Title: METHOD AND APPARATUS FOR INCREASING STABILITY IN A CLIENT- SERVER COMMUNICATION USING HTTP

Abstract: A method is presented for communicating between a first communication device (300) and a second communication device (310). The method comprises the steps of the first communication device sending a first message (342) to the second communication device over a connection between the first communication device and the second communication device at a first message sending time (351), the second communication device receiving the first message from the first communication device at a first message receiving time (352), and the second communication device sending a second message (344) to the first communication device, the second message being sent at a second message sending time (353) being at least a minimum response time (341) after the first message receiving time. A corresponding server, mobile communication terminal and computer program products are also presented.
before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
METHOD AND APPARATUS FOR INCREASING STABILITY IN A CLIENT-SERVER COMMUNICATION USING HTTP.

Field of the Invention

The present invention relates to data communication, and more particularly to improving stability of data communication.

Background of the Invention

Most data communication protocols in use today, such as Hypertext Transfer Protocol (HTTP), have been designed for use with wire-based communication. When protocols like these are subsequently used in new, originally unintended environments, such as mobile communication networks, problems arise. Due to the unpredictability of communicating with mobile terminals over radio, data frames are more often likely to be lost. Additionally, latency is higher in mobile communication, due to using appropriate extensive coding. More significantly, stability of communication using standard HTTP has proven to be unsatisfactory in practice.

A suite of protocols, known as Wireless Application Protocol (WAP), was designed, among other purposes, to address the problem of stability. While stability was improved with WAP, WAP had other issues associated with it, such as requiring a gateway to translate WAP protocols to more common protocols such as HTTP and Transmission Control Protocol (TCP).

There is thus a need for a new way to increase stability when performing data communication with a mobile terminal.

Summary of the Invention

In view of the above, an objective of the invention is to solve or at least reduce the problems discussed above. Generally, this objective has been achieved by
methods, devices and computer program products according to the enclosed independent patent claims.

According to a first aspect of the present invention, there is presented a method for communicating between a first communication device and a second communication device, comprising the steps of: the first communication device sending a first message to the second communication device over a connection between the first communication device and the second communication device at a first message sending time, the second communication device receiving the first message from the first communication device at a first message receiving time, and the second communication device sending a second message to the first communication device, the second message being sent at a second message sending time being at least a minimum response time after the first message receiving time.

The method may comprise the further step, after the step of the second communication device sending a second message to the first communication device, of: the first communication device closing the connection.

The step of the first communication device closing the connection may involve closing the connection at a connection closing time being no more than a maximum connection time after the first message sending time.

The connection may be a Hypertext Transfer Protocol (HTTP) connection. Thus, the step of the first communication device sending a first message to the second communication device may involve sending a Hypertext Transfer Protocol request. Correspondingly, the step of the second communication device sending a second message to the first communication device may involve sending a Hypertext Transfer Protocol response.

The first communication device may be a mobile communication terminal comprising a wireless interface to a mobile communication network, and the second communication device may be a server coupled to a data communica-
tion network, the data communication network being connected, directly or indirectly, to the mobile communication network.

A second aspect of the invention is a method performed by a second communication device, such as a server, for communicating with a first communication device, such as a mobile communication terminal, the method comprising the steps of: receiving a first message from the first communication device at a first message receiving time, and sending a second message to the first communication device, the second message being sent at a second message sending time being at least a minimum response time after the first message receiving time.

A third aspect of the invention is a method performed by a first communication device, such as a mobile communication terminal, for communicating with a second communication device, such as a server, the method comprising the steps of: sending a first message to the second communication device over a connection between the first communication device and the second communication device at a first message sending time, receiving a second message from the second communication device, and closing the connection at a connection closing time being no more than a maximum connection time after the first message sending time.

A fourth aspect of the invention is a server capable of communicating with a first communication device, comprising means for performing the method according to the second aspect.

A fifth aspect of the invention is a mobile communication terminal capable of communicating with a second communication device, comprising means for performing the method according to the third aspect.

The mobile communication terminal may comprise a communication manager client providing a Java application programming interface to a Java application.
The communication manager client may at least partly be implemented in Java.

A sixth aspect of the invention is a computer program product directly loadable into a memory of a controller of a server, where the computer program product comprises program code to be executed in the controller for performing the method according to the second aspect.

A seventh aspect of the invention is a computer program product directly loadable into a memory of a controller of a mobile communication terminal, where the computer program product comprises program code to be executed in the controller for performing the method according to the third aspect.

