Trials@uspto.gov
571-272-7822

Paper 83
Entered: May 15, 2014

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

EMC CORPORATION and VMWARE, INC.,
Petitioners,

v.

PERSONALWEB TECHNOLOGIES, LLC and
LEVEL 3 COMMUNICATIONS, LLC,
Patent Owners.

Case IPR2013-00082
Patent 5,978,791

Before KEVIN F. TURNER, JONI Y. CHANG, and
MICHAEL R. ZECHER, *Administrative Patent Judges*.

ZECHER, *Administrative Patent Judge*.

FINAL WRITTEN DECISION
35 U.S.C. § 318(a) and 37 C.F.R. § 42.73

I. BACKGROUND

EMC Corporation and VMware, Inc. (collectively, "EMC") filed a Petition on December 15, 2012, requesting an *inter partes* review of claims 1-4, 29-33, and 41 ("the challenged claims") of U.S. Patent No. 5,978,791 (Ex. 1001, "the '791 patent"). Paper 8 ("Pet."). PersonalWeb Technologies, LLC and Level 3 Communications, LLC (collectively, "PersonalWeb") timely filed a Patent Owner's Preliminary Response. Paper 15 ("Prelim. Resp."). Taking into account PersonalWeb's Preliminary Response, the Board determined that the information presented in the Petition demonstrated that there was a reasonable likelihood that EMC would prevail in challenging claims 1-4, 29-33, and 41 as unpatentable under 35 U.S.C. §§ 102(e) and 103(a). Pursuant to 35 U.S.C. § 314, the Board instituted this proceeding on May 17, 2013, as to the challenged claims of the '791 patent. Paper 21 ("Dec.").

During this proceeding, PersonalWeb timely filed a Patent Owner Response (Paper 47, "PO Resp."), and EMC timely filed a Reply to the Patent Owner Response (Paper 55, "Reply"). A consolidated oral hearing was held on December 16, 2013.[1]

We have jurisdiction under 35 U.S.C. § 6(c). This decision is a final written decision under 35 U.S.C. § 318(a) as to the patentability of the challenged claims. Based on the record before us, EMC has demonstrated

[1] This proceeding, as well as IPR2013-00083, IPR2013-00084, IPR2013-00085, IPR2013-00086, and IPR2013-00087, involve the same parties and similar issues. The oral arguments for all six *inter partes* reviews were merged and conducted at the same time. A transcript of the oral hearing is included in the record as Paper 82.

by a preponderance of the evidence that claims 1-4, 29-33, and 41 are unpatentable.

A. The Invention of the '791 Patent

The invention of the '791 patent relates to a data processing system that identifies data items using substantially unique identifiers, otherwise referred to as True Names, which depend on all the data in the data item and only on the data in the data item. Ex. 1001, 1:14-18, 3:29-32, and 6:6-10. According to the '791 patent, the identity of a data item depends only on the data and is independent of the data item's name, origin, location, address, or other information not directly derivable from the data associated therewith. Ex. 1001, 3:33-35. The invention of the '791 patent also examines the identities of a plurality of data items in order to determine whether a particular data item is present in the data processing system. Ex. 1001, 3:36-39.

B. Illustrative Claims

Claims 1, 30, and 33 are independent claims. Claims 2-4 and 29 depend directly or indirectly from independent claim 1. Claims 31, 32, and 41 depend directly or indirectly from independent claim 30. Independent claims 1, 30, and 33 are illustrative of the invention of the '791 patent and are reproduced below:

> 1. In a data processing system, an apparatus comprising:
>
> identity means for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier; and

existence means for determining whether a particular data item is present in the system, by examining the identifiers of the plurality of data items.

Ex. 1001, 39:14-23.

30. A method of identifying a data item present in a data processing system for subsequent access to the data item, the method comprising:

determining a substantial unique identifier for the data item, the identifier depending on and being determined using all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier; and

accessing a data item in the system using the identifier of the data item.

Ex. 1001, 42:58-67.

33. A method of duplicating a given data item present at a source location to a destination location in a data processing system, the method comprising:

determining a substantially unique identifier for the given data item, the identifier depending on and being determined using all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier;

determining, using the data identifier, whether the data item is present at the destination location; and

based on the determining whether the data item is present, providing the destination location with the data item only if the data item is not present at the destination.

Ex. 1001, 43:11-23.

C. Related Proceedings

EMC indicates that the ’791 patent was asserted against it in *PersonalWeb Technologies LLC v. EMC Corporation and VMware, Inc.*, No. 6:11-cv-00660-LED, pending in the United States District Court for the

Eastern District of Texas. Pet. 1. EMC also filed five other petitions seeking *inter partes* review of the following patents: (1) U.S. Patent No. 6,415,280 (*EMC Corp. and VMware, Inc. v. PersonalWeb Techs., LLC*, IPR2013-00083); (2) U.S. Patent No. 7,945,544 (*EMC Corp. v. PersonalWeb Techs., LLC*, IPR2013-00084); (3) U.S. Patent No. 7,945,539 (*EMC Corp. v. PersonalWeb Techs., LLC*, IPR2013-00085); (4) U.S. Patent No. 7,949,662 (*EMC Corp. v. PersonalWeb Techs.*, LLC, IPR2013-00086); and (5) U.S. Patent No. 8,001,096 (*EMC Corp. v. PersonalWeb Techs., LLC*, IPR2013-00087). *Id*.

D. Prior Art Relied Upon

EMC relies upon the following prior art reference:

Woodhill	US 5,649,196	July 15, 1997 (effectively filed July 1, 1993)	Ex. 1005

E. Grounds of Unpatentability

We instituted this proceeding based on the grounds of unpatentability set forth in the table below.

Claims	**Basis**	**Reference**
1-4, 29-33, and 41	§ 102(e)	Woodhill
1-4 and 29	§ 103(a)	Woodhill

II. ANALYSIS

A. Claim Construction

In an *inter partes* review, we construe a claim by applying the broadest reasonable interpretation in light of the specification of the patent in which it appears. 37 C.F.R. § 42.100(b); *see* Office Patent Trial Practice Guide, 77 Fed. Reg. 48,756, 48,766 (Aug. 14, 2012). Under the broadest

reasonable interpretation standard, claim terms are given their ordinary and customary meaning as would be understood by one of ordinary skill in the art in the context of the entire disclosure. *In re Translogic Tech. Inc.*, 504 F.3d 1249, 1257 (Fed. Cir. 2007). We must be careful not to read limitations from a particular embodiment appearing in the specification into the claim if the claim language is broader than that embodiment. *In re Van Geuns*, 988 F.2d 1181, 1184 (Fed. Cir. 1993). If a feature in the disclosure is not necessary to give meaning to what the inventor means by a claim term, it would be "extraneous" and, therefore, should not be read into the claim. *Renishaw PLC v. Marposs Societa' per Azioni*, 158 F.3d 1243, 1249 (Fed. Cir. 1998); *E.I. du Pont de Nemours & Co. v. Phillips Petroleum Co.*, 849 F.2d 1430, 1433 (Fed. Cir. 1988).

