Article
Iryna Mykhailova · Aug 21, 2021 9m read

Open Exchange

Transferring Files via REST to Store in a Property, Part 3

The first installment of this article series discussed how to read a big chunk of data from the raw body of an HTTP POST method and save it to a database as a stream property of a class. The second installment discussed how to send files and their names wrapped in a JSON format.

Now let ' s look closer at the idea of sending large files in parts to the server. There are several approaches we can use to do this. This article discusses using the Transfer-Encoding header to indicate chunked transfer.
The HTTP/1.1 specification introduced the Transfer-Encoding header, and the RFC 7230 section 4.1 described it, but it ' s absent from the HTTP/2 specification.

Transfer-Encoding Header

The objective of the Transfer-Encoding header is to specify the form of encoding used to transfer the payload body to the user safely. You use this header primarily to delimit a dynamically generated payload accurately and to distinguish payload encodings for transport efficiency or security from the characteristics of the selected resource.

You can use the following values in this header:

- Chunked
- Compress
- Deflate
- gzip

Transfer-Encoding Equals Chunked

When you set transfer encoding to chunked, the body of the message will consist of an unspecified number of regular chunks, a terminating chunk, a trailer part, and a final carriage return line feed (CRLF) sequence.

Each part starts with a chunk size represented by a hexadecimal number followed by an optional extension and CRLF. After that comes the body of the chunk with CRLF at the end of it. The extensions contain the metadata of the chunk. For example, metadata could include a signature, a hash, mid-message control information, and so on. The terminating chunk is a regular chunk with zero length. A trailer, which consists of (possibly empty) header fields, follows the terminating chunk.

To make it all easier to imagine, here is the structure of a message with Transfer-Encoding = chunked:

chunked_body	*chunk last_chunk trailer_part CRLF
chunk	chunk_size [chunk_ext] CRLF chunk_data CRLF
chunk_size	*size-of-current-chunk-in-HEX*
chunk_ext	*(";" chunk_ext_name ["=" chunk_ext_val])
chunk_ext_name	*token*
chunk_ext_val	*token / quoted-string*
chunk_data	*contents-of-current-chunk*
last_chunk	1*("0") [chunk_ext] CRLF
trailer_part	*(header_field CRLF)

An example of a short, chunked message looks like this:

```
13\r\n
Transferring Files \r\n
4\r\n
on\r\n
1A\r\n
community.intersystems.com
0\r\n
\r\n
```

This message body consists of three meaningful chunks. The first chunk has a length of nineteen octets, the second has four, and the third has twenty-six. You can see that the trailing CRLFs that mark the ends of the chunks don ' t count toward the chunk size. But, if you use CRLF as the end of line (EOL) marker, then the CRLF does count as a part of a message and takes two octets. The decoded message looks like this:

```
Transferring Files on
community.intersystems.com
```

Forming Chunked Messages in IRIS

For this tutorial, we ' ll use the method on the server created in the first article. This means that we are going to send the contents of the file directly to the body of the POST method. Since we are sending the contents of the file in the body, we send the POST to http://webserver/RestTransfer/file.

Now, let's look at how we can form a chunked message in IRIS. As specified in Sending HTTP Requests, under the section Sending a Chunked Request, you can send an HTTP request in chunks if you are using HTTP/1.1. The best part of this process is that %Net.HttpRequest automatically computes the content length of the entire message body on the server side so there is no need to change server side at all. Therefore, to send a chunked request, you need to follow these steps in the client only.

The first step is to create a subclass of %Net.ChunkedWriter and implement the OutputStream method. This method should get a stream of data, examine it, decide whether to split it into parts or not, how to split it, and

invoke the inherited methods of the class to write the output. In our case, we ' ll call the class RestTransfer.ChunkedWriter.

Next, in the client-side method responsible for sending data (called " SendFileChunked " here), you must create an instance of RestTransfer.ChunkedWriter class and fill it with the requested data you want to send. Since we are sending files, we ' ll do all the heavy lifting in the RestTransfer.ChunkedWriter class. We add a property named Filename As %String and a parameter named " MAXSIZEOFCHUNK = 10000. " Of course, you can decide to set a maximum allowed size for the chunk as a property and set it for each file or message.

Finally, set the EntityBody property of %Net.HttpRequest to be equal to the created instance of the RestTransfer.ChunkedWriter class and you ' re good to go.

These steps are just the new code you must write and replace in your existing method that sends files to a server.

The method looks like this:

```
ClassMethod SendFileChunked(aFileName) As %Status
{
  Set sc = $$$OK
  Set request = ..GetLink()
  set cw = ##class(RestTransfer.ChunkedWriter).%New()
  set cw.Filename = aFileName
  set request.EntityBody = cw
  set sc = request.Post("/RestTransfer/file")
  Quit:$System.Status.IsError(sc) sc
  Set response=request.HttpResponse
  do response.OutputToDevice()
  Quit sc
}
```

The %Net.ChunkedWriter class is an abstract stream class that provides an interface and has some implemented methods and properties. Here, we use the following property and methods:

- Property TranslateTable as %String forces automatic translation of the chunks when writing them into the output stream (EntityBody). We expect to receive raw data, so we must set TranslateTable to " RAW " .
- Method OutputStream is an abstract method overridden by a subclass to do all the chunking.

