Neutral Citation Number: [2018] EWCA Civ 2237

Case No: A3/2016/4537

IN THE COURT OF APPEAL (CIVIL DIVISION)
ON APPEAL FROM THE HIGH COURT OF JUSTICE
CHANCERY DIVISION, PATENTS COURT
Mr Roger Wyand QC, sitting as a Deputy High Court Judge
[2016] EWHC 2359 (Ch)

Royal Courts of Justice
Strand, London, WC2A 2LL

Date: 11/10/2018

Before:

LORD JUSTICE PATTEN
LORD KITCHIN
and
LORD JUSTICE FLOYD

- -

Between :

SSH COMMUNICATIONS SECURITY CORPORATION	**Appellant**
- and –	
SONY MOBILE COMMUNICATIONS AB and others	**Respondents**

- -
- -

Iain Purvis QC and Brian Nicholson (instructed by **Gowling WLG (UK) LLP**) for the **Appellant**
The **Respondents** were not present and were not represented

Hearing date: 24th July 2018

- -

Judgment Approved

Lord Kitchin:

1. This an appeal by SSH Communications Security Corporation ("SSH") against the decision of Mr Roger Wyand QC, sitting as a deputy judge of the High Court, which was handed down on 10 October 2016 ([2016] EWHC 2359).

2. SSH is the proprietor of European Patent EP (UK) 2,254,311 ("the patent") which is concerned with communications between computers in a data network and, in particular, computers connected to the Internet. It has a priority date of 15 June 1999, a time when the Internet was in its infancy.

3. In these proceedings three companies in the Sony group of companies (collectively "Sony") sought revocation of the patent on a series of grounds including, so far as relevant to this appeal, lack of novelty in light of two publications referred to as the "NAT Minutes" (or the "Minutes") and the "NAT Guidelines" (or the "Guidelines"), and obviousness in light of these publications and the common general knowledge. SSH responded with a counterclaim for infringement of the patent by Sony's Xperia mobile telephones.

4. The action came on for trial in July 2016 and lasted for four days. The deputy judge heard evidence from two experts: Professor Leslie, an academic in the Computer Science Department at Cambridge University, for Sony, and Mr Holdrege, a network engineer, for SSH. The deputy judge found the evidence of Professor Leslie balanced and helpful. Mr Holdrege's evidence was less satisfactory and so, where there was a direct conflict between them, the deputy judge preferred the evidence of Professor Leslie.

5. The deputy judge rejected the attack of obviousness over the common general knowledge but held the patent lacked novelty or was obvious over the NAT Minutes and Guidelines. He also found that the patent, if valid, would have been infringed.

6. On 5 February 2018 Sony and SSH entered into a worldwide confidential settlement and as a result Sony has played no part in this appeal and has not been represented. SSH pursues the appeal only for the purpose of overturning the deputy judge's finding that the patent is invalid.

7. In *Halliburton Energy Services Inc v Smith International (North Sea) Ltd and ors* [2006] EWCA Civ 185, [2006] RPC 26 the Court of Appeal considered just this kind of situation. It held that we must hear the appeal on its merits for it would not be right to restore a patent which has been held invalid by the court below unless that decision has been shown to be wrong. It also indicated that the Comptroller-General of Patents should be notified so that he can consider whether he wishes to be represented in order to assist the court. SSH has duly notified the Comptroller and he has responded that, after reviewing SSH's skeleton argument, he believes it is not necessary for him to intervene in the proceedings or be represented. So we have proceeded to hear SSH's appeal as a matter of substance and with only one side before us. Mr Iain Purvis QC has appeared with Mr Brian Nicholson on its behalf and we have had the benefit of their written and oral submissions. These submissions were characteristically clear and fair and we are grateful to them both.

The technical background

8. The technical background is set out in some detail in the judgment from [13] to [56], in an agreed technical primer and in SSH's written argument for this appeal. The following summary, which I hope is sufficient to understand the issues arising on this appeal, is drawn with gratitude from these sources.

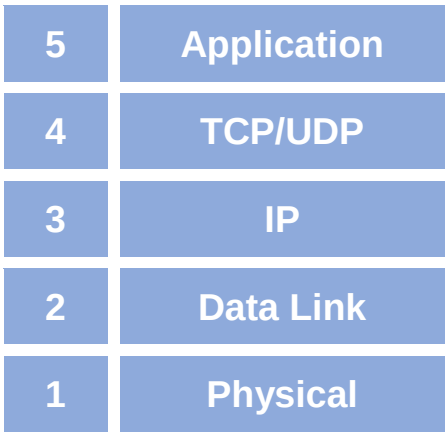
9. A computer network consists of a number of computers or, as they are sometimes called, nodes, each of which is connected to every other node in the network by means of a communications connection or link, such as a wired or wireless connection. For a node to communicate with other nodes, each must be uniquely addressable on the network.

10. Networks can be divided into two categories, local area networks ("LANs") and wide area networks ("WANs"). A LAN is a network in which all links are privately owned within a single building or site. In a WAN, some of the links span greater distances and may not be privately owned.

11. Communications over a network are specified by networking protocols which set out the procedures and data formats which should be used to convey data from one network element, such as a node, to another. Each node implements a protocol stack and data moves from the highest layer in the stack down to the lowest layer before being transmitted across a medium.