In its Petition, EMC identified five claim terms and provided a claim construction for those terms. Pet. 4-6. Those claim terms are listed as follows: (1) "substantially unique identifier"; (2) "using the identifier"; (3) "data" and "data item"; (4) "location"; and (5) "True Name, data identity, and data identifier." *Id*. In the Decision to Institute, we construed each claim term identified by EMC. Dec. 13-16.

In its Petition, EMC also identified several means-plus-function limitations that invoke 35 U.S.C. § 112, ¶ 6, and their corresponding structure for performing the claimed function. Pet. 6-8. Those means-plus-function limitations are listed as follows: (1) "identity means for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same

identifier"; (2) "existence means for determining whether a particular item is present in the system, by examining the identifiers of the plurality of data items"; (3) "local existence means for determining whether an instance of a particular data item is present at a particular location in the system, based on the identifier of the data item"; (4) "data associating means for making and maintaining, for a data item in the system, an association between the data item and the identifier of the data item"; and (5) "access means for accessing a particular data item using the identifier of the data item." In the Decision to Institute, we construed each means-plus-function limitation identified by EMC to cover the corresponding structure described in the specification of the '791 patent and equivalents thereof. Dec. 17-26, *see* 35 U.S.C. § 112, ¶ 6.

With one exception, PersonalWeb agrees with our claim constructions in the Decision to Institute. PO Resp. 1-3 (quoting Dec. 13-16, 20-25). PersonalWeb proposes an alternative claim construction for the following means-plus-function limitation recited in independent claim 1:

> identity means for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier.

Id. at 13-14 (quoting Ex. 1001, 39:16-21).

We will address PersonalWeb's alternative claim construction for this means-plus-function limitation below.

> 1. "*identity means for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier" (Claim 1)*

As we indicated in the Decision to Institute, both parties agreed that the claimed function of this means-plus-function limitation is "determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier." Dec. 18 (quoting Ex. 1001, 39:16-21). We then identified a data processor programmed to perform a hash function, e.g., MD5 or SHA, as the corresponding structure for performing the claimed function. *Id*. at 20.

In its Patent Owner Response, PersonalWeb contends that our construction of the corresponding structure for performing the claimed function is overly broad. PO Resp. 14. PersonalWeb argues that the specification of the '791 patent discloses that the identity calculating mechanism "must" have at least five properties, and "must" be employed on a system wide basis. *Id*. (citing Ex. 1001, 12:61-13:9, 13:15-19). PersonalWeb further argues that the specification of the '791 patent describes corresponding structure that is necessary to determine a substantially unique identifier "for any of a plurality of data items present in the system," as claimed. *Id*. (citing Ex. 1001, 14:12-31). PersonalWeb asserts that our claim construction in the Decision to Institute does not reflect the aforementioned features disclosed in the specification of the '791

patent that are necessary to perform the claim function. *Id*. Based on those arguments, PersonalWeb identifies at least one processor programmed to perform the Calculate True name mechanism as the corresponding structure for performing the claimed function. *Id*. (citing Ex. 1001, 7:62-63, 12:54-13:19, 14:1-39).

To the extent that PersonalWeb argues that we failed to consider the use of the term "must" in the specification of the '791 patent when we previously construed the corresponding structure that performs the claimed function, we disagree. As we explained above, PersonalWeb directs us to the disclosure in the specification of the '791 patent that states, "[t]he function MD *must* have the following properties . . . [and] [t]hese functions (or algorithms) include MD4, MD5, and SHA." Ex. 1001, 12:61-13:14 (emphasis added). PersonalWeb also directs us to the disclosure in the specification of the '791 patent that states "[i]n the presently preferred embodiments, either MD5 or SHA is employed as the basis for the computation of True Names. Whichever of these two message digest functions is employed, the same function *must* be employed on a system-wide basis." Ex. 1001, 13:15-19 (emphasis added).

Although the specification of the '791 patent uses the absolute term "must" when describing MD hash functions generally, and hash functions MD5 and SHA specifically, it nonetheless describes these hash functions in the context of "preferred embodiments." Therefore, we included the hash functions MD5 and SHA in our claim construction as examples only. Dec. 18-20. Our determination in that regard incorporates the features necessary to perform the claimed function, yet does not conflict with the principle that this means-plus-function limitation is to be given its broadest

reasonable interpretation. *In re Donaldson*, 16 F.3d 1189, 1194 (Fed. Cir. 1994). In other words, we did not view hash functions MD5 and SHA as "necessary" to perform the claimed function because they were part of the preferred embodiments disclosed in the specification of the '791 patent. Claim interpretation under § 112, ¶ 6, does not "permit incorporation of structure from the written description beyond that necessary to perform the claimed function." *Micro Chem., Inc. v. Great Plains Chem. Co.*, 194 F.3d 1250, 1258 (Fed. Cir. 1999).

We also are not persuaded by PersonalWeb's argument that the specification of the '791 patent describes corresponding structure that is necessary to determine a substantially unique identifier "for any of a plurality of data items present in the system," as claimed. The alleged corresponding structure referenced in PersonalWeb's argument is the embodiment illustrated in Figure 10(b), which is a flowchart depicting the operations associated with calculating the True Name of an arbitrary, i.e., simple or compound, data item. Ex. 1001, 14:1-3, 13-15. PersonalWeb does not explain adequately why the steps illustrated in that embodiment are necessary to perform the claimed function, nor does PersonalWeb explain why such steps must be part of the algorithm that provides the necessary structure under § 112, ¶ 6. We do not find the steps illustrated in Figure 10(b) as "necessary" to perform the claimed function and, therefore, such steps should not be read into the corresponding structure for performing the claimed function. *See Micro Chem.*, 194 F.3d at 1258.

Applying the broadest reasonable interpretation standard, we maintain that the corresponding structure identified in the specification of the '791 patent for performing the claimed function of "determining, for any of a

plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier" is a data processor programmed to perform a hash function, e.g., MD5 or SHA.

B. The Level of Ordinary Skill in the Art

In determining the level of one with ordinary skill in the art, we note that various factors may be considered, including "type of problems encountered in the art; prior art solutions to those problems; rapidity with which innovations are made; sophistication of the technology; and educational level of active workers in the field." *In re GPAC*, 57 F.3d 1573, 1579 (Fed. Cir. 1995) (citing *Custom Accessories, Inc. v. Jeffrey-Allan Indus., Inc.*, 807 F.2d 955, 962 (Fed. Cir. 1986)). There is sufficient evidence in the record before us that reflects the knowledge level of a person with ordinary skill in the art. PersonalWeb's expert, Dr. Robert B.K. Dewar, attests that a person with ordinary skill in the art would be an individual with a bachelor's degree in computer science who possesses ten to fifteen years of teaching or work experience in the field of data processing systems. Ex. 2013 ¶ 18.

C. Anticipation by Woodhill—Claims 1-4, 29-33, and 41

EMC contends that claims 1-4, 29-33, and 41 are anticipated under § 102(e) by Woodhill. Pet. 51-59. In support of that alleged ground of unpatentability, EMC provides explanations as to how Woodhill describes each claim limitation. *Id*. (citing Ex. 1041). EMC also submits the declarations of Dr. Douglas W. Clark (Ex. 1009 ¶¶ 81-95; Ex. 1081) to support its positions. Upon reviewing EMC's Petition and supporting

evidence, as well as PersonalWeb's Patent Owner Response and supporting evidence, we determine that EMC has demonstrated by a preponderance of the evidence that claims 1-4, 29-33, and 41 are anticipated by Woodhill.