- Method WriteSingleChunk(buffer As %String) writes the Content-Length HTTP header followed by the entity-body as a single chunk. We check to see if the size of the file is smaller than the MAXSIZEOFCHUNK method, in which case, we use this method.
- Method WriteFirstChunk(buffer As %String) writes the Transfer-Encoding header followed by the first chunk. It should always be present. Zero or more calls to write more chunks may follow it, a compulsory call to write the last chunk with the empty string follows. We check that the length of the file is greater than the MAXSIZEOFCHUNK method and call this method.
- Method WriteChunk(buffer As %String) writes consequent chunks. Check to see if the rest of the file after the first chunk is still greater than MAXSIZEOFCHUNK then use this method to send data. We keep doing it until the size of the last part of the file is less than MAXSIZEOFCHUNK.
- Method WriteLastChunk(buffer As %String) writes the last chunk followed by a zero-length chunk to mark the end of the data.

Based on everything above, our class RestTransfer.ChunkedWriter looks like this:

```
Class RestTransfer.ChunkedWriter Extends %Net.ChunkedWriter
{
  Parameter MAXSIZEOFCHUNK = 10000;
  Property Filename As %String;
```

```
  Method OutputStream()
  {
    set ..TranslateTable = "RAW"
    set cTime = $zdatetime($Now(), 8, 1)
    set fStream = ##class(%Stream.FileBinary).%New()
    set fStream.Filename = ..Filename
    set size = fStream.Size
    if size < ..#MAXSIZEOFCHUNK {
      set buf = fStream.Read(.size, .st)
      if $$$ISERR(st)
      {
        THROW st
      } else {
        set ^log(cTime, ..Filename) = size
        do ..WriteSingleChunk(buf)
      }
    } else {
      set ^log(cTime, ..Filename, 0) = size
      set len = ..#MAXSIZEOFCHUNK
      set buf = fStream.Read(.len, .st)
      if $$$ISERR(st)
      {
        THROW st
      } else {
        set ^log(cTime, ..Filename, 1) = len
        do ..WriteFirstChunk(buf)
      }
      set i = 2
      While 'fStream.AtEnd {
        set len = ..#MAXSIZEOFCHUNK
        set temp = fStream.Read(.len, .sc)
    if len<..#MAXSIZEOFCHUNK
    {
      do ..WriteLastChunk(temp)
    } else {
          do ..WriteChunk(temp)
        }
        set ^log(cTime, ..Filename, i) = len
        set i = $increment(i)
      }
    }
  }
}
```

To see how these methods split the file into parts, we add a global ^log with the following structure:

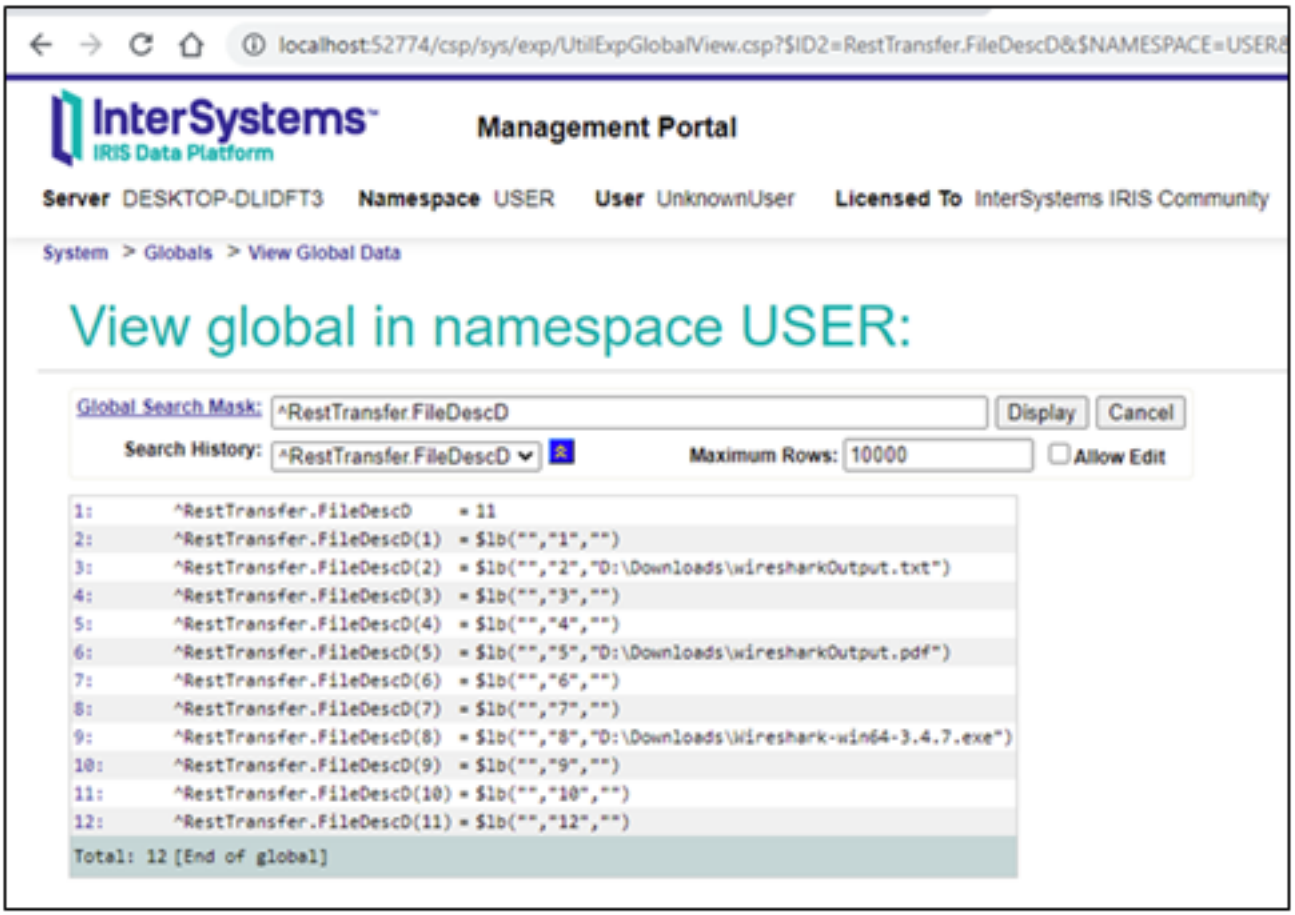
```
//for transfer in a single chunk
^log(time, filename) = size_of_the_file
//for transfer in several chunks
^log(time, filename, 0) = size_of_the_file
^log(time, filename, idx) = size_of_the_idx’s_chunk
```