Internet Protocol Suite

12. The Internet Protocol Suite underpins the Internet and defines a five layer protocol stack in which each layer represents the functionality necessary to allow communication between two nodes in a network. It can be depicted schematically like this:



13. The functions of these layers are as follows:

    i) Layer 5, the application layer, ensures that one application program can communicate effectively with another in the network.

ii) Layer 4, the transport layer, is concerned with the reliable transfer of information across a potentially unreliable network. Two well-known transport protocols are Transmission Control Protocol ("TCP") and User Datagram Protocol ("UDP").

iii) Layer 3, the network layer, is concerned with the transfer of packets between nodes anywhere on the network, and the addressing of those nodes. The network layer in the Internet Protocol Suite is the Internet Protocol or, as it is also known, the IP.

iv) Layer 2, the data link layer, is concerned with the establishment and maintenance of actual communications between directly connected nodes. It is concerned with how data should be packaged to allow it to be transmitted over the physical medium between the nodes.

v) Layer 1, the physical layer, is concerned with the actual encoding of data onto an electromagnetic signal.

14. In the Internet Protocol Suite, data to be transmitted is created and formatted by the application layer, passed to the TCP/UDP layer which generates appropriate TCP/UDP packets. These packets are then encapsulated into IP packets for transmission over the network. An IP packet (or datagram) consists of a payload portion containing the data to be transmitted, and a header portion which contains addressing information to ensure delivery of the packet to its intended destination.

The IP Protocol

15. I must now say a little more about the IP Protocol. There are two versions of it, IPv4 and IPv6. In 1999, IPv4 was used almost exclusively. Even now, IPv4 remains dominant. Each IPv4 address is 32 binary digits (bits) long and is denoted numerically as having four octets (groups of eight bits), with each octet taking a value between zero and 255. So it takes the following format: "xxx.xxx.xxx.xxx". In theory, this allows for approximately four billion unique addresses, and so Internet nodes which can be recognised, one by another.

TCP

16. As I have explained, the IP layer facilitates the routing of an IP packet from a source node A to a destination node B. But the node may have multiple applications or processes which are connected to the IP network at the same time. For example, a desktop computer may be using a web browser and an email manager simultaneously. An IP packet arriving at node B must be directed to the relevant process.

17. This is achieved by the TCP layer which assigns a destination port number to the IP packet. This destination port number is included in the TCP header contained within the IP packet. An IP address and port number can therefore be written in the format xxx.xxx.xxx.xxx.yy where the 'x's represent the IP address and the 'y's the port number.

18. Before substantive transmission of a sequence begins, TCP sends an initial message containing header information which requires the receiving node to acknowledge it has

received that message. When this acknowledgement has been received, the sending node TCP layer begins substantive transmission of the data sequence, and at the same time acknowledges that it has received the receiving node's acknowledgement. This is known as a three-way handshake. Transmission continues until the sending node has transmitted the last segment of data. At this point it sends a FIN message to the receiving node which must then reply, acknowledging receipt and with its own FIN message which the transmitting node then acknowledges. The connection is then closed.

UDP

19. UDP operates with port numbers in the same way as TCP. But, unlike TCP, it is a connectionless protocol and does not maintain sessions. Instead, it relies on the characteristics and performance of the underlying IP network to deliver packets on a best effort basis to the receiving node. It does not specify a mechanism for signalling when a UDP transmission begins, or when it ends.

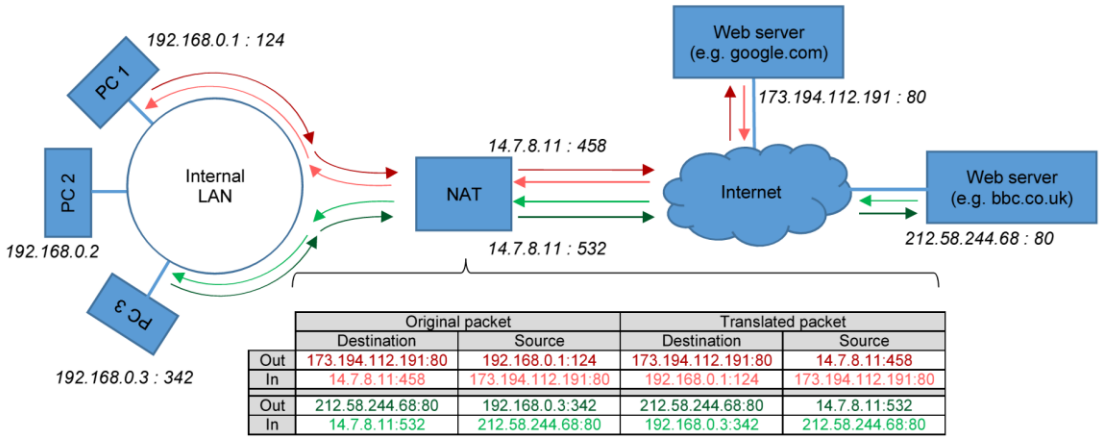
NAT

20. As I have mentioned, IPv4 allows for around four billion addresses but the number available in practice is much smaller as a result of the way in which addresses have been allocated by the central addressing authority. Generally speaking, larger institutions such as universities and substantial corporations have been allocated what effectively amounts to a block of in excess of 65,000 addresses. Smaller entities have been allocated a block of in excess of 250 addresses. Once allocated, blocks are rarely, if ever, released for reassignment.

21. A consequence of this approach to allocation has been that large numbers of addresses have been allocated to entities that have not used them, and it was becoming apparent in the late 1990s that the available IPv4 address space would run out quickly. It was known at that time that IPv6 was a long-term solution but, for various reasons which I need not go into, a solution to the IPv4 address space problem was also required.

22. The solution was provided by a system called Network Address Translation ("NAT"). This offered a way for multiple devices to connect to the Internet using a single public IP address and could be installed, for example, at the interface between a home or company and the public Internet. Nowadays, virtually every router that connects a home or small business to the Internet incorporates a NAT.