We begin our analysis with the principles of law that generally apply to a ground of unpatentability based on anticipation, followed by a brief discussion of Woodhill, and then we turn to the arguments presented by both EMC and PersonalWeb that are directed towards each challenged claim.

1. Principles of Law

To establish anticipation under § 102(e), "all of the elements and limitations of the claim must be shown in a single prior reference, arranged as in the claim." *Karsten Mfg. Corp. v. Cleveland Golf Co.*, 242 F.3d 1376, 1383 (Fed. Cir. 2001). "A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference." *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631 (Fed. Cir. 1987). We analyze the ground of unpatentability based on anticipation by Woodhill with the above-stated principles in mind.

2. Woodhill

Woodhill generally relates to a system and method for distributed storage management on a networked computer system that includes a remote backup file server in communication with one or more local area networks. Ex. 1005, 1:11-17. Figure 1 of Woodhill, which is reproduced below, illustrates networked computer system 10. Ex. 1005, 2:56-58.

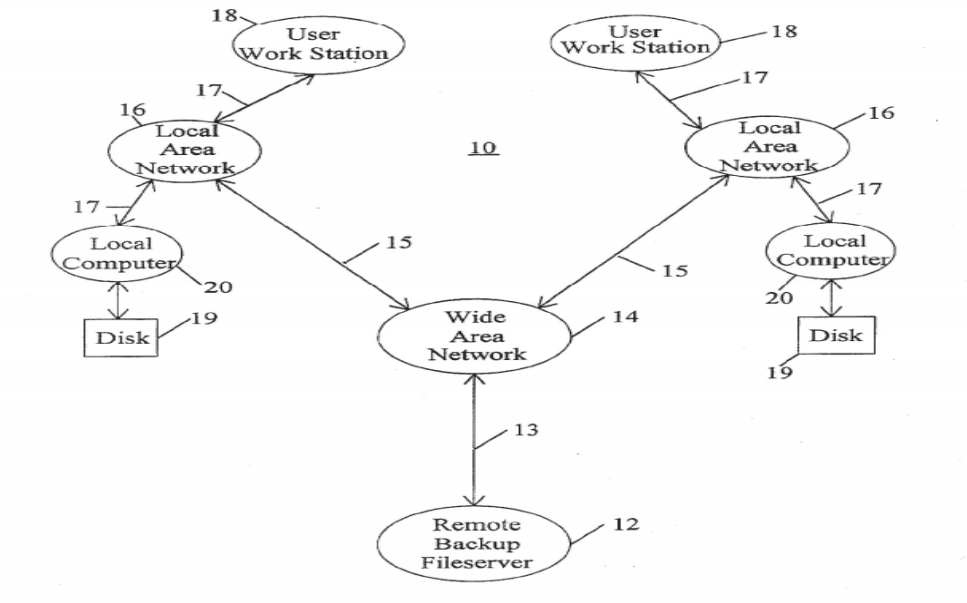

FIG. 1

As shown in Figure 1 of Woodhill, remote backup file server 12 communicates with wide area network 14 via data path 13, wide area network 14 communicates with a plurality of local area networks 16 via data paths 15, and each local area network 16 communicates with multiple user workstations 18 and local computers 20 via data paths 17. Ex. 1005, 3:12-31. The storage space on each disk drive 19 on each local computer 20 is allocated according to the hierarchy illustrated in Figure 2. Ex. 1005, 3:31-44.

Figure 2 of Woodhill, which is reproduced below, illustrates Distributed Storage Manager program 24 that allocates storage space on each of the storage devices in networked computer system 10. Ex. 1005, 2:59-62.

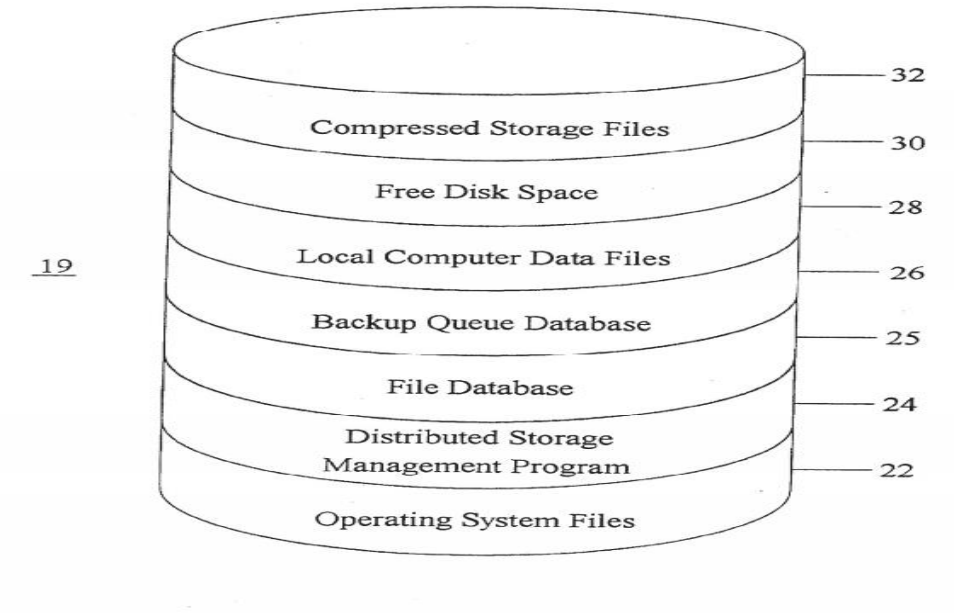

FIG. 2

As shown in Figure 2 of Woodhill, Distributed Storage Manager program 24 builds and maintains File Database 25 on the one or more disk drives 19 on each local computer 20 in networked computer system 10. Ex. 1005, 3:45-49. Distributed Storage Manager program 24 views a file as a collection of data streams. Ex. 1005, 4:13-15. Woodhill defines a data stream as a distinct collection of data within a file that may change independently from other distinct collections of data within the file. Ex. 1005, 4:15-18. For instance, Woodhill discloses that a file may contain both its normal data and any extended attribute data. Ex. 1004, 4:18-19. Depending on the size of the data stream, Distributed Storage Manager program 24 divides each data stream into one or more binary objects. Ex. 1005, 4:21-30.

Figure 3 of Woodhill, which is reproduced below, illustrates File Database 25 used by Distributed Storage Manager program 24. Ex. 1005, 2:63-64.

