United States Court of Appeals for the Federal Circuit

02-1201, -1225

RESONATE INC.,

Plaintiff-Appellant,

v.

ALTEON WEBSYSTEMS, INC.,

Defendant-Cross Appellant.

Robert P. Feldman, Wilson Sonsini Goodrich & Rosati, of Palo Alto, California, argued for plaintiff-appellant. With him on the brief were Elizabeth M. Saunders, Roger J. Chin, and Maura L. Rees.

H. Mark Lyon, Gibson, Dunn & Crutcher LLP, of Palo Alto, California, argued for defendant-cross appellant. With him on the brief were Jonathan C. Dickey, David A. Zonana, and Gillian Thackray.

Appealed from: United States District Court for the Northern District of California

Judge Claudia Wilken

United States Court of Appeals for the Federal Circuit

02-1201, -1225

RESONATE INC.,

Plaintiff-Appellant,

v.

ALTEON WEBSYSTEMS, INC.,

Defendant-Cross Appellant.

_____

DECIDED: August 5, 2003

_____

Before BRYSON, <u>Circuit Judge</u>, PLAGER, <u>Senior Circuit Judge</u>, and PROST, <u>Circuit Judge</u>.

PLAGER, <u>Senior Circuit Judge</u>.

Resonate Inc. owns U.S. Patent No. 5,774,660, entitled "World-Wide-Web Server with Delayed Resource-Binding for Resource-Based Load Balancing on a Distributed Resource Multi-Node Network." Resonate filed suit in the United States District Court for the Northern District of California against Alteon Websystems, Inc. alleging infringement of claim 6 of the '660 patent. The district court held a Markman hearing and entered a claim construction order. Thereafter, based on the court's interpretation of a particular limitation in claim 6, Resonate stipulated that it could not prevail on its

claim of infringement, either literal or under the doctrine of equivalents. Subsequently, the district court entered final judgment of noninfringement in favor of Alteon. Because the trial court erred in its construction of the limitation at issue, we vacate the judgment and remand for further proceedings.

BACKGROUND

A.

The Internet is a global network connecting millions of computers in more than 100 countries. The World Wide Web, a collection of files, or 'web pages,' containing text, graphics, audio, and video, as well as 'hyperlinks' to other web pages, has become a central part of the Internet. Consumers typically access the web using client software applications known as web browsers that run on their personal computers.

Every web page is identified by a unique Uniform Resource Locator (URL). Web pages are stored on 'web sites,' locations on the World Wide Web comprising one or more computers, known as servers. Every web site has a home page, which is identified by a URL and is the first document users see when they first connect to the web site. Also associated with each web site is a domain name, usually part of the URL.

Each web server typically is identified by a unique 32-bit numeric address known as an Internet Protocol address, or IP address. When a user requests a web page by entering a URL into a browser, the URL is sent to a domain name system (DNS) server, which uses a look-up table to translate the domain name in the URL into the IP address of a server associated with the web site being accessed. That IP address is returned to the browser, which then uses the address to initiate a communications session with the server that contains the desired web page.

While smaller web sites may be served by a single computer, larger web sites that receive a large number of requests from clients (i.e., users accessing the web) may require multiple computers acting as servers. Such a configuration is referred to as a 'server farm' or 'web farm.' If a web site resides on

more than one server, a mechanism is needed for distributing client requests among the multiple servers so that no single server is overloaded. This process is referred to as 'load-balancing.' In a prior art approach described in the '660 patent, each server in a server farm contains a copy of the entire web site. Using a method known as 'round-robin load-balancing,' requests from users to access web pages on the web site are sent to the servers in a rotating fashion. This is implemented by including multiple IP addresses, one for each server, in the look-up table of the DNS server, and returning the IP addresses to clients in a round-robin fashion.

