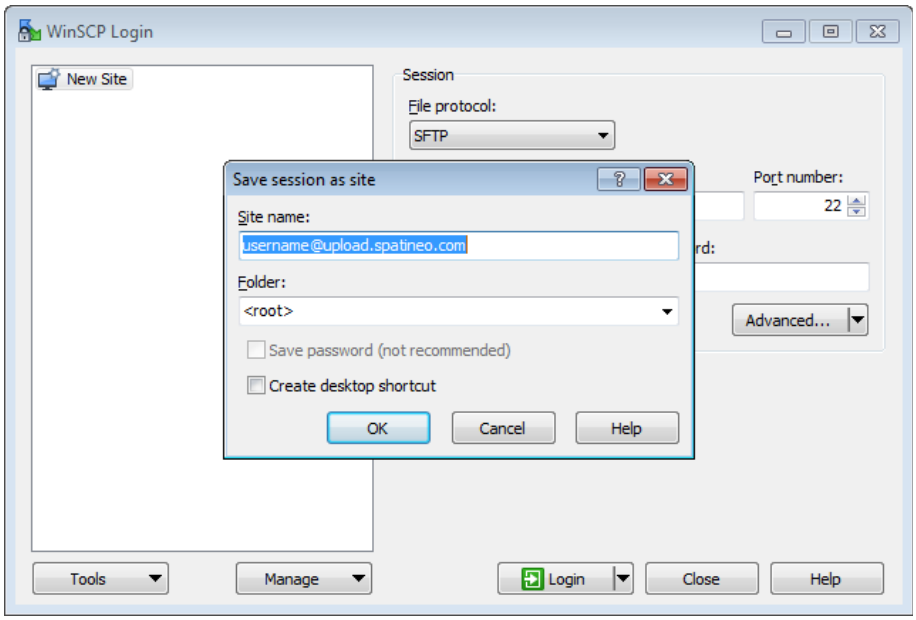
4. Open WinSCP and fill in the host and user names. Click advanced



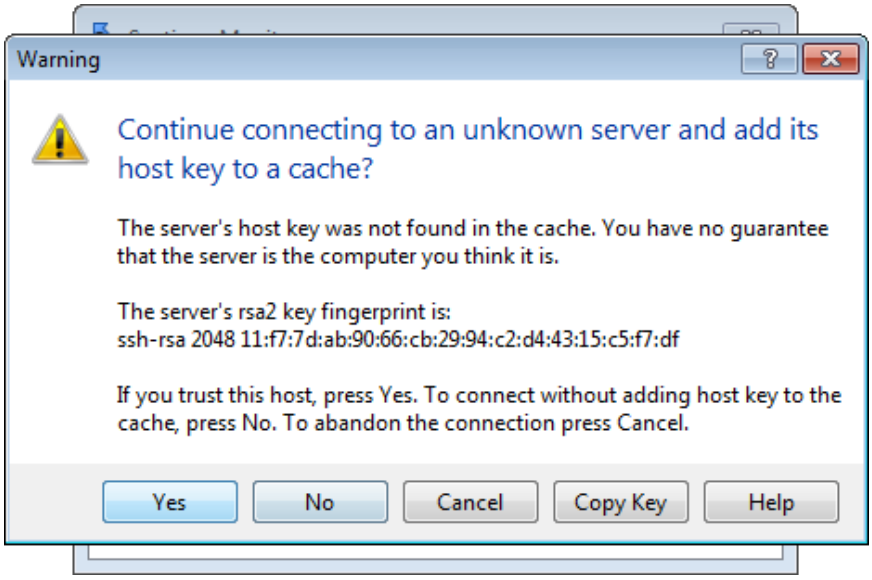
5. In SSH -> Authentication, choose the private key file you saved in step 3.



6. Save the site configuration



7. The first time you connect, you will need to confirm the Spatineo upload service key



After this, you will see the remote log folders and you will be able to transfer your log files into the folders.