The present invention thus provides a more stable way of communicating, particularly in situations where mobile communication is involved. A risk of the software crashing due to unsynchronized communication is thereby reduced, which reduces end user frustration when using data communication.

Other objectives, features and advantages of the present invention will appear from the following detailed disclosure, from the attached dependent claims as well as from the drawings.

Generally, all terms used in the claims are to be interpreted according to their ordinary meaning in the technical field, unless explicitly defined otherwise herein. All references to "a/an/the [element, device, component, means, step, etc.]" are to be interpreted openly as referring to at least one instance of said element, device, component, means, step, etc., unless explicitly stated otherwise. The steps of any method disclosed herein do not have to be performed in the exact order disclosed, unless explicitly stated.
Brief Description of the Drawings

The present invention will now be described in more detail, reference being made to the enclosed drawings.

Fig 1 is a diagram showing an environment where an embodiment of the present invention can be implemented.

Fig 2 is a diagram showing the upper parts of communication stacks in a client and a server according to an embodiment of the present invention.

Fig 3 is a sequence diagram showing communication between a client and a server according to an embodiment of the present invention.

Detailed Description of an Embodiment

Fig 1 is a diagram showing an environment where an embodiment of the present invention can be implemented. A user 101 is interfacing with a mobile terminal 100 using an input device such as keypad 105 or a touch sensitive screen (not shown). Output from the mobile terminal 100 to the user 101 is effected by means of a display 103, speaker (not shown), vibrator (not shown), etc. The mobile terminal also comprises a controller, such as a central processing unit, microcontroller, etc, (not shown), and memory, such as RAM memory, ROM memory, EEPROM memory, flash memory, hard drive memory, optical memory or any combination thereof (not shown), capable of storing software instructions to be executed by the controller. The mobile terminal 100 has a wireless interface for communication over a wireless connection 102 with a base station 104 connected to a mobile communication network 106 in accordance with a mobile communication protocol. The mobile communication protocol can be any standardized or proprietary protocol, including but not limited to GSM, UMTS, CDMA2000, FOMA, PDC, TD-SCDMA, HSPDA and HSUPA. The mobile communication network 106 is connected to a wide area data communication network 108, such as the Internet, and a Public Switched Telephone Network (PSTN) (not shown). Connected
to the Internet 108 is a server 110 and a client computer 112. By means of this configuration, the mobile terminal 100 is capable of communicating with the server 110.

The server 110 comprises a controller 113, such as a central processing unit, and a memory 114, such as RAM memory, ROM memory, EEPROM memory, flash memory, hard drive memory, optical memory, or any combination thereof. The memory 114 is used by the controller 113 for several purposes, among them to store software instructions, such as an operating system and server applications, to be executed by the controller 113. The server 110 may run any of a plurality of operating systems, such as UNIX, Linux, BSD, FreeBSD, Windows Server, Mac OS X Server, etc.

Fig 2 is a diagram showing the upper parts of communication stacks in a client and a server according to an embodiment of the present invention. The left side of the diagram shows the upper part of a communication stack for a mobile communication terminal 200, such as mobile communication terminal 100 of Fig 1, while the right side of the diagram shows a corresponding part of a communication stack for the server 210, such as server 110 of Fig 1. At the uppermost level in the communication stack of the mobile terminal 200, a Java application 220 is running and wishes to communicate with a server application 226 at the uppermost level in the communication stack of the server 110. The Java application 220 accesses a communication manager client 222 at a lower level in the communication stack of the mobile terminal 200 through a Java application programming interface (API) which may for example be implemented as an interface which is similar to a TCP socket. The communication manager client 222 may be entirely or partly implemented in Java as well. This brings the advantage of allowing simple network access for any Java program when the communication manager client 222 is included.
The communication manager client 222 in turn uses an HTTP API with an HTTP client 224 at yet a lower level in the communication stack of the mobile terminal 200 to implement the communication with the server 210. In a Java implementation, the HTTP client 224 may be implemented in a Kilobyte Virtual Machine (KVM) which provides a run-time environment for all Java applications in the mobile terminal 200. The KVM may support the specification known as Mobile Information Device Profile 1 (MIDP 1) or Mobile Information Device Profile 2 (MIDP 2), where MIDP 2 is a superset of MIDP 1. MIDP 1 only requires networking in the form of HTTP, while MIDP 2 also mandates support for TCP sockets. As many mobile terminals support only MIDP 1, the communication manager client 222 only uses API's which are mandated by MIDP 1. Consequently, the communication manager client 222 works with any mobile terminal compliant with either MIDP 1 or MIDP 2 and presumably with later released versions of MIDP.