One drawback of round-robin load-balancing is that if one of the servers associated with a web site 'crashes,' clients will receive error messages when attempting to access the defective IP address of the crashed server because it may take several hours or even days before the defective IP address is removed from DNS server caches and browser caches, which temporarily store IP addresses. An alternative prior art solution, 'router-based load-balancing,' mitigates this problem by placing a router as an intermediary between browser clients and a web site's server farm. The IP address of the router is the only IP address available to DNS servers as the IP address corresponding to the web site. Thus all requests from client browsers to access the web site are sent to the router, which then uses an algorithm to determine which server in the server farm should service the request. The router retransmits the browser's request to the selected server, which transmits the requested file back to the router, which in turn retransmits the file to the browser. Each transmission is accomplished via a separately established connection.[1]

The '660 patent identifies two problems with the prior art router-based load-balancing approach. First, each server in a server farm must store a complete copy of the entire web site, i.e., the content must be mirrored. This implementation may be expensive, especially for certain web applications such as multimedia and video that consume large amounts of disk space. A second disadvantage of the router-based system is that all data transfers go through the router. When large amounts of data (e.g., video files) are sent from a server to a client browser through the router, a bottleneck may occur at the router and adversely affect system performance.

The invention described in the '660 patent, referred to as 'resource-based' or 'content-based' load-balancing, solves the first problem by distributing the web site's content among the servers. A router, also referred to in the patent as a load balancer, first determines what type of information the client is requesting and then selects a server to handle the request based on the content requested. The patent refers to this feature as 'delayed resource-binding.' As described in the preferred embodiment, a connection is first established between the client and the load balancer. Once the load balancer has selected the appropriate server, the connection is transferred to the server, and the server provides the requested file to the client.

In the preferred embodiment of the '660 patent, delayed resource-binding is accomplished by specific modifications to the TCP/IP protocol. The methodology disclosed as the preferred embodiment also solves the second problem identified in the prior art—the potential bottleneck at the router caused when large amounts of data are sent from the server to the client browser via the router. As described in the patent, the process permits data transmitted from the selected server back to the client to bypass the load balancer. This 'bypass' feature eliminates the bottleneck that might otherwise occur at the load balancer.

<div align="center">B.</div>

The dispute between the parties centers on whether claim 6, the only claim at issue, by its terms requires the data transmission path from the selected server back to the client to bypass the load balancer. Claim 6 provides (emphasis added to highlight the disputed claim limitation):

> 6. A computer-implemented method of servicing requests for resources from a client by nodes containing different resources, the computer-implemented method comprising the steps of:
> making a connection and setting up a session between the client and a load balancer at a web site for servicing requests from clients;
> waiting for a URL request from the client once the load balancer has made the connection with the client;
> receiving the URL request from the client and decoding the URL request to determine a requested resource;
> comparing an identifier for the requested resource to identifiers for resources located on a plurality of nodes and determining a first subset of the plurality of nodes which contain the requested resource and a second subset of the plurality of nodes which

        do not contain the requested resource;

        assigning the URL request to an assigned node in the first subset of the nodes which contain the requested resource, by determining the assigned node to be a server in the first subset of the nodes which is least busy processing requests, wherein the assigned node is not in the second subset;

        transferring the connection and the session setup to the assigned node containing the requested resource by storing packets received from the client when establishing the connection and by transmitting the packets to the assigned node after the URL request is received;

        reading the requested resource on the assigned node and transmitting the requested resource to the client,

        whereby the assigned node is selected based on a location of the requested resource determined from the URL request and load balancing is performed among nodes having the requested resource and the connection is transferred from the load balancer to the assigned node by re-transmitting the packets to the assigned node.

The key phrase "transmitting the requested resource to the client" lies at the heart of the dispute. The district court construed the phrase to mean "transmitting outbound data packets from the server directly to the client using the connection with the client which was transferred to the server, causing the outbound data to bypass the load balancer." (Emphasis added.) In other words, the district court held that claim 6 requires the bypass feature.

In the accused devices, Alteon's Web Switch products running WebOS software, all data transmitted from the selected server to the client passes through the load balancer. Under the district court's construction, the accused devices could not infringe claim 6. Accordingly, the patentee Resonate stipulated that it could not prevail under the district court's construction of the disputed limitation; a judgment of noninfringement was entered.

Resonate timely appeals the final judgment entered by the district court. Alteon supports the district court's construction and judgment, and in addition has filed what it denominates as a cross-appeal. In its cross-appeal Alteon states that, in the event we disagree with the district court's construction of the phrase "transmitting the requested resource to the client," and the matter is returned to the district court for possible trial before a jury, we should "supplement" the court's construction of two additional limitations in claim 6—"making a connection" and "transferring the connection."