Now that the programming is complete, let ’ s see how all three approaches work for different files. We write a simple class method to make calls to the server:

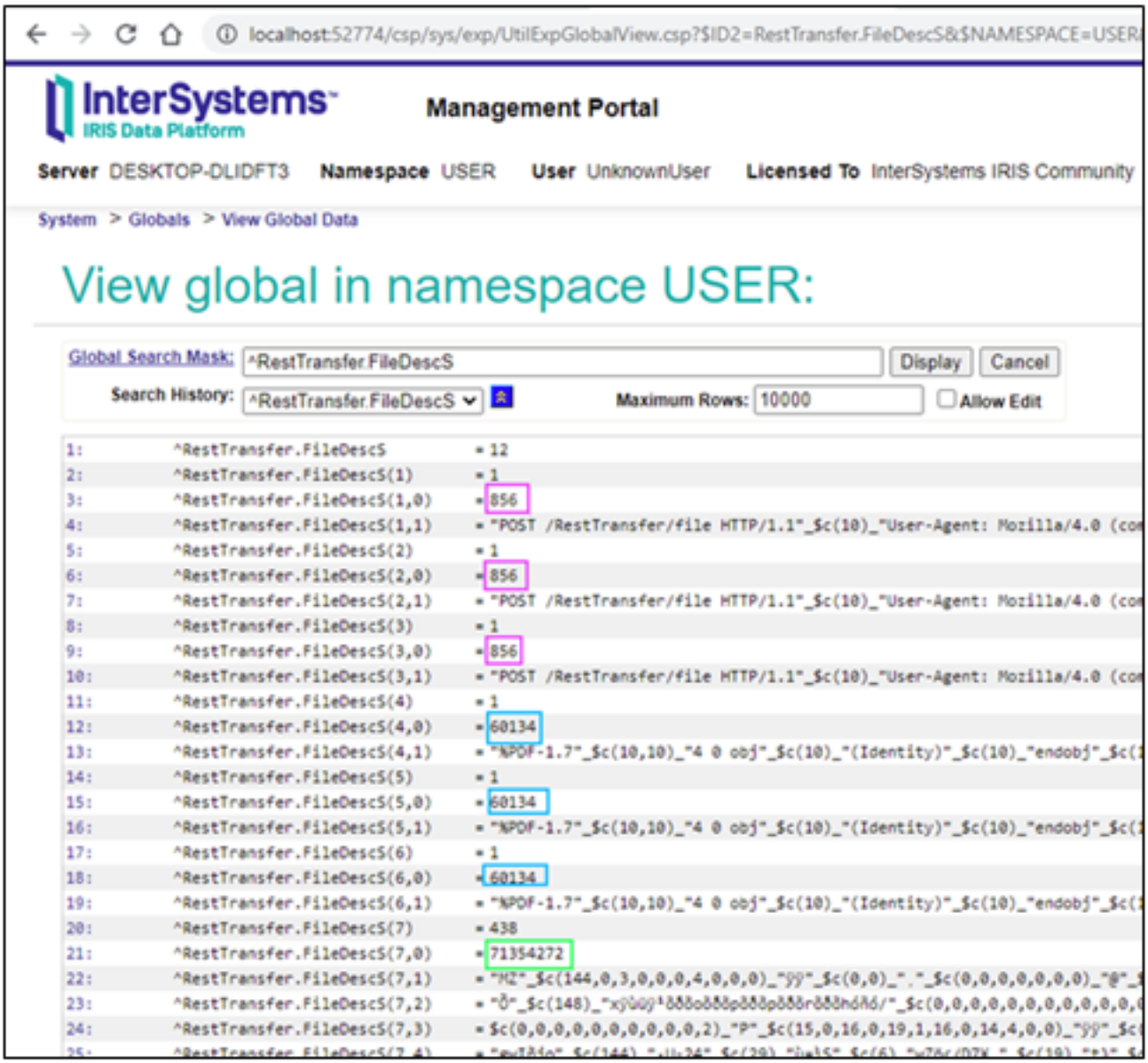
```
ClassMethod Run()
{
```

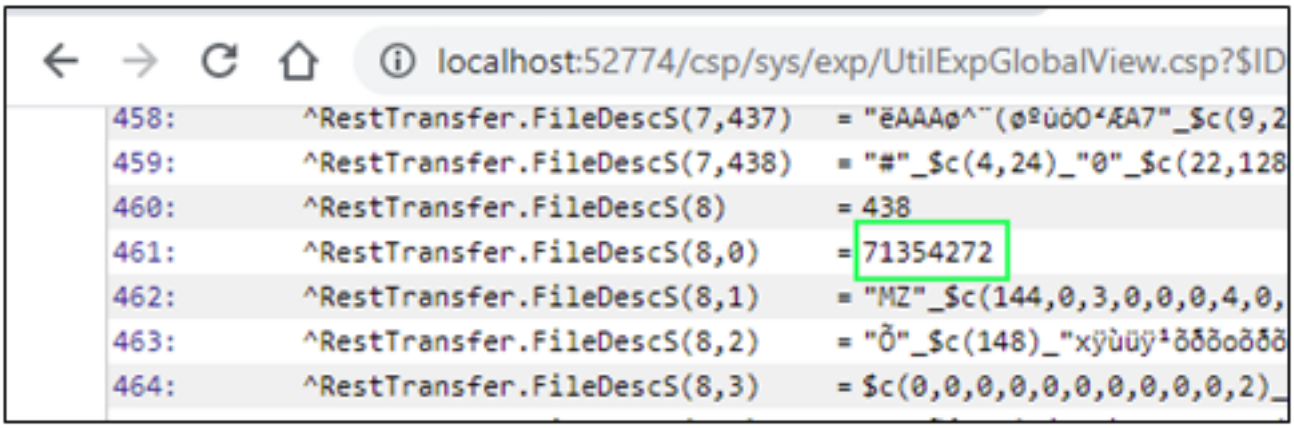
```
  // First, I am deleting globals.
  kill ^RestTransfer.FileDescD
  kill ^RestTransfer.FileDescS
  // Then I form a list of files I want to send
  for filename = "D:\Downloads\wiresharkOutput.txt", // 856 bytes
     "D:\Downloads\wiresharkOutput.pdf", // 60 134 bytes
     "D:\Downloads\Wireshark-win64-3.4.7.exe", // 71 354 272 bytes
     "D:\Downloads\IRIS_Community-2021.1.0.215.0-win_x64.exe" //542 370 224 bytes
  {
    write !, !, filename, !, !
    // And call all three methods of sending data to server side.
    set resp1=##class(RestTransfer.Client).SendFileChunked(filename)