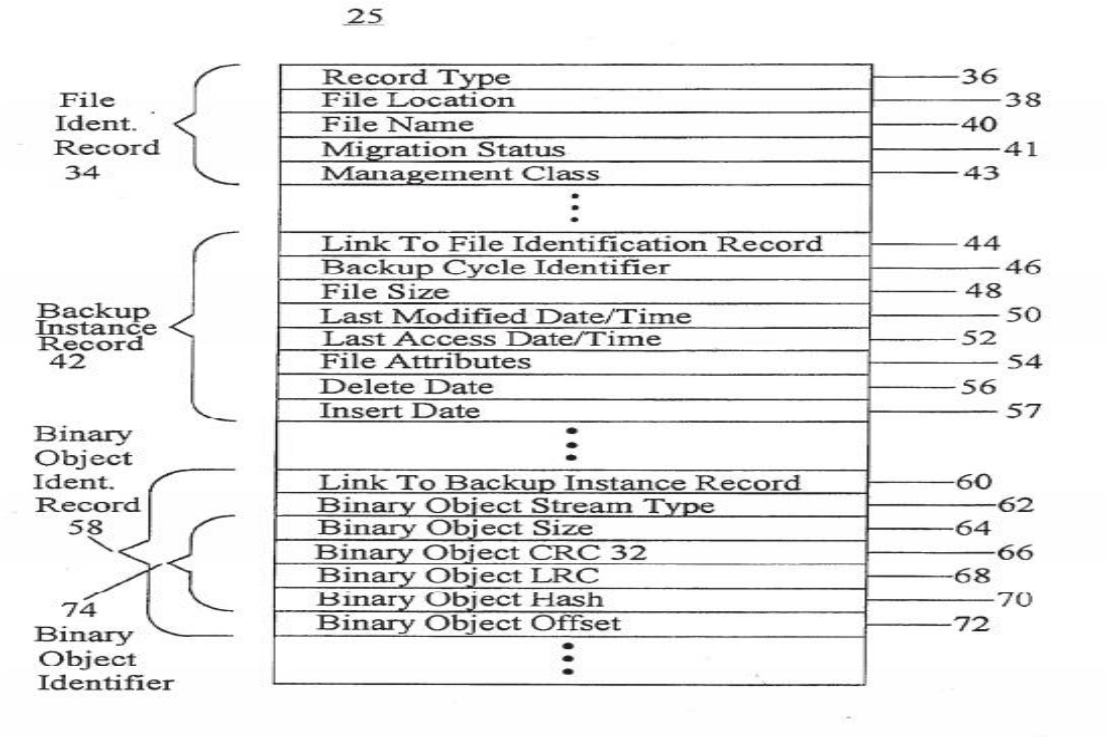

FIG. 3

As shown in Figure 3 of Woodhill, File Database 25 includes the following three levels of records organized according to a predefined hierarchy: (1) File Identification Record 34; (2) Backup Instance Record 42; and (3) Binary Object Identification Record 58. Ex. 1005, 3:54-4:47. Binary Object Identification Record 58 includes, amongst other things, Binary Object Identifier 74 that comprises Binary Object Size 64, Binary Object CRC32 66, Binary Object LRC 68, and Binary Object Hash 70. Ex. 1005, 4:45-47, 7:64-8:1. Binary Object Identifier 74 is a unique identifier for each binary object that is backed up. Ex. 1005, 4:45-47.

Although Woodhill discloses calculating Binary Object Identifier 74 in various ways, e.g., using a binary hash algorithm (Ex.1005, 8:1-31), the key notion is that Binary Object Identifier 74 is calculated from the content of the data instead of from an external or arbitrary source. Ex. 1005, 8:38-

42. In other words, Woodhill recognizes that the critical feature in creating Binary Object Identifier 74 is that the identifier should be based on the contents of the binary object, such that Binary Object Identifier 74 changes when the contents of the binary object changes. Ex. 1005, 8:58-62. Therefore, duplicate binary objects, even if resident on different types of computers in the network, may be recognized by their identical Binary Object Identifiers 74. Ex. 1005, 8:62-65.

Woodhill discloses that Distributed Storage Manager program 24 performs two backup operations concurrently. Ex. 1005, 9:30-31. First, Distributed Storage Manager program 24 stores a compressed copy of each binary object that it needs to restore disk drive 19 on each local computer 20 somewhere on local area network 16 other than on local computer 20 where the binary object originally resided. Ex. 1005, 9:31-36. Second, Distributed Storage Manager program 24 transmits new or changed binary objects to remote backup file server 12. Ex. 1005, 9:36-38.

Woodhill also discloses that Distributed Storage Manager program 24 performs auditing and reporting functions on a periodic basis to ensure that binary objects, which already have been backed up, may be restored. Ex. 1005, 18:11-13. Distributed Storage Manager program 24 initiates a restore of a randomly selected binary object identified by a Binary Object Identification Record 58 stored in File Database 25. Ex. 1005, 18:16-19.

3. Claim 1

a. "determining whether a particular data item is present in the system"

Independent claim 1 recites, in relevant part, an "existence means for *determining whether a particular data item is present in the system*, by

examining the identifiers of the plurality of data items." Ex. 1001, 39:21-23 (emphasis added).

In its Petition, EMC contends that Woodhill determines the existence of a binary object by examining Binary Object Identifiers 74 on local computers 20, which are part of local area networks 16, and on remote backup file server 12. Pet. 57 (citing Ex. 1009 ¶¶ 85-87; Ex. 1005, 8:62-9:23). For example, EMC argues that Woodhill uses Binary Object Identifier 74 to check whether a binary object has changed since it was last backed up, as well as to check whether a local copy of a binary object is available to be restored. *Id*. (citing Ex. 1009 ¶ 86; Ex. 1005, 9:14-22).

In its Patent Owner Response, PersonalWeb contends that Woodhill only determines whether a particular Binary Object Identifier 74 for a binary object is present for the most-recently backed up version of a single file, and therefore, Woodhill cannot "determine[] whether a particular data item is present in the system," as claimed. PO Resp. 4-7. In particular, PersonalWeb argues that the ability to determine whether a particular file is present in the system requires the ability to look at information for more than one file. *Id*. at 5. PersonalWeb further argues that one would need the ability to look at information regarding all the files in a system in order to determine if a particular file is present in the system. *Id*. PersonalWeb alleges that Woodhill cannot determine whether a particular binary object is present in its system because it is incapable of determining if that particular binary object is present in any of the thousands, if not millions, of files in the system. *Id*. at 8. PersonalWeb relies upon the declaration of Robert B.K. Dewar to support its positions. Ex. 2013 ¶¶ 21-28.

In its Reply, EMC contends that PersonalWeb's argument is predicated on the notion that one would need to have the ability to look at all the files in a system in order to determine whether a file is present in the system. Reply 1-2. EMC argues that independent claim 1 is not that specific, but instead it generally refers to determining whether a data item is present in the system. *Id*. at 2. EMC directs our attention to related district court litigation, where the court was not persuaded by a similar argument presented by PersonalWeb. *Id*. at 2 (citing Ex. 1074, 35). We are not persuaded by PersonalWeb's arguments because they are based on an overly narrow claim construction.

As we explained in the Decision to Institute, we identified the corresponding structure for performing the claimed function of "determining whether a particular data item is present in the system, by examining the identifiers of the plurality of data items" to be a data processor programmed according to step S232 illustrated in Figure 11 or step S260 illustrated in Figure 14. Dec. 20-22. With respect to step S232 illustrated in Figure 11, the specification of the '791 patent discloses "look[ing] for an entry for the True Name in the True File registry 126 (Step S232) and determin[ing] whether a True Name entry, record 140, exists in the True file registry 126." Ex. 1001, 14:53-56. With respect to step S260 illustrated in Figure 14 of the '791 patent, the specification of the '791 patent discloses, "if desired, confirm[ing] that the True Name exists locally by searching for it in the True Name registry or local directory extensions table 135 (Step S260)." Ex. 1001, 15:54-56.