23. This representation of a NAT appears in the technical primer and the judgment:



	Original packet		Translated packet	
	Destination	Source	Destination	Source
Out	173.194.112.191:80	192.168.0.1:124	173.194.112.191:80	14.7.8.11:458
In	14.7.8.11:458	173.194.112.191:80	192.168.0.1:124	173.194.112.191:80
Out	212.58.244.68:80	192.168.0.3:342	212.58.244.68:80	14.7.8.11:532
In	14.7.8.11:532	212.58.244.68:80	192.168.0.3:342	212.58.244.68:80

Figure 4: Example NAT system (numbers are largely arbitrary)

24. In this representation, the NAT sits between an internal LAN and the Internet. The internal LAN has three personal computers (PC1 to PC3). Each personal computer has its own private IPv4 address (192.168.0.[1, 2 or 3]) and the NAT has a single public IPv4 address (14.7.8.11). In addition and as we have seen, each address can have a port number. So the complete Internet address for PC1 is 192.168.0.1:124, where 124 is the port number.

25. The following description of how the NAT in this figure works is taken with only minor changes from an explanation helpfully provided to us by counsel:

    i) PC1 (in the private network) wants to get web pages from the public Google web server (google.com) at IP:Port 173.194.112.191:80. The web-browser on the computer is able to work out both the public IP address and port number.

    ii) PC1 accordingly sends a message onto the internal LAN addressed to 173.194.112.191:80. It sends it from its own port 124 (chosen effectively at random) and states its return address in the message (so that Google can reply) as 192.168.0.1:124.

    iii) The NAT is connected to the private network and receives PC1's message.

    iv) The NAT changes the return address in the message to its own public address and (effectively randomly) chooses a unique port number (in the example, 458). It then forwards PC1's message to Google (at 173.194.112.191:80) with a public return address of 14.7.8.11:458. The NAT keeps a mapping of the translation it has made. This is shown in the first line of the translation table in Figure 4, above. The second line of the table is created at the same time; it is the same information in reverse order.

    v) When Google responds, it sends the reply to the public reply address supplied by the NAT (14.7.8.11:458).

vi) When the packet addressed to the NAT (14.7.8.11) arrives at its port 458, the NAT checks its translation table (the second line) and changes the outgoing address and port number of the message to that used by PC1 to send the original message to Google (i.e. 192.168.0.1:124), and forwards the message into the local network.

vii) A packet seamlessly arrives at PC1's port 124, purporting to come from the right address at Google (173.194.112.191:80).

26. The mapping table is maintained in the NAT until, after a certain period of inactivity, it is allowed to "time out" and in this way prevent the internal memory of the NAT being filled with mapping data that is no longer required.

27. As the judge explained, the introduction of NATs meant that, for the first time, one of the core principles of the Internet, that each node has a unique address so that it can be contacted by any other node, was broken. Counsel also submit and I accept that the potential problems that this might cause were exacerbated by a lack of standardisation of NAT operation and functionality. So each implementer devised its own approach and system for managing the translation tables to allow as much flexibility and capacity as possible without using more memory than necessary.

The patent

28. The specification explains that the invention relates to the field of secure communications between computers in packet-switched data transmission networks. Much of the specification is concerned with a protocol called Internet Protocol Security or IPSec.

29. The purpose of IPSec was to verify messages being sent over the Internet by authenticating the sender so that the recipient could be sure that the message had indeed come from the source it purported to be from. It carried this out by comparing the IP address of the computer that the message had come from (which was in the header of the message) with an encrypted version of the address which was hidden in the message itself. If the two matched, the recipient would know the message was genuine. If they did not, the recipient would be alerted to the possibility of fraud and the message rejected.

30. As the specification explains at [0007], it was known that the IPSec protocol did not work well over NATs. The reason is this. Once a computer is hidden behind a NAT, the internal address of the computer sending the message does not match the public-facing address of the NAT. The patent describes a method of addressing this problem in which the sending computer works out what address, including the port, is being used by the NAT to send the message and then incorporates that address in the encrypted part of the message. In this way the addresses match and a genuine connection is authenticated.

31. The successful operation of this method does, however, depend upon the information in the NAT translation table remaining the same. If it does not then the next packet of data sent through the NAT will be rejected by the receiving computer because the sender's IP address will no longer match that contained in the encrypted part of the message.

32. So this gave rise to a further problem, namely how to ensure that NAT devices that use timeouts for mappings do not lose the mapping for the communicating computers. It is the solution to this problem which forms the basis for the invention claimed in the patent. The problem itself is described in the specification at [0039]:

> "[0039] Next we will address the "keepalive" aspect of the invention, i.e. ensuring that the network address translations performed in the network do not change after the translations that occur have been determined. Network address translators cache the information about address mapping, so that they can reverse the mapping for reply packets. If TCP is used, the address translator may look at the FIN bit of the TCP header to determine when it can drop a particular mapping. For UDP, however, there is no explicit termination indication for flows. For this reason, many NATs will time out mappings for UDP quite fast (even as fast as in 30 seconds). Thus, it becomes necessary to force the mapping to be maintained."