<div align="center">DISCUSSION</div>

<div align="center">A.</div>

The single issue raised on appeal by Resonate is the correctness of the district court's claim construction of the disputed phrase in claim 6. Claim construction is a matter of law over which we exercise independent review. Cybor Corp. v. FAS Techs., Inc., 138 F.3d 1448, 1456 (Fed. Cir. 1998) (en banc).

As always, we begin our claim construction analysis with the language of the claim. There is a "heavy presumption" that the terms used in claims "mean what they say and have the ordinary meaning that would be attributed to those words by persons skilled in the relevant art." Tex. Digital Sys., Inc. v. Telegenix, Inc., 308 F.3d 1193, 1201 (Fed. Cir. 2002). After identifying the ordinary meaning of a disputed claim term, we turn to the patent's written description and drawings to determine whether that meaning is inconsistent with the patentee's use of the term, Rexnord Corp. v. Laitram Corp., 274 F.3d 1336, 1342 (Fed. Cir. 2001), for example whether the patentee has specially defined the term or otherwise limited the scope of the claim. SciMed Life Sys., Inc. v. Advanced Cardiovascular Sys., Inc., 242 F.3d 1337, 1344 (Fed. Cir. 2001).

However, the written description is not a substitute for, nor can it be used to rewrite, the chosen claim language. Though understanding the claim language may be aided by the explanations contained in the written description, it is important not to import into a claim limitations that are not a part of the claim. For example, a particular embodiment appearing in the written description may not be read into a claim when the claim language is broader than the embodiment. Electro Med. Sys., S.A. v. Cooper Life Scis., Inc., 34 F.3d 1048 (Fed. Cir. 1994). In addition to the written description, prosecution history may be relevant to the claim construction process because statements made by a patentee during prosecution also may affect the scope of the claimed invention. Rexnord, 274 F.3d at 1343.

The disputed language in claim 6—"transmitting the requested resource to the client"—specifies nothing regarding the transmission path over which the requested data must be sent. It merely requires the requested resource to be transmitted from the "assigned node" (the server identified as the "least busy processing requests" from among those that contain the requested resource) to the client that

initiated the request. The limitation contains no language regarding the load balancer's involvement in the transmission of the requested resource to the client; by its terms, the limitation fails to specify whether data must pass through or bypass the load balancer. Thus, based on the plain language of the disputed limitation, any transmission path from the selected server to the client appears to be within the scope of claim 6. The question then becomes whether there is sufficient reason—in the other language contained in the claim, in the written description, or in the prosecution history—to read into the claim the further limitation that the data must bypass the load balancer on its return to the client.

The district court based its claim construction primarily on its observation that every other step of claim 6 is described in detail, including every interaction between the client, load balancer, and assigned server. From this, the court inferred that if the requested data passed through the load balancer on its way to the client, that step would have been detailed in the claim as well. Because the limitation simply requires the server to transmit the requested resource to the client, the district court held that the language implies that the load balancer is bypassed.

The district court's 'level of detail' analysis does not withstand close scrutiny. The patentee's apparent choice not to specify a transmission path from the server to the client led the district court to add a limitation that the requested resource be transmitted directly to the client. But patentees are not required to claim each part of an invention with the same amount of detail; indeed, such a rule likely would prove unworkable. Courts may not rewrite claim language based on what has been omitted from a claim, and the district court's attempt to do so here was legal error. See K-2 Corp. v. Salomon S.A., 191 F.3d 1356, 1364 (Fed. Cir. 1999) ("Courts do not rewrite claims; instead, we give effect to the terms chosen by the patentee."); Autogiro Co. of Am. v. United States, 384 F.2d 391, 396 (Ct. Cl. 1967) ("Courts can neither broaden nor narrow the claims to give the patentee something different than what he has set forth.").