Contrary to PersonalWeb's arguments, the claimed function of "determining whether a particular data item is present in the system" does

not encompass searching all the files in a system. Instead, according to the specification of the '791 patent, it simply includes determining whether a file exists in a registry or table. *See, e.g.*, Ex. 1001, 14:53-56, 15:54-56, fig. 11, step S232, fig. 14, step S260. Woodhill recognizes duplicate binary objects residing on different types of computers in the network by their identical Binary Object Identifiers 74. Ex. 1005, 8:62-65. During Woodhill's backup procedure, Binary Object Identifiers 74 are calculated for each binary object and then compared against their counterparts in File Database 25. Ex. 1005, Ex. 9:14-16. For example, Woodhill discloses that Distributed Storage Manager program 24 compares a newly calculated Binary Object Identifier 74 for a particular binary object with Binary Object Identifier 74 associated with the most recent version of that binary object. Ex. 1005, 9:16-22. Dr. Clark testifies that this comparison of Binary Object Identifiers 74 is just one relevant example of determining whether and where a particular binary object is present in its system. Ex. 1009 ¶ 86.

In summary, we agree with EMC that Woodhill's backup procedure, which includes calculating Binary Object Identifiers 74 for each binary object and then comparing them against their counterparts in File Database 25, describes the function of "determining whether a particular data item is present in the system," as recited in independent claim 1.

b. "existence means"

Independent claim 1 recites, in relevant part, an "*existence means* for determining whether a particular data item is present in the system, by examining the identifiers of the plurality of data items." Ex. 1001, 39:21-23 (emphasis added).

In its Patent Owner Response, PersonalWeb relies upon essentially the same argument presented above with respect to the claimed function of "determining whether a particular data item is present in the system." *Compare* PO Resp. 4-10 *with* PO Resp. 10-13. That is, PersonalWeb contends that Woodhill does not search for the newly calculated Binary Object Identifier 74 in a registry or table that includes a plurality of Binary Object Identifiers 74 associated with different files in the system. *Id*. at 11-12. Therefore, PersonalWeb asserts that Woodhill does not disclose a structure tantamount to the corresponding structure for the claimed "existence means" because Woodhill does not perform the identical function in substantially the same way to achieve substantially the same results. *Id*.

In its Reply, EMC contends that Woodhill's File Database 25 is equivalent to True file registry 126 described in the specification of the '791 patent. Reply 4 (citing Ex. 1001, 9:36-67). EMC reiterates that we properly identified the corresponding structure for the "existence means" in the Decision to Institute (Dec. 20-22), and then contends that both the '791 patent and our construction simply require confirming whether an identifier exists in a database that has a plurality of identifiers. Reply 4-5.

As we explained above, the claimed function associated with the "existence means" simply encompasses determining whether a file exists in a registry or table. *See, e.g.*, Ex. 1001, 14:53-56, 15:54-56, fig. 11, step S232, fig. 14, step S260. It does not include searching all the files in a system. According to Woodhill, both its system and method for managing storage space on network computer system 10 include comparing the current value of the binary object identifier associated with a particular binary object to *one or more* previous values of the binary object identifier associated with

that particular binary object. Ex. 1005, 2:14-17, 33-36 (emphasis added); *see also* Ex. 2007, 22.[2] Independent claim 1 of Woodhill further recites, in relevant part, "means for comparing said current name of a particular binary object to *one or more* previous names of said binary object." Ex. 1005, 22:5-7 (emphasis added); *see also* Ex. 2007, 62 (originally presented independent claim 1). In our view, these disclosures in Woodhill apply to its backup procedure and, in particular, support a finding that File Database 25 stores a plurality of Binary Object Identifiers 74 associated with different binary objects or files that have been backed up in the system (*see, e.g.*, Ex. 1005, 3:49-52, 4:30-34, 9:8-22). *Cf. In re Preda*, 401 F.2d 825, 826 (CCPA 1968) ("[I]t is proper to take into account not only specific teachings of the references but also the inferences which one skilled in the art would reasonably be expected to draw therefrom.").

As EMC explains in its Reply, Dr. Clark testifies that Woodhill determines whether a binary object or file is present in the system by confirming that its Binary Object Identifier 74 already exists among the plurality of Binary Object Identifiers 74 stored in File Database 25. Reply 5 (citing Ex. 1081 ¶ 14). We credit Dr. Clark's testimony in that regard because it is consistent with Woodhill's summary of its own invention, the broader disclosure provided by independent claim 1, and the description of File Database 25.

---

[2] Exhibit 2007 includes excerpts from the file history of Woodhill. PersonalWeb did not provide any page numbers for this Exhibit. For purposes of this decision, page 1 is the page that includes "Exhibit PersonalWeb 2007" in the lower, right-hand corner. The remaining pages are numbered consecutively therefrom.

In light of our analysis above, we agree with EMC that Woodhill discloses a structure equivalent to the corresponding structure for the claimed "existence means" because it performs an identical function in substantially the same way to achieve substantially the same results. *See, e.g.*, *Odetics, Inc. v. Storage Tech. Corp.*, 185 F.3d 1259, 1267 (Fed. Cir. 1999). In other words, Woodhill's backup procedure, which includes calculating Binary Object Identifiers 74 for each binary object and then comparing them against a plurality of Binary Object Identifiers 74 stored in File Database 25, describes the function of "determining whether a particular data item is present in the system, by examining the identifiers of the plurality of data items," as recited in independent claim 1.

c. "identity means"

Independent claim 1 recites, in relevant part,

> *Identity means* for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier.

Ex. 1001, 39:16-21.

In its Petition, EMC contends that Woodhill's Binary Object Identifiers 74 constitute the claimed "substantially unique identifiers" determined using the contents in the binary object. Pet. 56-57 (citing Ex. 1005, 7:60-8:1, fig. 3; Ex. 1009 ¶¶ 83-84). EMC argues that, when calculating the Binary Object Identifiers 74, Woodhill uses "all of" the data of a binary object and "only" the data of the binary object. *Id*. at 57 (citing Ex. 1009 ¶ 84; Ex. 1005, 8:1-31). EMC argues that two identical binary

objects in Woodhill's system will have the same Binary Object Identifier 74 because each Binary Object Identifier 74 is based on the data of the binary object associated therewith. *Id*.

In its Patent Owner Response, PersonalWeb proposes an alternative claim construction for the claimed "identity means," and then contends that Woodhill fails to disclose the corresponding structure identified in its alternative claim construction. *Id*. at 14-15. In particular, PersonalWeb argues that, when the '791 patent determines whether a data item is compound, the claimed "identity means" requires a cryptographic hash of cryptographic hashes ("a hash of hashes"). *Id*. PersonalWeb argues that, although Woodhill discloses calculating Binary Object Identifier 74 for a binary object by applying a hash function to the binary object, it does not apply a hash function to Binary Object Identifier 74, itself. *Id*. In addition, PersonalWeb contends that Woodhill does not disclose a cryptographic hash, such as MD5, SHA, or anything equivalent thereto. *Id*.