    if $$$ISERR(resp1) do $System.OBJ.DisplayError(resp1)
    set resp1=##class(RestTransfer.Client).SendFile(filename)
    if $$$ISERR(resp1) do $System.OBJ.DisplayError(resp1)
    set resp1=##class(RestTransfer.Client).SendFileDirect(filename)
    if $$$ISERR(resp1) do $System.OBJ.DisplayError(resp1)
  }
}
```

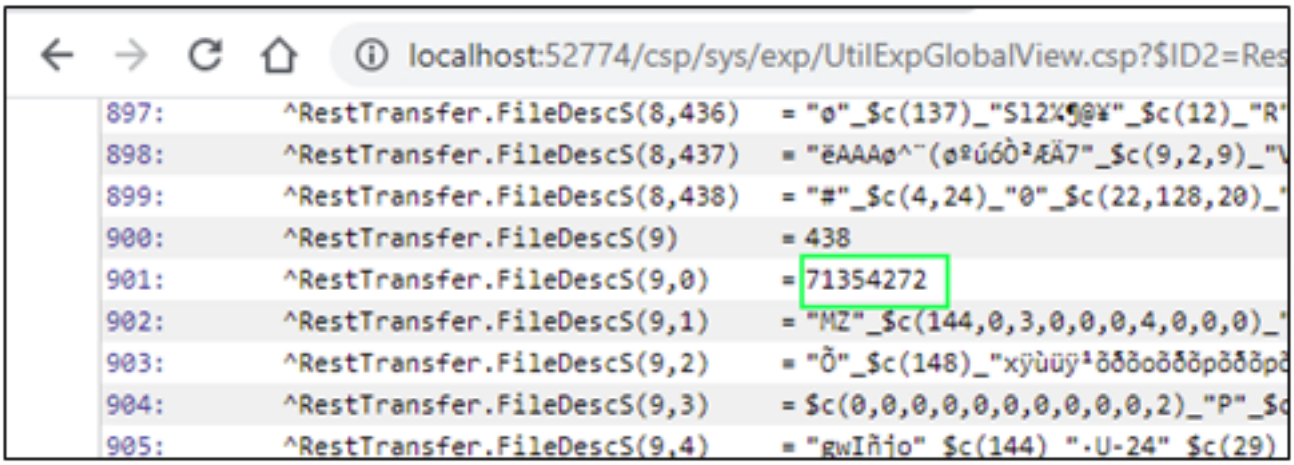
After running the class method Run, in the output for the first three files, the status was okay. But for the last file, while the first and last calls worked, the middle one returned an error: 5922, Timed out waiting for response. If we look in our globals method, we see that the code didn ’ t save the eleventh file. This means that ##class(RestTransfer.Client).SendFile(filename) failed — or to be precise, the method that unwraps data from JSON didn ’ t succeed.