33. As can be seen, this is cast in general terms and is said to be a particular problem with UDP. The solution is described in the next two paragraphs and it is, in summary, to send "keepalive" packets frequently enough to ensure that the mappings are retained in the memory of the NAT:

> "[0040] A possible way of ensuring the maintaining of mappings is to send keepalive packets frequently enough that the address translation remains in the cache. When computing the required frequency, one must take into account that packets may be lost in the network, and thus multiple keepalives must be sent within the estimated shortest period in which NATs may forget the mapping. The appropriate frequency depends on both the period the mappings are kept cached and on the packet loss probability of the network; optimal frequency values for various context may be found through experimenting.
>
> [0041] Keepalive packets do not need to contain any meaningful information other than the necessary headers that are equal to the data packet headers to ensure that the keepalive packets will be handled exactly in the same way as the actual data packets. A keepalive packet may contain an indicator that identifies it as a keepalive packet and not a data packet; however it may also be determined that all packets that do not contain meaningful payload information are interpreted to be keepalive packets. … It should be noted that the use of keepalive packets is not needed at all if actual data packets are transmitted frequently enough and/or the connection is to remain valid only for such a short time (e.g. a few seconds) that it is improbable that any intermediate device would delete the mapping information from its cache. Keepalive packets need to be transmitted in one direction only, although they may be transmitted also bidirectionally; the drawback resulting from their bidirectional transmission is the resulting increase in unnecessary network

> traffic. The invention does not limit the direction(s) in which keepalive packets (if any) are transmitted."

34. The patent has 15 claims and the judge was faced with three sets of conditional amendments. We are not concerned on this appeal with any of those conditional amendments, however; and counsel told us on instructions that we need only address claims 1 and 3 as granted. These read:

> "1. A method of maintaining communication of datagrams in a communication system where address translation is provided by a network address translator (305) for communication of datagrams between a first device and a second device, **characterised by** maintaining a determined network address translation for communication of datagrams between the first device and the second device by sending (306) from the first device or the second device at least one keepalive packet before a timeout of the determined network address translation.
>
> 3. A method according to claim one or two, wherein the at least one keepalive packet comprises a header that equals with the headers of the datagrams."

35. One of the issues at trial was the meaning of the expression "keepalive packet". The deputy judge accepted that its meaning was that contended for by SSH, namely a packet which is transmitted repeatedly for the purpose of maintaining the IP address and port mapping at the NAT.

The common general knowledge, the prior art and this appeal

36. As I mentioned at the outset of this judgment, Sony asserted the patent was (i) obvious in light of the common general knowledge and (ii) lacked novelty or was obvious in light of the NAT Minutes and Guidelines. The judge rejected (i) but accepted (ii). The first of these findings is not the subject of this appeal, but counsel submitted that it is nevertheless important to have in mind the judge's reasoning in relation to it because it has a bearing on the correctness of the second of his findings. I accept this may be the case and so I will summarise that reasoning.

The common general knowledge

37. Sony contended that the following four matters were common general knowledge and rendered the patent obvious:

    i) NATs track associations between packets and store information to allow them to translate outgoing packets and incoming reply packets; in other words, NATs cache information about address mappings;

    ii) such mappings are typically dynamic and are removed by the NAT if not refreshed by the translation of a packet; in other words, the mappings have a timeout;

    iii) keepalives provide periodic traffic and are frequently sent to refresh timers; and

iv) such applications require the consistent use of one IP address and port number in order to operate without error.

38. SSH accepted that the first of these matters was common general knowledge but disputed that any of the others were. The deputy judge agreed. So far as the third was concerned, the deputy judge accepted that certain keepalives formed part of the common general knowledge but these were either "hello" messages sent from one point in a network to another point saying "I am alive", or "probe" messages asking for an acknowledgment that that the other point was still alive. These keepalives were not of the same character as the keepalives of the patent, however.

39. The deputy judge dealt with the second and fourth matters together. He noted and apparently accepted SSHs criticism that Sony had not identified any application which required the consistent use of one IP address or port, or any application in which timeout had been a problem.

40. Sony had therefore failed to establish the foundation for its argument that the claims were obvious over the common general knowledge and this attack fell away.

NAT Minutes and Guidelines

41. On the 9 December 1998 Mr Senie presented a paper to the 43rd meeting of the Internet Engineering Taskforce in Orlando in Florida. It was entitled "NAT Friendly Application Design Guidelines". This is the Guidelines document. The relevant parts of it read as follows:

> "Preface
>
> While many common Internet applications will operate cleanly in the presence of Network Address Translators, others suffer from a variety of problems when crossing these devices. This document discusses those things which application designers might wish to consider when designing new protocols. Guidelines are presented herein to help ensure new protocols and applications will, to the extent possible, be compatible with NAT.
>
> 1. Introduction
>
> Other documents which describe Network Address Translation (NAT) discuss the Terminology and Considerations [Srisuresh1] and Protocol Issues [Holdrege] or discuss the perceived implications of NAT [Hain] [Rekhter]. All of those relate to various issues with the NAT mechanism and its effects on protocols which already exist. It is the focus of this document to instruct authors of new protocols what to think about when designing new protocols such that special handling is not required at NAT gateway points.
>
> …
>
> 2 Discussion

Network Address Translation presents a challenge to some existing applications. It should be possible for developers of new applications to avoid problems if they understand the issues involved. This document aims to provide the application designer with information on what to do, and what to avoid, when building applications.

The proliferation of NAT, especially in homes and small offices cannot be dismissed. The emerging technologies for providing high bandwidth to these types of locations often allow only a single IP address per location. As such NAT is a clear choice for connecting more than a single system per location.

Clearly the most common problem associated with NAT implementations is the passing of addressing data between stations. Where possible, applications should find alternatives to such schemes. Studying a few existing protocols will serve to highlight the different approaches possible.