As we explained previously, we disagree with the alternative claim construction for "identity means" proposed by PersonalWeb in its Patent Owner Response. To the extent PersonalWeb now argues that the claimed "identity means" requires a hash of hashes, we also disagree. Similar to our explanation in the Decision to Institute, PersonalWeb's argument in that regard is not commensurate in scope with our claim construction of "identity means." *See* Dec. 27-28. The corresponding structure for performing the claimed function of "determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will

have the same identifier" is a data processor programmed to perform a hash function, e.g., MD5 or SHA. Neither the specification of the '791 patent, nor the claim itself, indicates that the "identity means" requires determining a substantial unique identifier for a compound data item, much less using a hash of hashes when determining whether a data item is compound.

We also are not persuaded by PersonalWeb's argument that Woodhill does not disclose a cryptographic hash, such as an MD5, SHA, or anything equivalent thereto. Woodhill discloses various ways to calculate Binary Object Identifier 74 for a particular binary object, including using a binary hash algorithm. Ex.1005, 8:1-31. The key notion in Woodhill is that Binary Object Identifier 74 is calculated based on the content of each binary object instead of from an external or arbitrary source. Ex. 1005, 8:38-42. In other words, Woodhill creates Binary Object Identifier 74 for a binary object based on the contents of the binary object, such that Binary Object Identifier 74 changes when the contents of the binary object changes. Ex. 1005, 8:58-62. Based on these cited disclosures, Woodhill's binary hash algorithm relies on "all of" the data of a binary object and "only" the data of the binary object when calculating Binary Object Identifier 74.

In summary, we agree with EMC that Woodhill's disclosure of calculating Binary Object Identifier 74 for a particular binary object describes the corresponding structure for performing the claimed function associated with the "identity means," as recited in independent claim 1. For the foregoing reasons, we conclude that that EMC has demonstrated by a preponderance of the evidence that independent claim 1 is anticipated by Woodhill.

4. Claim 2

a. "determining whether . . . a particular data item is present at a particular location in the system"

Dependent claim 2 recites a "local existence means for *determining whether* an instance of *a particular data item is present at a particular location in the system*, based on the identifier of the data item." Ex. 1001, 39:25-29 (emphasis added).

In its Petition, EMC contends that Woodhill discloses a "local existence means," as claimed. Pet. 58 (citing Ex. 1005, 9:5-23). In particular, EMC argues that Woodhill's remote backup server 12 constitutes the claimed "particular location in the system." Ex. 1041, 15. EMC also offers the testimony of Dr. Clark to support its position. Ex. 1009 ¶¶ 88, 89.

In its Patent Owner Response, PersonalWeb presents essentially the same arguments discussed above with respect to independent claim 1. PO Resp. 16-17. That is, PersonalWeb argues that Woodhill only determines whether Binary Object Identifier 74 for a particular binary object is present for the most-recent version of a particular file at remote backup server 12, and does not determine whether that particular file is present in the many other files stored at remote backup server 12. *Id*.

As we explained in the Decision to Institute, we identified the corresponding structure for performing the claimed function of "determining whether an instance of a particular data item is present at a particular location in the system, based on the identifier of the data item" to be a data processor programmed according to step S260 illustrated in Figure 14. Dec. 21-22. With respect to step S260 illustrated in Figure 14, the specification of the '791 patent discloses, "if desired, confirm[ing] that the

True Name exists locally by searching for it in the True Name registry or local directory extensions table 135 (Step S260)." Ex. 1001, 15:54-56.

Contrary to PersonalWeb's arguments, the claimed function of "determining whether . . . a particular data item is present at a particular location in the system" does not encompass searching all the files in a system. Instead, according to the specification of the '791 patent, it simply includes determining whether a file exists in a registry or table. *See, e.g.*, Ex. 1001, 15:54-56, fig. 14, step S260.

To support its position regarding dependent claim 2, EMC, once again, directs our attention to Woodhill's backup procedure. During Woodhill's backup procedure, Distributed Storage Manager program 24 determines whether a particular binary object has changed using the version of the binary object that previously was backed up. Ex. 1005, 9:6-9. Dr. Clark testifies that File Database 25 contains a list of Binary Object Identifiers 74 for binary objects recently backed up and stored in the system, including the binary objects backed up and stored in remote backup file server 12. Ex. 1009 ¶ 89. Dr. Clark also testifies that, when comparing Binary Object Identifier 74 calculated during the current backup cycle with those stored in File Database 25, Distributed Storage Manager program 24 essentially determines the existence, at remote backup file server 12, of the particular binary object being processed. Ex. 1009 ¶ 89; *see also* Ex. 1081 ¶¶ 5-8. We credit Dr. Clark's testimony because it is consistent with Woodhill's description of the backup procedure.

PersonalWeb further contends that, although it agrees with our claim construction of the claim term "location," a single file in Woodhill does not constitute the claimed "location." PO Resp. 17 (citing Dec. 15-16; Ex. 2013

¶ 39). PersonalWeb's argument that a single file in Woodhill does not constitute the claimed "particular location in the system" is misplaced. There is no indication in the record before us that EMC takes the position that Woodhill's disclosure of a single file constitutes the claimed "particular location in the system." Instead, as explained above, EMC takes the position that Woodhill's disclosure of remote backup server 12 constitutes the claimed "particular location in the system." Ex. 1041, 15.

In summary, we agree with EMC that Woodhill's backup procedure, which includes examining Binary Object Identifiers 74 stored in File database 25 to determine if the most recent version of a binary object is present at remote backup file server 12, describes the function of "determining whether . . . a particular data item is present at a particular location in the system," as recited in dependent claim 2.

b. "local existence means"

In its Patent Owner Response, PersonalWeb relies upon essentially the same arguments presented above with respect the claimed function of this means-plus-function limitation, "determining whether . . . a particular data item is present at a particular location in the system," to rebut EMC explanations as to how Woodhill describes the claimed "local existence means." *Compare* PO Resp. 16-18 *with* PO Resp. 18-20. That is, PersonalWeb contends that Woodhill does not search for the newly calculated Binary Object Identifier 74 in a registry or table that includes a plurality of Binary Object Identifiers 74 associated with different files in the system. PO Resp. 19. Therefore, PersonalWeb asserts that Woodhill does not disclose a structure tantamount to the corresponding structure for the claimed "local existence means" because Woodhill does not perform the

identical function in substantially the same way to achieve substantially the same results. *Id*.

As we explained above, the claimed function associated with the "local existence means" simply encompasses determining whether a file exists in a registry or table. *See, e.g.*, Ex. 1001, 15:54-56, fig. 14, step S260. It does not include searching all the files in a system. Moreover, Woodhill discloses that File Database 25 stores a plurality of Binary Object Identifiers 74 associated with different binary objects or files that have been backed up in the system. *See*, *e.g.*, Ex. 1005, 3:49-52, 4:30-34, 9:8-22. Dr. Clark testifies that Woodhill determines whether a binary object or file is present at remote backup file server 12 by confirming that its Binary Object Identifier 74 already exists among the plurality of Binary Object Identifiers 74 stored in File Database 25. Ex. 1009 ¶ 89; *see also* Ex. 1081 ¶ 14. We credit Dr. Clark's testimony in that regard because it is consistent with Woodhill's description of File Database 25, as well as its description of the backup procedure.