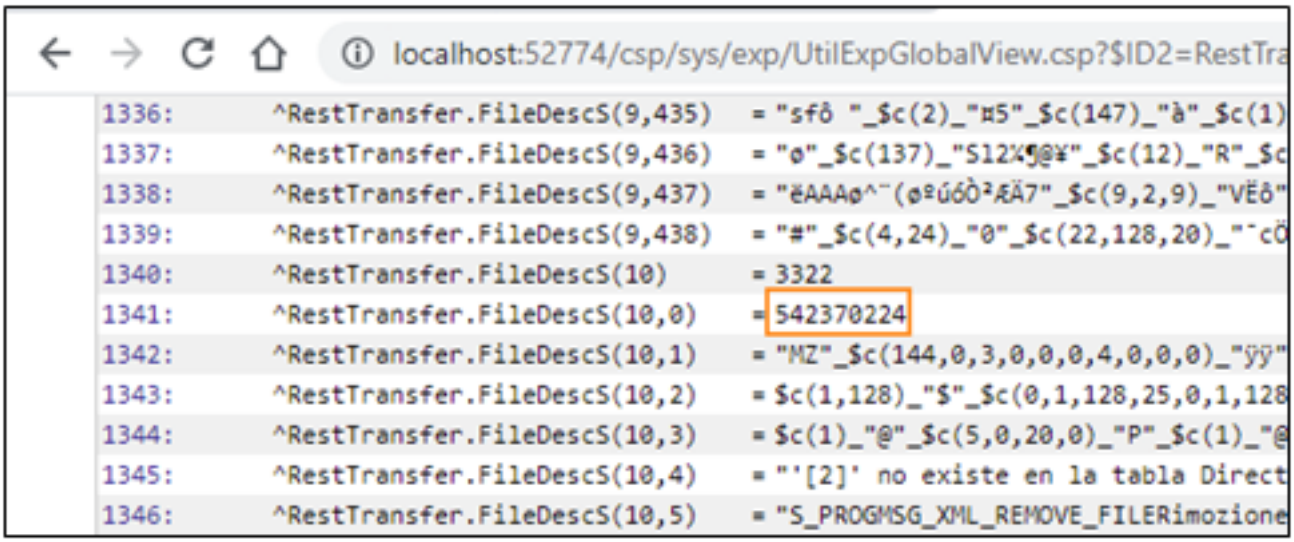


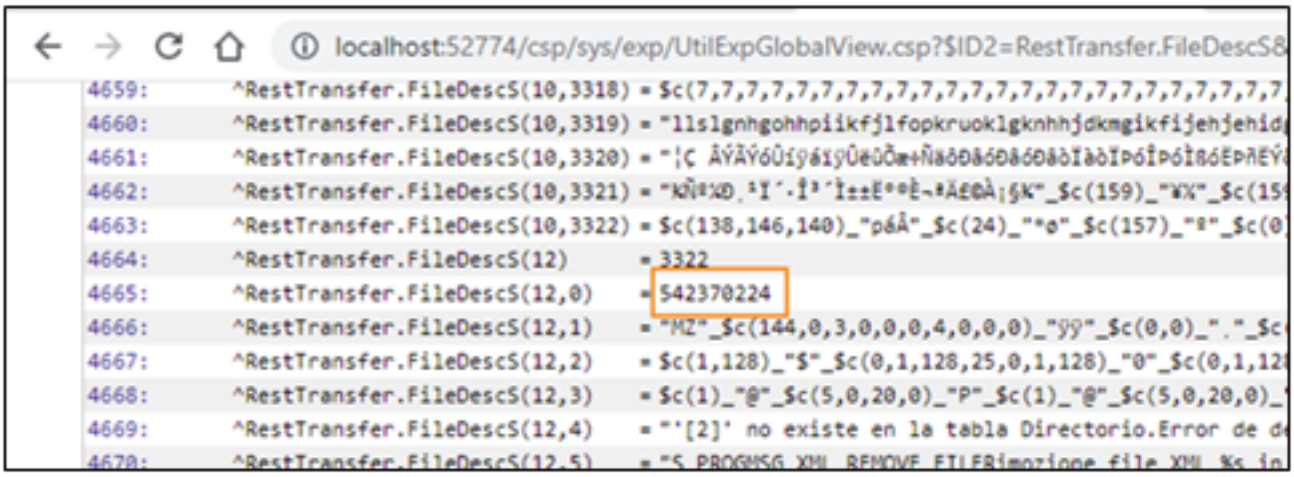
Now, if we look at our streams, we see that all the successfully saved files have the correct sizes.