Two common forms of NAT exist. With Basic NAT, only the IP addresses of packets are altered by the NAT implementation. Many applications will operate correctly with Basic NAT. The other common form is Network Address Port Translation. With NAPT, both the IP addresses and the source and destination ports (for TCP and UDP) are potentially altered by the gateway. As such, applications which pass only port number information will work with basic NAT, but not with NAPT

Application designers should ensure compatibility with NAPT, as this form of NAT is the most widely deployed. This is also the form of NAT which will likely see the greatest penetration in homes and small offices.

…

3.1 Avoid Session Bundles

Independent sessions, such as used by HTTP, are preferred to protocols which attempt to manage a bundle of related sessions, such as FTP.

In the FTP protocol, port information is passed over one TCP connection and is used to construct a second TCP connection for passing the actual data. While using a separate connection to pass the files being transferred makes determination of the end of data quite simple, other schemes could be envisioned.

The HTTP protocol, for example, uses a header and content length approach to passing data. In this model, all data is transferred over the single TCP connection, with the header portion indicating the length of the data to follow. HTTP has

> evolved to allow multiple objects to be passed on a single connection (thereby cutting the connection establishment overhead). Clearly a new file transfer function could be built that would perform most of the functions of FTP without the need for additional TCP connections.
>
> It is clear that the lesson here is to keep to single connections where possible. This keeps us from needing to pass addressing information of any sort across the network. Since addressing issues are limited to the establishment of the TCP session, standard NAT functionality is sufficient.
>
> …
>
> 3.5 TCP preferred over UDP
>
> NAT implementations must track which sessions are alive, and flush old sessions. TCP has clear advantages in this area, since there are specific beginning and end of session indicators in the packets (SYN and FIN packets). While UDP works for some types of applications with NAT, there can be issues when that data is infrequent. Since there is no clean way to know when an end station has finished using a UDP session, NAT implementations use timeouts to guess when a UDP session completes. If an application doesn't send data for a long period of time, the NAT translation may time out.
>
> 3.6. Single Sessions Preferred Over Multiple Sessions
>
> Resource utilization on the NAT gateway should be considered. An application which opens and closes many TCP connections, for example, will use up more resources on the NAT router than a similar application which performs all transfers over a single TCP connection. HTTP 1.0 opened a connection for each object on a web page, whereas HTTP 1.1 permits the TCP session to be held open for additional objects which may need to be transferred. Clearly the latter imposes a lower overhead on the NAT gateway, as it is only maintaining state on a single connection instead of multiple connections. …"

42. Sony argued at trial that a number of points emerge from the Guidelines. First, they alert application designers to the fact that, while some applications will work across NATs, others will suffer from a variety of problems. Secondly, they advise these designers that there are various things they might wish to consider when designing new protocols and that they should try and ensure compatibility with both NAT and NAT with port mapping ("NAPT"). Thirdly, they warn these designers that, while UDP works for some types of applications with NAT, there can be issues when data transfer is infrequent because NAT implementations use timeouts to guess when a UDP session is complete.

43. Minutes of the meeting were taken and were themselves published. These are the Minutes. They set out the agenda for the meeting as a whole and record that Mr Senie presented his paper. They also contain some notes of what he actually said which include the following:

> "o Operational Reliability
>
> * TCP preferred over UDP since NAT can track TCP sessions more easily and know when sessions end.
>
> * UDP sessions are tracked by timeouts.
>
> ALG's can overcome this problem, but we'd rather design applications to not need this processing."

44. Here I should interpose that it was explained to us at the appeal hearing that ALGs - application layer gateways – are pieces of software that sit in the router governing the NAT and can be addressed directly with, for example, an instruction that the session is to be kept alive. Reverting to the Minutes, these record that, after the presentation, there was a question and answer session with members of the audience. Here the relevant parts of the Minutes read:

> "Questions from the Audience:
>
> [Eliot Lear - UDP session management. UDP may make it more difficult to maintain the mapping]
>
> An application may maintain keep-alives to make this less of a problem.

45. Sony, supported by the evidence of Professor Leslie, contended at trial that Eliot Lear reiterated in his question the problem that Mr Senie had identified in his Guidelines at paragraph 3.5 and that Mr Senie then gave the answer to it, namely that it could be addressed by using keepalives. This, Sony continued, was the solution provided by the patent, and so the claims of the patent in issue were anticipated or obvious.

46. SSH, supported by the evidence of Mr Holdrege, focused on paragraphs 2, 3.1, 3.5 and 3.6 of the Guidelines and the advice they contain that the most common problem associated with NAT implementations is the passing of addressing data between stations. The Guidelines then make several different recommendations, one of which focuses on HTTP 1.1 which allows multiple objects to be transferred on a single TCP connection, and so imposes a lower burden on the NAT gateway. The Minutes then make the same point. This, SSH continued, is what Mr Lear's question and Mr Senie's answer were directed to. Mr Senie was urging authors of applications to ensure that, so far as possible, they used only a single session to send all related objects (such as all those on a single webpage) rather than open a new session for each object. This would mean that a single TCP or UDP port could be used for all traffic and in this way the use of resources in the NAT could be minimised.

47. SSH continued that this could be managed in a relatively straightforward way using TCP and HTTP. The HTTP 1.1 protocol used the term "Keep-Alive" in the context of session management to identify a system which involved the sending of a "Keep-Alive

connection token" together with another indicator called "content-length" which marked the end of each message. But in UDP there was no equivalent protocol and every new set of data was mapped differently. It was in this context that Mr Lear's question and Mr Senie's answer fell to be considered and it was likely that Mr Senie was speculating about how a client and server could communicate over a UDP connection in such a way that a single session could be indicated, maintained and managed. The parties to the UDP connection would be able to tell each other that they were beginning a session which would then be kept open long enough for all of the data to be sent.