The HTTP client 224 utilizes lower network, link and physical layers, as is known in the art, to connect to a HTTP server 230 layer of the server 210. A communication manager server 228 in the server 210 corresponds to the communication manager client 222 in the mobile terminal 200 and provides communication with the server application 226. Thus, a network interface between a client application 220 and a server application 226 is provided. The server application is here shown to be in the same location (i.e. the same physical server computer) as the communication manager server 228, but this is not necessarily the case. The communication manager client 222 may indicate to the communication manager server 228 that the intended server application is located on a server with a specific uniform resource locator (URL) or IP address.

Fig 3 is a sequence diagram showing communication between a client and a server according to an embodiment
of the present invention. The entities performing the communication in the embodiment of Fig 3 are the communication manager client 222 and the communication manager server 228 of Fig 2. To initiate communication, the mobile communication terminal 300 sends a request 342 at a request sending time 351 (referred to as first message sending time in other sections of this document) to the server 310, which receives the request at a request receiving time 352 (first message receiving time). In the disclosed embodiment, this is a HTTP request, whereby a connection between the mobile communication terminal to the server 310 is opened before the request 342 is sent. Upon receiving the request 342, the server 310 performs necessary processing and sends a response 344 at a response sending time 353 (second message sending time) to the mobile communication terminal 300, which receives the response at a response receiving time 354. However, regardless of how quickly the server performs the necessary processing, the server 310 always makes sure that at least a specific amount of time has passed since the server 310 received the request. This minimum response time 341 enables the mobile communication terminal 300 to perform any processing required after sending the request 342, making the communication more stable. While the minimum response time 341 may be configured to be any value, it has been found in practice for the disclosed embodiment that a value between 1 and 3 seconds is a good balance between stability and latency. A value of about 2 seconds has been found to be particularly well balanced.

Subsequently, the mobile communication terminal 300 sends a close command 346 at a connection closing time 355 to terminate the connection with the server 310. Counting from the time 351 of the request 342, there is a maximum connection time 345, after which the connection needs to be closed for stability reasons. While the maximum connection time 345 may be configured to be any
value, it has been found in practice for the disclosed embodiment that a good value is between 3 and 7 seconds, with a particularly good time being about 5 seconds.

Each subsequent communication from the mobile communication terminal 300 to the server 310 requires a new connection, using a new request 348, a new response and a new close in accordance with the previously described connection.

It is to be noted that the present invention is not dependent on absolute times. In other words, neither the client (e.g. terminal 300) nor the server (310) needs to be aware of the timings of the other party, allowing for asynchronous communication in accordance with HTTP.

Furthermore, it is to be noted that when Java is mentioned herein, it is for mere illustrative purposes; the invention is not limited to any particular programming language and may for example be implemented in C, C++, C#, Assembler code, etc.

The invention has mainly been described above with reference to an embodiment. However, as is readily appreciated by a person skilled in the art, other embodiments than the one disclosed above are equally possible within the scope of the invention, as defined by the appended patent claims.
Claims

1. A method for communicating between a first communication device (300) and a second communication device (310), the method comprising the steps of:

- said first communication device sending a first message (342) to said second communication device over a connection between said first communication device and said second communication device at a first message sending time (351),

- said second communication device receiving said first message from said first communication device at a first message receiving time (352), and

- said second communication device sending a second message (344) to said first communication device, said second message being sent at a second message sending time (353) being at least a minimum response time (341) after said first message receiving time.

2. The method according to claim 1, comprising the further step, after said step of said second communication device (310) sending a second message (344) to said first communication device (300), of:

- said first communication device closing (346) said connection.

3. The method according to claim 2, wherein said step of said first communication device (300) closing said connection involves closing said connection at a connection closing time (355) being no more than a maximum connection time (345) after said first message sending time (351).

4. The method according to any one of the previous claims, wherein said minimum response time (341) is between 1 second and 3 seconds.
5. The method according to any one of the previous claims, wherein said minimum response time (341) is about 2 seconds.

6. The method according to any one of the previous claims, wherein said maximum connection time (345) is between 3 seconds and 7 seconds.

7. The method according to any one of the previous claims, wherein said maximum connection time (345) is about 5 seconds.