localhost:52774/csp/sys/exp/UtilExpGlobalView.csp?$ID2=RestTransfer.FileDescS&$NAMESPACE=USER
InterSystems
IRIS Data Platform
Management Portal
Server DESKTOP-DLIDFT3 Namespace USER User UnknownUser Licensed To InterSystems IRIS Community
System > Globals > View Global Data
View global in namespace USER:
Global Search Mask: ^RestTransfer.FileDescS
Display
Cancel
Search History: ^RestTransfer.FileDescS
Maximum Rows: 10000
Allow Edit
1: ^RestTransfer.FileDescS = 12
2: ^RestTransfer.FileDescS(1) = 1
3: ^RestTransfer.FileDescS(1,0) = 856
4: ^RestTransfer.FileDescS(1,1) = "POST /RestTransfer/file HTTP/1.1"_$c(10)_"User-Agent: Mozilla/4.0 (co
5: ^RestTransfer.FileDescS(2) = 1
6: ^RestTransfer.FileDescS(2,0) = 856
7: ^RestTransfer.FileDescS(2,1) = "POST /RestTransfer/file HTTP/1.1"_$c(10)_"User-Agent: Mozilla/4.0 (co
8: ^RestTransfer.FileDescS(3) = 1
9: ^RestTransfer.FileDescS(3,0) = 856
10: ^RestTransfer.FileDescS(3,1) = "POST /RestTransfer/file HTTP/1.1"_$c(10)_"User-Agent: Mozilla/4.0 (co
11: ^RestTransfer.FileDescS(4) = 1
12: ^RestTransfer.FileDescS(4,0) = 60134
13: ^RestTransfer.FileDescS(4,1) = "%PDF-1.7"_$c(10,10)_"4 0 obj"_$c(10)_"(Identity)"_$c(10)_"endobj"_$c(
14: ^RestTransfer.FileDescS(5) = 1
15: ^RestTransfer.FileDescS(5,0) = 60134
16: ^RestTransfer.FileDescS(5,1) = "%PDF-1.7"_$c(10,10)_"4 0 obj"_$c(10)_"(Identity)"_$c(10)_"endobj"_$c(
17: ^RestTransfer.FileDescS(6) = 1
18: ^RestTransfer.FileDescS(6,0) = 60134
19: ^RestTransfer.FileDescS(6,1) = "%PDF-1.7"_$c(10,10)_"4 0 obj"_$c(10)_"(Identity)"_$c(10)_"endobj"_$c(
20: ^RestTransfer.FileDescS(7) = 438
21: ^RestTransfer.FileDescS(7,0) = 71354272
22: ^RestTransfer.FileDescS(7,1) = "MZ"_$c(144,0,3,0,0,0,4,0,0,0)_"ÿÿ"_$c(0,0)_"
23: ^RestTransfer.FileDescS(7,2) = "Õ"_$c(148)_"xÿùüÿ
24: ^RestTransfer.FileDescS(7,3) = $c(0,0,0,0,0,0,0,0,0,0,2)_"P"_$c(15,0,16,0,19,1,16,0,14,4,0,0)_"ÿÿ"_$c