48. The deputy judge preferred the evidence of Professor Leslie over that of Mr Holdrege. He concluded that the skilled person reading the Minutes and Guidelines together would understand them to be teaching that NATs used timeouts to clear their memory of unused mappings and that this could cause problems with UDP connections. The skilled person would also understand the question and answer session to be teaching the use of keepalives to maintain NAT mapping and in this way to solve the problem of timeouts in the NAT. Indeed, the deputy judge observed, one only had to state the problem of timeouts to arrive at the solution, particularly when the word keepalive was used.

49. Further, the deputy judge continued, Mr Holdrege's explanation was complicated and would require the use of HTTP Keep-Alive to be common general knowledge, and there was no evidence of this. The deputy judge also referred to and derived support for his conclusion from the following statement in Mr Holdrege's second report concerning HTTP:

> "This, of course, does not pose an obvious solution to the question of renewing the timeout of the network address translation. This is because not all applications know the amount of data which they are sending and, in any event, as the signaling is at the application layer, the NAT would not be able to read it."

50. The deputy judge then expressed his conclusion about the disclosure and found that it anticipated claims 1 and 2:

> "114. On the basis of Professor Leslie's evidence as to the meaning of the Minutes and the Guidelines, as they would be understood by the skilled person, I find that Claims 1 and 2 are anticipated by the disclosure which clearly teaches the use of keepalive packets to maintain a NAT mapping for the communication of datagrams by sending a keepalive packet before a timeout of the mapping by the NAT device."

51. The deputy judge turned next to obviousness on the assumption he was wrong about the teaching being clear. Here he adopted the approach explained by the Court of Appeal in *Pozzoli SpA v BDMO SA* [2007] EWCA Civ 588, [2007] FSR 37 at [14] to [23]. He considered the inventive concept of claim 1 to be the sending of keepalive packets between two devices communicating with each other through the NAT, in order to maintain the mapping in the NAT, and that such keepalive packets should be sent before timeout. He perceived the difference, if any, between the Minutes and Guidelines and the claimed invention to be that the Minutes and Guidelines do not

specify that keepalive packets must be sent between devices sufficiently frequently that timeout does not remove the mapping prematurely. But it was his view that, on the evidence, this would have been obvious to the skilled addressee.

52. As for claim 3, this adds the feature that the header used for the keepalive message is the same as that used for other datagrams. In other words, the keepalive is addressed to the same endpoint and in the same way as ordinary data sent over the connection.

53. The deputy judge found that this particular feature is not disclosed in the Guidelines or the Minutes but was nevertheless obvious. He accepted the evidence of Professor Leslie that keepalives are sent on the communication path of a normal message and that the skilled person would have known that they had to go through the NAT. Without the headers being equal, the NAT would not be able to translate the keepalive packets and corresponding datagrams at all and so the NAT would not be refreshed. The implementation of the invention of claim 1 would therefore have required the use of the invention of claim 3 and this would have been obvious.

54. Upon this appeal counsel for SSH have argued that the deputy judge fell into error in numerous respects. They contend that in order to get to the invention from the NAT Minutes and Guidelines, the deputy judge had to read at least three things into the words as a matter of disclosure:

    i) He had to find that the words "UDP may make it more difficult to maintain the mapping" were a reference to the fact that UDP mapping would be lost by timeouts in certain implementations.

    ii) He had to construe the words "an application may maintain keep-alives" as meaning the sending of a packet between two devices communicating with each other, through the NAT, in order to maintain the network address translation at the NAT.

    iii) He had to construe the words "to make this less of a problem" as disclosing the sending of a packet at intervals sufficiently short to prevent the timeout affecting the connection.

55. Counsel continue that there is an obvious difficulty with this analysis which the deputy judge failed to recognise, namely that, for a claim to be anticipated, the prior art citation must disclose the invention clearly and unambiguously or clearly and unmistakeably. They argue that the first of the three propositions set out at [54] above is just about manageable, even though the loss of UDP mapping through timeouts in NATs was not part of the common general knowledge. But they say that the second and third propositions have no basis in the Minutes, the Guidelines or the common general knowledge. In the circumstances, they continue, it is impossible to conclude that the skilled team would have understood the term "keep-alives" as a reference to the keepalives the subject of the claimed invention.

56. The position is still worse for Sony, counsel argue, because the term "Keep-Alive" was used in the context of session management in the HTTP 1.1 protocol. This protocol is referred to in paragraphs 3.1, 3.5 and 3.6 of the Guidelines. Mr Holdrege explained that in this context the term denoted the sending of the "Keep-Alive connection token" together with "content-length" which allowed parties to a connection to tell each other

that they were beginning a session which would be kept open long enough for a fixed number of bits to be sent, and in this way reduce the number of connections and so preserve the resources in the NAT. Indeed, said Mr Holdrege, this is precisely what is described in 3.6 of the Guidelines and it has nothing to do with timeout in the NAT.

57. The deputy judge's reasons for rejecting the evidence of Mr Holdrege are, counsel argue, completely flawed. The first was that it had not been established that the HTTP Keep-Alive was a matter of common general knowledge and there was no evidence about it. In fact, however, the HTTP 1.1 Protocol was extensively referred to in the Guidelines and was in any event a "standards track" protocol which had been in widespread use since 1990. Mr Holdrege said in unchallenged evidence that these were likely to be referred to by the skilled person as a matter of course and that this person would have considered them sufficiently reliable to use. What is more, HTTP was referred to as an important protocol for key Internet applications in the technical primer and this was itself accepted to be common general knowledge.