Alteon argues that the district court's reading of the plain language of the claim is correct when the claim is considered as a whole. According to Alteon, bypassing the load balancer is the natural

consequence of other limitations set forth earlier in the claim. While it is true that a disputed claim limitation is construed in the context of other words and limitations in the claim, see Brookhill-Wilk 1, LLC v. Intuitive Surgical, Inc., No. 02-1145, 2003 WL 21488142, at *1 (Fed. Cir. June 27, 2003), that canon of claim construction does not help Alteon here.

Most of the steps set out in claim 6 relate to the process of selecting a server based on the resource requested by the client. After the initial step of "making a connection" between the client and load balancer, the claim includes several steps detailing the server selection process, followed by the step of "transferring the connection" to the selected server. These steps constitute the 'delayed resource-binding' aspect of the invention; a connection between the client and a server is not made until the load balancer determines which server should receive the client's request. Only the final step of claim 6 involves transmitting the requested resource back to the client.

The crux of Alteon's argument is that once the connection is transferred to the server, transmission of data from the server to the client must occur over that connection, and use of that connection must necessarily cause the data to bypass the load balancer. This line of reasoning is flawed. First, nothing in claim 6 requires that the requested resource be transmitted to the client using the transferred connection. Second, and the issue focused on by the parties, is that even if data transmitted from the server to the client uses the transferred connection, no language in the claim indicates that data transmitted via the transferred connection necessarily bypasses the load balancer.

Alteon fills in the missing language by arguing that a proper understanding of the term "connection" in the TCP/IP context requires that the transferred connection must bypass the load balancer. We agree with Resonate that Alteon's view is overly narrow for at least two reasons. First, as the district court correctly held, nothing in the language of claim 6 requires the connection to be defined in terms of TCP/IP. In holding that a connection is established by the exchange of information uniquely identifying the connection, the district court accepted a portion of Alteon's proposed construction, to which Resonate eventually acceded. The court, however, declined to restrict its construction to the specific unique information Alteon argues is required by a TCP connection.

Second, even if it is proper to consider the TCP/IP context when construing the term "connection," there is no basis for limiting the interpretation to the particular TCP connection Alteon proposes. As Resonate correctly points out, the written description of the '660 patent teaches that TCP/IP can be modified to achieve desired routing of data packets. Therefore, even if one interpreted the term "connection" in the context of TCP/IP, it would be improper to restrict the meaning to any particular implementation of TCP/IP. We conclude that the language of claim 6, both the disputed phrase and the other language of the claim, does not require data transmitted from the server to the client to bypass the load balancer.

The next question to be addressed is whether the written description limits the scope of the claim. Alteon argues that the written description establishes that the claimed invention must include the bypass feature to avoid the potential bottleneck at the load balancer when large files are transmitted from the server to the client via the load balancer. In support of its position, Alteon refers us to several passages in various parts of the written description, including the abstract, "Background of the Invention," and "Summary of the Invention" sections, all of which discuss the bottleneck problem. In response, Resonate cites portions of the written description, including a section entitled "Advantages of the Invention" and an overview of the inventive method in the "Summary of the Invention" section, from which any mention of bypass is conspicuously missing.

Alteon accuses Resonate of ignoring bypass as one of the main inventive features by selectively citing parts of the written description. Resonate responds by criticizing Alteon for focusing on the bypass feature to the exclusion of the other main aspect of the invention, arguing that the key novel feature of the invention is the delayed resource-binding, which allows for content-based load-balancing and eliminates the need to mirror content on the multiple servers comprising a web site's server farm. The bypass feature, Resonate contends, is an additional inventive aspect, not included within the particular terms of claim 6.

The issue at this point may be stated thus: when the written description sets out two different problems present in the prior art, is it necessary that the invention claimed, and thus each and every

claim in the patent, address both problems? We conclude that on the record in this case, the answer is no. While not explicitly mentioning the prior art problem of having to store complete copies of the web site's content on multiple servers, claim 6 contains provisions that refer to the solution to this problem: the preamble language indicating that the nodes (servers) contain different resources, and the steps relating to the delayed resource-binding aspect of the invention. The claim does not, however, make any reference to the bottleneck problem or to the bypass solution of that problem. Therefore, without explicit claim language relating to the bottleneck problem, we see no reason why the invention as recited in claim 6 must include the bypass feature described in the written description as a solution to that problem. See Honeywell, Inc. v. Victor Co. of Japan, 298 F.3d 1317 (Fed. Cir. 2002) (holding that claimed invention did not need to solve both problems in the prior art when claim included language addressing only one problem).