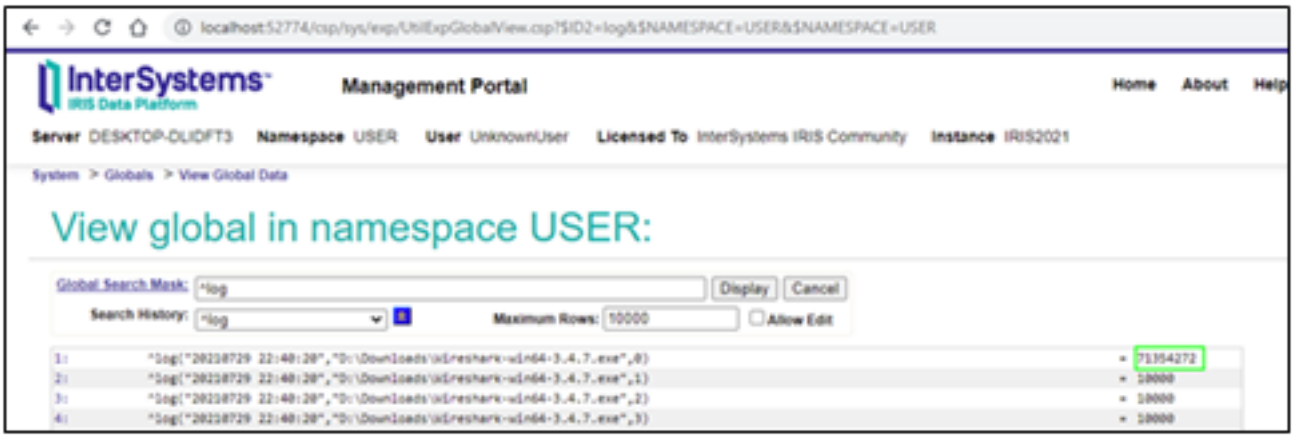

localhost:52774/csp/sys/exp/UtilExpGlobalView.csp?$ID
458: ^RestTransfer.FileDescS(7,437) = "ëAAAø^¨(ø
459: ^RestTransfer.FileDescS(7,438) = "#"_$c(4,24)_"0"_$c(22,128
460: ^RestTransfer.FileDescS(8) = 438
461: ^RestTransfer.FileDescS(8,0) = 71354272
462: ^RestTransfer.FileDescS(8,1) = "MZ"_$c(144,0,3,0,0,0,4,0,
463: ^RestTransfer.FileDescS(8,2) = "Õ"_$c(148)_"xÿùüÿ
464: ^RestTransfer.FileDescS(8,3) = $c(0,0,0,0,0,0,0,0,0,0,2)_

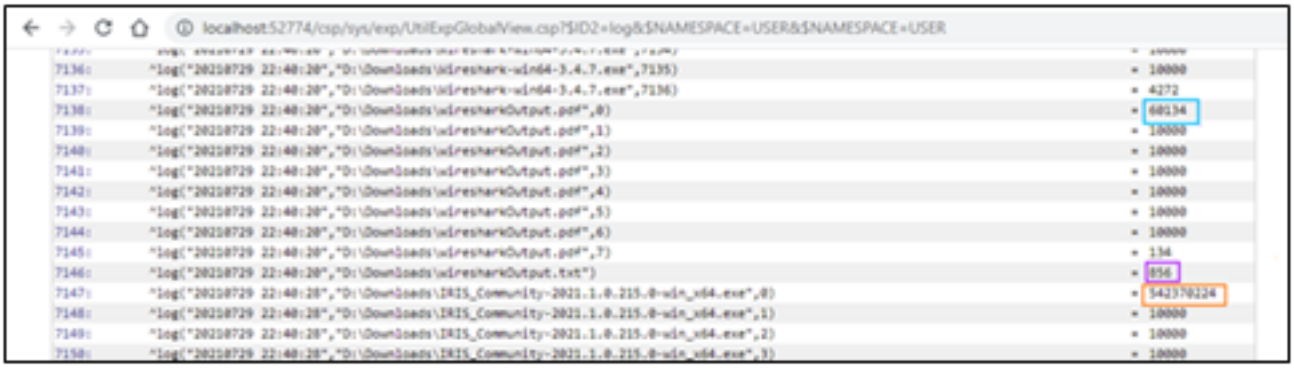






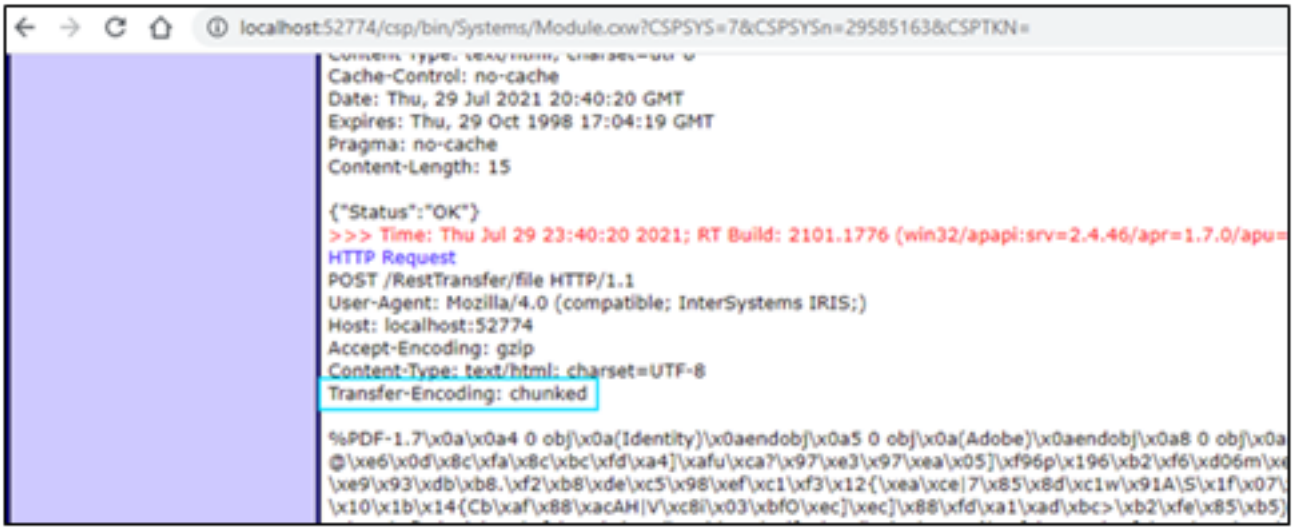
If we look at the ^log global, we see how many chunks the code created for each file:





You ' d probably like to see the bodies of the actual messages. Eduard Lebedyuk suggested in the article Debugging Web that it ' s possible to use CSP Gateway Logging and Tracing.