In light of our analysis above, we agree with EMC that Woodhill discloses a structure tantamount to the corresponding structure for the claimed "local existence means" because it performs an identical function in substantially the same way to achieve substantially the same results. *See, e.g.*, *Odetics*, 185 F.3d at 1267. In other words, Woodhill's backup procedure, which includes examining Binary Object Identifiers 74 stored in File database 25 to determine if the most recent version of a binary object is present at remote backup file server 12, describes the function of "determining whether an instance of a particular data item is present at a particular location in the system, based on the identifier of the data item," as

recited in dependent claim 2. For the foregoing reasons, we conclude that that EMC has demonstrated by a preponderance of the evidence that dependent claim 2 is anticipated by Woodhill.

5. Claim 3

a. "examining the identifiers of the plurality of data items at said particular location in the system"

Dependent claim 3 recites, in relevant part, "wherein said local existence means for determining whether a particular data item is present at a particular location in the system by *examining the identifiers of the plurality of data items at said particular location in the system*." Ex. 1001, 39:31-35 (emphasis added).

In its Petition, EMC contends that Woodhill discloses the "local existence means," as claimed. Pet. 58 (citing Ex. 1005, 9:5-23). In particular, EMC argues that Woodhill's Distributed Storage Manager program 24, which executes on a computer, accesses and checks Binary Object Identification Records 58 in File Database 25 to determine whether a local copy of a particular binary object is present on the local system before restoring a remote copy. Ex. 1041, 15. According to EMC, Woodhill's Distributed Storage Manager program 24 performs this function by examining Binary Object Identifiers 74 for the plurality of binary objects stored in File Database 25. *Id.* EMC also offers the testimony of Dr. Clark to support its position. Ex. 1009 ¶¶ 90, 91.

In its Patent Owner Response, PersonalWeb reiterates that, during Woodhill's backup procedure, Distributed Storage Manager program 24 compares Binary Object Identifier 74 for a newly processed binary object with only a single prior version of Binary Object Identifier 74. PO Resp. 21-

22. PersonalWeb maintains that Woodhill does not compare the newly created Binary Object Identifier 74 for a binary object or file with a plurality of previous Binary Object Identifiers 74 for that file. *Id*. at 21. Once again, we are not persuaded by PersonalWeb's argument.

Woodhill discloses that File Database 25 stores a plurality of Binary Object Identifiers 74 associated with different binary objects or files that have been backed up in the system. *See, e.g.*, Ex. 1005, 3:49-52, 4:30-34, 9:8-22. We agree with EMC that Woodhill's Distributed Storage Manager program 24 determines whether a binary object or file is present at remote backup file server 12 by examining Binary Object Identifiers 74 for the plurality of binary objects or files stored in File Database 25. Ex. 1041, 15. Dr. Clark further confirms EMC's position in that regard. Ex. 1009 ¶ 91; *see* Ex. 1008 ¶ 14. We credit Dr. Clark's testimony because it is consistent with Woodhill's description of File Database 25, as well as its description of the backup procedure.

In summary, we agree with EMC that Woodhill's backup procedure, which includes examining Binary Object Identifiers 74 stored in File database 25 to determine if a particular binary object is present at remote backup file server 12, describes the function of "examining the identifiers of the plurality of data items at said particular location in the system," as recited in dependent claim 3. For the foregoing reasons, we conclude that that EMC has demonstrated by a preponderance of the evidence that dependent claim 3 is anticipated by Woodhill.

6. Claim 4

a. "accessing a particular data item using the identifier for the data item"

Dependent claim 4 recites, in relevant part, "access means for *accessing a particular data item using the identifier of the data item*." Ex. 1001, 39:40-41 (emphasis added).

In its Petition, EMC contends that Woodhill discloses the "access means," as claimed. Pet. 58 (citing Ex. 1005, 7:60-8:65, 18:11-23). In particular, EMC argues that Woodhill's Distributed Storage Manager program 24 executes a self-audit procedure on a computer that accesses binary objects using their Binary Object Identifiers 74 during the backup/restore routine. Ex. 1041, 16. EMC also argues that Woodhill's Distributed Storage Manager program 24 performs self-audits by initiating a restore of a randomly selected binary object using its Binary Object Identification record 58, which includes, amongst other things, Binary Object Identifier 74. *Id*. EMC offers the testimony of Dr. Clark to support its position. Ex. 1009 ¶¶ 94, 95.

In its Patent Owner Response, PersonalWeb contends that Woodhill does not use Binary Object Identifier 74, which is part of Binary Object Identification record 58, to access a particular binary object. PO Resp. 24-29. In particular, PersonalWeb argues that, during Woodhill's self-audit procedure, Binary Object Identifier 74 is used merely for comparison purposes after the particular binary object already has been accessed to determine if the audit restore worked properly. *Id*. at 25 (citing Ex. 1005, 18:28-38; Ex. 2013 ¶ 54). PesonalWeb further argues that Woodhill's File Location 38 and File Name 40 in File Identification Record 34 are used to access a file containing a particular binary object, whereas Binary Object

Stream Type 62 and Binary Object Offset 72 in Binary Object Identification record 58 are used to locate the binary object in that file. *Id*. at 25-26 (Ex. 1005, 9:18-20, fig. 2; Ex. 2013 ¶ 55).

In its Reply, EMC contends that, contrary to PersonalWeb's arguments, Woodhill uses Binary Object Identifier 74 to name and restore binary objects. Reply 8 (citing Ex. 1005, 18:13-19, 22:3-4; Ex. 1081 ¶¶ 20-25). EMC directs us to Dr. Clark's testimony that there was no need to explain which subfields of Binary Object Identification Record 58 are used to access a binary object because, for such a basic and well known operation, a person of ordinary skill in the art would have understood that Binary Object Identifier 74 is used to look up a binary object. *Id*. at 9 (citing Ex. 1081 ¶ 20). EMC asserts this is why Woodhill's Binary Object Identifier 74 is referred as an "identifier," and why independent claim 1 of Woodhill refers to it as the "name" of a binary object. *Id*. EMC further contends that, instead of using Woodhill's File Location 38 and File Name 40 in File Identification Record 34 to access a binary object from remote backup file server 12, Woodhill uses Binary Object Identifier 74—the key component of Binary Object Identification record 58—to access the binary object from remote backup file server 12. *Id*. at 9-10.

As we explained in the Decision to Institute (Dec. 29), Woodhill discloses that Distributed Storage Manager program 24 performs auditing and reporting functions on a periodic basis in order to ensure that the binary objects, which already have been backed up, may be restored. Ex. 1005, 18:11-13. According to Woodhill, Distributed Storage Manager program 24 initiates a restore of a randomly selected binary object identified by Binary Object Identification Record 58 stored in File Database 25. Ex. 1005, 18:16-

19. Binary Object Identification Record 58 includes, amongst other things, a Binary Object Identifier 74, which is a unique identifier for each binary object. Ex. 1005, 4:35-47, 7:64-8:1.