Alteon also relies on the "Detailed Description" portion of the written description and the preferred embodiment contained therein for support of its position that claim 6 must include the bypass feature. It is undisputed that in the preferred embodiment of the '660 patent, data transmitted from the server to the client bypasses the load balancer. Indeed, the "Detailed Description" section contains a subsection entitled "Outgoing Data Bypasses Load-Balancer." However, as noted earlier, limitations may not be read into a claim from a preferred embodiment when the claim language is broader than that embodiment. See Electro Med. Sys., 34 F.3d at 1054. In this case, claim 6 is broader than the preferred embodiment because, as discussed, there is no language in claim 6 relating to the bottleneck problem or the bypass solution of that problem. Therefore, we conclude that the bypass feature, though shown as part of the preferred embodiment and discussed in the written description as a solution to the potential bottleneck problem in the prior art, is not properly a limitation of claim 6.

Nor is this conclusion undercut by anything in the prosecution history. The only argument in that regard that Alteon makes is that the examiner required the inventors to combine both the steps of delayed load-balancing and transferring the connection before he allowed the patent to issue. Even if this statement is true, which Resonate disputes, it is irrelevant because the thrust of Alteon's argument stands on its contention that bypass is an inevitable result of transferring the connection, an argument we

have rejected. Therefore, we see nothing in the prosecution history to alter our conclusion that claim 6 does not require that data, transmitted from the server to the client, bypass the load balancer.

<div align="center">B.</div>

Alteon filed what it denominated as a cross-appeal in this case. In the event this court concludes that further proceedings are required in the district court, Alteon asks that we "supplement" the district court's claim construction with respect to two other limitations in claim 6. Alteon notes that it does not disagree with the district court's basic construction of the phrases at issue in these other two limitations, but believes it would be helpful if we would provide a "refinement" of the district court's work. On its face, this is not a cross-appeal. We agree with Resonate that the cross-appeal must be dismissed, and Alteon's reply brief based on its cross-appeal stricken from the record.

The question of whether a cross-appeal is properly before us implicates our jurisdiction to decide the issues raised in the cross-appeal. No cross-appeal is needed in order for a prevailing party to present any legitimate argument in support of the judgment below, even if the argument was rejected or ignored by the trial court. As we pointed out in the court's Order in <u>Bailey v. Dart Container Corp. of Michigan</u>, 292 F.3d 1360, 1362 (Fed. Cir. 2002), a cross-appeal is necessary and appropriate only when a party seeks to enlarge its own rights under the judgment under review, or to lessen the rights of its adversary under the judgment. Here Alteon's base case is in support of the judgment as issued; its "cross-appeal" is a suggestion that, should it fail in that effort, it would like us to render an advisory opinion regarding matters not properly before us. This we may not do. (As it turns out, we necessarily addressed at least in part the matter Alteon attempted to raise. In our earlier discussion we considered the impact of construing the "transmitting" limitation in the context of the claim as a whole, and concluded that the district court correctly declined to restrict any part of claim 6, including the "making a connection" and "transferring the connection" limitations, to the TCP/IP context.)

We have considered the parties' other arguments, and find them to be unpersuasive or unnecessary for resolution of this appeal.

CONCLUSION

For the foregoing reasons, we hold that the district court erred in construing the limitation "transmitting the requested resource to the client" in claim 6 to require that data transmitted from the server to the client must bypass the load balancer. Accordingly, we vacate the district court's judgment of noninfringement of the '660 patent, and remand for further proceedings not inconsistent with this opinion.

VACATED and REMANDED

---

[1]   In the context of the Internet, a 'connection' is not a physical one but instead is a logical circuit allowing for the exchange of information between two computers. The standard protocol suite used on the Internet is TCP/IP, named after the two main protocols included in the suite, TCP (Transmission Control Protocol) and IP (Internet Protocol). As already mentioned, the IP protocol provides the addressing scheme used on the Internet; it also specifies the format of data packets. The TCP protocol governs the exchange of information via TCP/IP packets necessary to establish a connection between two computers.