58. Counsel also argue that the deputy judge's second reason is no better. Here he referred to and apparently derived support for his conclusion from the following statement in Mr Holdrege's second report:

> "This, of course, does not pose an obvious solution to the question of renewing the timeout of the network address translation. This is because not all applications know the amount of data which they are sending and, in any event, as the signaling is at the application layer, the NAT would not be able to read it."

59. They continue that this is in fact entirely consistent with what Mr Senie said. He did not claim that it would solve the problem in the way that the patented solution does. Rather, it would make it less of a problem.

60. Drawing the threads together, counsel submit that the deputy judge's findings on disclosure of the invention by the Minutes and Guidelines were fatally undermined by his failure to apply the law requiring a clear and unambiguous disclosure; by his plain mistake in finding that HTTP 1.1 was not part of the common general knowledge; and by his misunderstanding of the significance of Mr Holdrege's evidence about the utility of the HTTP 1.1 session management solution. In all these circumstances the deputy judge should have found either that Mr Holdrege's reading of the Minutes and Guidelines was correct, or, at the very least, that the disclosure failed to give clear and unambiguous directions to make something falling within claim 1 or claim 3.

61. Turning to obviousness, counsel submit that the deputy judge ought to have considered this on the basis that SSH's interpretation of the disclosure of the Minutes and Guidelines was to be preferred. Alternatively, if the teaching of the Minutes and Guidelines was not clear, the deputy judge did not explain the basis of his finding that the claims would have been obvious. The critical question why it would have been obvious for the skilled person to take the necessary steps was simply not addressed. SSH was entitled to a properly reasoned decision and the deputy judge failed to provide one. A proper consideration of the issue would have required an acknowledgment of the peculiar nature and context of this disclosure, and it would also have required some recognition of the fact that "keepalive" was not a term of art with a specific meaning.

The deputy judge ought to have found that, without hindsight, the invention was simply not obvious.

62. Counsel also submit that the deputy judge had no proper basis upon which to find that the invention the subject of claim 3 obvious, and that this finding was based upon a misunderstanding of the evidence and unsound reasoning.

63. It is apparent from the submissions made to us on this appeal and my review of the arguments developed before the deputy judge that the difference between the parties was founded upon the fundamentally different interpretations of the Guidelines and Minutes for which each of them contended. The first task for the deputy judge was therefore to identify what information those documents conveyed to the public at the priority date, and this was a matter of construction through the eyes of the notional skilled person. The deputy judge had to decide, assisted where necessary by the evidence of the experts as to the meaning of the technical language, what the documents disclosed.

64. The next question for the deputy judge was whether this disclosure deprived claim 1 or claim 3 of novelty. It is of course well established that, to make good its case, Sony had to show that the documents disclosed subject matter which would, if performed, necessarily result in an infringement of the patent if carried out after its grant. They had to contain a clear description of, or clear instructions to do or to make, something that would infringe.

65. The third question for the deputy judge, on the assumption the disclosure did not deprive either claim of novelty, was whether it nevertheless rendered each claim obvious.

66. Reverting to the first question, I am satisfied that the deputy judge made no error in the way he approached the interpretation of the Guidelines and Minutes and that the conclusion he reached was not only open to him but also correct. At the outset he recognised that the common general knowledge as to the structure and means of operation of NATs was limited in that it comprised only the fact that NATs cache information about address mappings. It did not extend to timeouts or any relevant keepalives.

67. The deputy judge then recited the relevant parts of the Guidelines and the Minutes, full details of which I have set out earlier in this judgment. It is in my view important to note that the purpose of the Guidelines was to provide instructions to authors of new protocols as what to think about when designing those protocols such that special handling was not required at NAT gateway points.

68. To this end, the Guidelines contain a description of the challenges that NATs present, and information as to what to do and what to avoid when building applications. There follows a description of the two common forms of NAT, basic NAT and NAPT. Then, in section 3, the Guidelines explain the benefits of TCP and HTTP, outline the approach used by HTTP to the passing of data and describe its evolution to allow multiple objects to be passed on a single connection. They teach the reader to keep to single connections where possible.

69. Sections 3.5 and 3.6 are clearly of the greatest importance. As we have seen, section 3.5 states in terms that TCP is preferred over UDP because it has beginning and end of session indicators in the packets, and it then describes in concise but clear terms how UDP suffers from the problem of timeouts.

70. Section 3.6, on the other hand, is directed to a different issue, namely the benefit of single sessions over multiple sessions. Here it is explained that the benefit of HTTP 1.1 is that it permits the TCP session to be held open for additional objects which may need to be transferred and in this way imposes a lower overhead on the NAT gateway.

71. Turning now to the relevant parts of the Minutes, these echo section 3.5 of the Guidelines and record expressly that TCP is preferred over UDP because of the ability of the NAT to track sessions more easily and know where they end. They also explain that UDP sessions are, by contrast, tracked by timeouts; and that although ALGs can overcome this problem (i.e. timeouts), they are not the most desirable solution. Then, critically, there follows the question and answer exchange with Eliot Lear. Here the Minutes are, to my mind, clear. Mr Senie provided a solution or, more accurately, a partial solution to the problem identified in the Guidelines that UDP may make it more difficult to maintain mapping in the NAT, and that solution is to use "keep-alives". It is true that there is no record of precisely what Mr Senie meant by "keep-alives" but it was, in my judgment, self-evidently the sending of periodic data to refresh the timer and prevent timeout and loss of the mapping tables in the NAT. The question concerned the loss of mapping in the NAT and the answer was to "maintain keep-alives". Furthermore, it is clear from the context that this has nothing to do with HTTP "Keep-Alive" tokens which operated over TCP and in the application layer of the protocol.