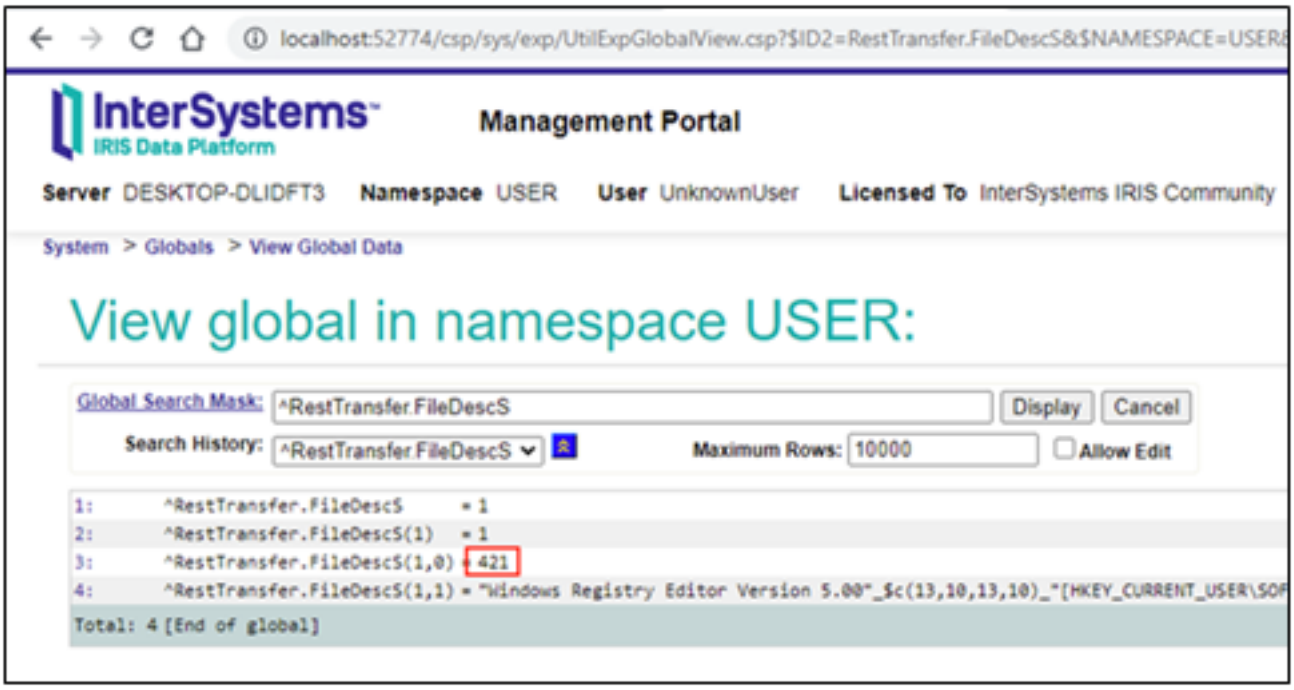
If we look in the Event Log for the second chunked file, we see that the value of the Transfer-Encoding header is indeed " chunked. " Unfortunately, the server has already glued the message together, so we don ' t see the actual chunking.



Using the Trace feature doesn ' t show a lot more information, but it clarifies that there is a gap between the penultimate and the last request.

To see the actual parts of the messages, we copy the client to another computer to use a network sniffer. Here we ' ve chosen to use Wireshark because it is free and it has the necessary functions. To better show you how the code splits the file into chunks, we can change the value of MAXSIZEOFCHUNK to 100 and chose to send a small file. So now, we can see the following result:

We see that the lengths of all but the last two chunks equal 64 in HEX (100 in DEC), the final chunk with data equals 21 DEC (15 in HEX), and we can see the size of the last chunk is zero. Everything looks OK and accords with the specification. The overall length of the file equals 421 (4x100+1x21), which we can also see in globals:



Wrapping Up

Overall, we can see that this approach works and enables sending large files without problems to the server. Additionally, if you ' re sending large amounts of data to a client, you might want to familiarize yourself with the Web Gateway Operation and Configuration, section Application Path Configuration Parameters, parameter Response Size Notification. It specifies Web Gateway behavior when sending large amounts of data depending on the version of HTTP used.

The code for this approach is added to the previous version of this example on GitHub and InterSystems Open Exchange.

While on the topic of sending files in chunks, it is also possible to use the Content-Range header with or without the Transfer-Encoding header to indicate which exact part of the data is being transferred. Furthermore, you can use a completely new concept of streams available with the HTTP/2 specification.

As always, if you have any questions or suggestions, please don ' t hesitate to write them in the comments section.

#REST API #InterSystems IRIS
Check the related application on InterSystems Open Exchange

Source URL:https://community.intersystems.com/post/transferring-files-rest-store-property-part-3