8. The method according to any one of the previous claims, wherein said connection is a Hypertext Transfer Protocol (HTTP) connection.

9. The method according to claim 8, wherein said step of said first communication device (300) sending a first message (342) to said second communication device (310) involves sending a Hypertext Transfer Protocol request (342).

10. The method according to claim 8 or 9, wherein said step of said second communication device (310) sending a second message (344) to said first communication device (300) involves sending a Hypertext Transfer Protocol response (344).

11. The method according to any one of the previous claims, wherein said first communication device is a mobile communication terminal (100, 200, 300) comprising a wireless interface to a mobile communication network (106), and said second communication device is a server (110, 210, 310) coupled to a data communication network (108), said data communication network being connected, directly or indirectly, to said mobile communication network.
12. A method performed by a second communication device (310), for communicating with a first communication device (300), the method comprising the steps of:

- receiving a first message (342) from said first communication device at a first message receiving time (352), and
- sending a second message (344) to said first communication device, said second message being sent at a second message sending time (353) being at least a minimum response time (341) after said first message receiving time.

13. A method performed by a first communication device (300), for communicating with a second communication device (310), the method comprising the steps of:

- sending a first message (342) to said second communication device over a connection between said first communication device and said second communication device at a first message sending time (351),
- receiving a second message (344) from said second communication device, and
- closing (346) said connection at a connection closing time (355) being no more than a maximum connection time (345) after said first message sending time.

14. A server (110, 210, 310) capable of communicating with a first communication device, comprising means for performing the method according to claim 12.

15. A mobile communication terminal (100, 200, 300) capable of communicating with a second communication device, comprising means for performing the method according to claim 13.

16. The mobile communication terminal (100, 200, 300) according to claim 15, comprising a communication
13
manager client (222) providing a Java application
programming interface to a Java application (220).

17. The mobile communication terminal (100, 200, 300) according to claim 16, wherein said communication
manager client (222) is at least partly implemented in Java.

18. A computer program product directly loadable
into a memory (114) of a controller (113) of a server (110, 210, 310), where said computer program product
comprises program code to be executed in said controller (113) for performing the method according to claim 12.

19. A computer program product directly loadable
into a memory of a controller of a mobile communication
terminal (100, 200, 300), where said computer program
product comprises program code to be executed in said
controller for performing the method according to claim
13.
Fig 1
Fig 2
Fig 3
INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE2007/000117

A. CLASSIFICATION OF SUBJECT MATTER

In/Pt/Cod: 70EC/70ent/International classification
A 1 1 F 9 C (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L, H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE, DK, FI, NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI, DATA, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 20060026169 A1 (PASQUA, R D), 2 February 2006 (02.02.2006), paragraphs 0033-0035; 0040-0043; 0062; 0075-0088, claims 1-8, abstract</td>
<td>1-19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. X See patent family annex.

* Special categories of cited documents:
  "A" document defining the general state of the art which is not considered to be of particular relevance
  "E" earlier application or patent but published on or after the international filing date
  "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  "O" document referring to an oral disclosure, use, exhibition or other means
  "P" document published prior to the international filing date but later than the priority date claimed
  "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
  "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
  "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
  "&" document member of the same patent family

Date of the actual completion of the international search: 29 June 2007

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. +46 8 666 02 86

Authorized officer
Rune Bengtsson /LR
Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (April 2007)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 20040111516 A1 (CAIN, S), 10 June 2004 (10.06.2004), claims 1-8, abstract</td>
<td>1-19</td>
</tr>
</tbody>
</table>
International patent classification (IPC)
H04L 29/06 (2006.01)
H04L 29/08 (2006.01)

Download your patent documents at www.prv.se
The cited patent documents can be downloaded at www.pirv.se by following the links:
  • In English/Searches and advisory services/Cited documents (service in English) or
  • e-tjanster/anförda dokument (service in Swedish).
Use the application number as username.
The password is BUTQLQIFYD.

Paper copies can be ordered at a cost of 50 SEK per copy from PRV InterPat (telephone number 08-782 28 85).

Cited literature, if any, will be enclosed in paper form.
US 2006026169 A1 02/02/2006
AU 2003301781 A 00/00/0000
EP 1561163 A 10/08/2005
WO 2004042571 A 21/05/2004

US 2004011516 A1 10/06/2004 NONE

Form PCT/ISA/210 (patent family annex) (April 2005)