72. I therefore reject the submission that the deputy judge fell into error in the way he interpreted the words of the Minutes and Guidelines. For the reasons I have given, the skilled person would have understood the words "an application may maintain keep-alives to make this less of a problem" as being directed to the issue of timeouts in UDP and the loss of mapping in the NAT. The skilled person would also have understood the disclosed solution to be the sending of packets of information at intervals sufficiently short to prevent or at least ameliorate timeout affecting the connection. Indeed, Mr Holdrege accepted as much in the course of his cross-examination. He was asked to assume that the skilled person either knew or would find out that UDP transmissions through a NAT may timeout and then this interchange took place:

> "Q: If he [the skilled person] knows there is a timeout, in some applications, that is not going to matter because if it is a continuous stream, it will not timeout. If he is concerned, because maybe he has got intermittent data, then he knows, does he not, that a way of stopping that timeout is to send a packet through to prevent the timeout?
>
> A: yes."

73. I accept that the term "Keep-Alive" was used in the context of session management in the HTTP 1.1 protocol to refer to the tokens used in the application layer as part of the HTTP packet. I accept too that these protocols are referred to in the Guidelines. But these protocols are concerned with the use of HTTP and TCP; and the Keep-Alive tokens in HTTP have nothing to do with UDP and the problem of timeouts and the

consequential loss of mapping tables in the NATs. As Professor Leslie explained, the NAT is completely unaware of the Keep-Alive indicator because it is a feature of a higher level protocol and has no effect on the mapping or its timeout.

74. I also reject the submission that the deputy judge made any error when assessing the common general knowledge and, in particular, in failing to find that the HTTP Keep-Alive was common general knowledge. I have to say it is far from clear to me that the deputy judge was ever asked to make a specific finding that it was common general knowledge. Be that as it may, the particular point made by Sony in its closing submissions was that it had not been shown that it was a matter of common general knowledge that the term "Keep-Alive" to identify this feature of HTTP was common general knowledge. I accept that it is referred to in HTTP 1.1 and that this was a standards track protocol. I also accept that HTTP was an important protocol and that the protocol is referred to in the technical primer. But this does not mean that that every feature in the protocol was common general knowledge; nor does it mean that it was a matter of common general knowledge that each feature in the protocol had the name used there to identify it. The protocol is long and detailed and as Professor Leslie explained in his second report, the use of the term Keep-Alive in this context was an oddity resulting from early versions of HTTP.

75. The deputy judge was also entitled to attach some weight to the evidence given by Mr Holdrege in his second report which I have set out at [49] above. The point he was making was entirely fair. Even if the use by Mr Senie of the term "keep-alive" did bring the HTTP Keep-Alive to mind, the skilled person would realise this was not what Mr Senie was referring to for the reasons Mr Holdrege gave. The HTTP Keep-Alive is sent at the application layer at the beginning of the session. But there are two problems: first, not all applications know the amount of data which they are sending; and secondly, the NAT cannot read the signaling and so timeout will still take place.

76. It was also argued before the deputy judge and put to Professor Leslie that the use of a keepalive in the invention would be a complete solution to the problem of timeout and loss of mapping yet the Minutes say that Mr Senie indicated it would simply be less of a problem. The deputy judge did not think much of this point and neither do I. He pointed out, entirely fairly, that different devices might have different timeout periods and so the skilled person would realise it might be necessary to optimise the rate at which the keepalives are sent.

77. The judge therefore came to the right conclusion about the disclosure of the Guidelines and the Minutes. Read together they clearly and unmistakeably disclose a method of maintaining communication of datagrams between two devices using a NAT in which the mapping is maintained in the NAT by sending from one device to the other at least one keepalive packet before the timeout of the NAT. The judge was also right to find that this disclosed the invention of claim 1.

78. In these circumstances it is not necessary to address the deputy judge's further finding that if he was wrong about the disclosure being clear then any difference between that disclosure and the invention of claim 1 would have been an obvious step to take. Nevertheless and since we heard argument on the point, I will state my view, albeit briefly. In my judgment the criticisms made of the deputy judge's reasoning are wide of the mark. As he explained, any difference could only be that those documents did not clearly disclose the sending of the keepalive packets sufficiently often to prevent

timeout. In my view the deputy judge was entitled to find that this was an obvious step for the skilled person to take and it was a finding which was indeed supported by the evidence of Professor Leslie. The Guidelines say that if an application (using UDP) does not send data for a long period of time then the NAT translation may timeout and, as Professor Leslie said, the skilled person would realise that if an application had no actual data to send, an obvious way to fill the gaps would be with keepalives.

79. As for claim 3 and as we have seen, this adds the feature that a header used for the keepalive message is the same as that used for other datagrams. In other words, the keepalive is addressed to the same endpoint and in the same way as ordinary data sent over the connection. The deputy judge found that this particular feature was not disclosed in the Guidelines or the Minutes but was nevertheless obvious. Here he accepted the evidence of Professor Leslie that keepalives are sent on the communication path of a normal message and that the skilled person would know that they have to go through the NAT. Without the headers being equal, the NAT would not be able to translate the keepalive packets and corresponding datagrams and so the NAT would not be refreshed. The implementation of the invention of claim 1 would therefore have required the use of the invention of claim 3 and this would have been obvious. There is no flaw in this analysis. Indeed I agree with it.

Conclusion

80. For all the reasons I have given, I would dismiss this appeal.

Lord Justice Floyd

81. I agree.

Lord Justice Patten

82. I also agree.