We are not persuaded by PersonalWeb's argument that Woodhill does not use Binary Object Identifier 74, which is part of Binary Object Identification record 58, to access a particular binary object. Upon reviewing Woodhill's description of Binary Object Identification record 58, the only part of the record that uniquely identifies the binary object associated therewith is Binary Object Identifier 74. Ex. 1005, 4:45-47, 8:33-65. Therefore, during Woodhill's self-auditing procedure, we determine that Distributed Storage Manager program 24 uses Binary Object Identifier 74 to access a randomly selected binary object by retrieving its corresponding Binary Object Identification record 58 in File Database 25. *See* Ex. 1005, 18:16-19. Dr. Clark confirms such an operation was routine because it was old and well-known to access objects using their identifiers. *See* Ex. 1081 ¶ 20. We credit Dr. Clark's testimony because it is consistent with a general understanding of how one with ordinary skill in the art would use an identifier for basic file management functions, e.g., using an identifier to access a record stored in a database.

Next, we are not persuaded by PersonalWeb's argument that, during the self-auditing procedure, Binary Object Identifier 74 merely is used for comparison purposes after the particular binary object already has been accessed to determine if the audit restore worked properly. As we explained above, the only part of Binary Object Identification record 58 that uniquely identifies the binary object associated therewith is Binary Object Identifier 74. Ex. 1005, 4:45-47, 8:33-65. Consequently, during Woodhill's self-

auditing procedure, Binary Object Identifier 74 serves the following two purposes: (1) Distributed Storage Manager program 24 uses Binary Object Identifier 74 to access a randomly selected binary object by retrieving its corresponding Binary Object Identification record 58 in File Database 25 (*see* Ex. 1005, 18:16-19); and (2) Binary Object Identifier 74, which is stored as part of the randomly selected Binary Object Identification record 58, is compared with Binary Object Identifier 74, previously calculated by Distributed Storage Manager program 24, in order to confirm whether the audit restore was successful (Ex. 1005, 18:28-38).

We also are not persuaded by PersonalWeb's argument that Woodhill's File Location 38 and File Name 40 in File Identification Record 34 are used to access a file containing a particular binary object, whereas Binary Object Stream Type 62 and Binary Object Offset 72 in Binary Object Identification record 58 are used to locate the binary object in that file. Although we recognize that a file containing a particular binary object may be accessed using File Location 38 and File Name 40 (Ex. 1005, 3:56-63), we nonetheless are persuaded that EMC has presented sufficient evidence to support a finding that a particular binary object or file also may be accessed using its Binary Object Identifier 74 (*see, e.g.*, Ex. 1005, 4:45-47, 8:33-65, 18:10-38).

Consequently, we agree with EMC that Woodhill's self-auditing procedure, which includes using Binary Object Identifier 74 to access a randomly selected binary object by retrieving its corresponding Binary Object Identification record 58 in File Database 25, describes the function of "accessing a particular data item using the identifier of the data item," as recited in dependent claim 4.

b. "access means"

Dependent claim 4 recites, in relevant part, "*access means* for accessing a particular data item using the identifier of the data item." Ex. 1001, 39:40-41 (emphasis added).

In its Patent Owner Response, PersonalWeb contends that Woodhill fails to disclose the corresponding structure identified for the claimed "access means." PO Resp. 29-30. In particular, PersonalWeb argues that, according to the specification of the '791 patent, the corresponding structure for this means-plus-function limitation includes looking to True File registry 126 for the record of a corresponding True Name. *Id*. at 29. PersonalWeb alleges that Woodhill fails to disclose such structure or anything equivalent thereto. *Id*.

In its Reply, EMC contends that PersonalWeb's argument is predicated on an improper claim construction for the claimed "access means." Reply 10. In particular, EMC argues that PersonalWeb attempts to add a plurality of True Names for a plurality of files in True File registry 126 to the corresponding structure for the claimed "access means," as well as limit this means-plus-function limitation to both the format of the records and number records to be checked. *Id*.

As explained in our Decision to Institute, we identified the corresponding structure for performing the claimed function of "accessing a particular data item using the identifier of the data item" to be a data processor programmed according to steps S292 and S294 illustrated in Figure 17(a). Dec. 25-26. With respect to steps S292 and S294 illustrated in Figure 17(a), the specification of the '791 patent discloses "look[ing] to the True File registry 126 for a True File entry record 140 for a corresponding

True Name (Step S292). . . . If there is already a True File ID for the entry (Step S294), this mechanism's task is complete." Ex. 1001, 17:10-23.

Although we agree with PersonalWeb that the claimed function of "accessing a particular data item using the identifier of the data item" encompasses looking to True File registry 126 for the record of a corresponding True Name, we nonetheless are persuaded that EMC has presented sufficient evidence to support a finding that Woodhill's Distributed Storage Manager program 24 may look to File Database 25 for Binary Object Identification record 58 of a corresponding Binary Object Identifier 74. As we explained previously, Woodhill's Distributed Storage Manager program 24 may access a particular binary object or file by using its Binary Object Identifier 74 to retrieve its corresponding Binary Object Identification record 58 in File Database 25. *See, e.g.*, Ex. 1005, 4:45-47, 8:33-65, 18:10-38.

Consequently, we agree with EMC that Woodhill's self-auditing procedure, which includes accessing a randomly selected binary object by using its Binary Object Identifier 74 to retrieve its corresponding Binary Object Identification record 58 in File Database 25, describes the corresponding structure for performing the function of "accessing a particular data item using the identifier of the data item," as recited in dependent claim 4.

c. "data associating means"

Dependent claim 4 recites, in relevant part, "*data associating means* for making and maintaining, for a data item in the system, an association between the data item and the identifier of the data item. Ex. 1001, 39:37-39 (emphasis added).

In its Petition, EMC contends that Woodhill discloses the "data associating means," as claimed. Pet. 58 (citing Ex. 1005, 7:60-8:65, 18:11-23). In particular, EMC argues that Woodhill's Distributed Storage Manager program 24, which executes on a computer, accesses and checks Binary Object Identification records 58 in File Database 25 to break up a plurality of files into one or more data streams, each of which is divided into one or more binary objects. Ex. 1041, 16 (citing Ex. 1005, fig. 5A). EMC also argues that Woodhill's Distributed Storage Manager program 24 executes on a computer to make and maintain Binary Object Identification record 58, which associates each binary object with its Binary Object Identifier 74. *Id*. EMC offers the testimony of Dr. Clark to support its position. Ex. 1009 ¶¶ 92, 93.

In its Patent Owner Response, PersonalWeb contends that the claimed function associated with the "data associating means" includes deleting a file in response to comparing True Names. PO Resp. 30. PersonalWeb argues that Woodhill fails to disclose this deletion function or anything equivalent thereto. *Id*. PersonalWeb further argues that Woodhill teaches away from this deletion function because Woodhill discloses that the purpose of the backup procedure is to ensure that backup copies of the binary objects are saved—not deleted or lost. *Id*. at 30-31.

In its Reply, EMC contends that the critical aspect of the claimed function associated with the "data association means" is to avoid unwanted duplicates. Reply 11. EMC argues that Woodhill performs this function because it detects and avoids unwanted duplicates in File Database 25. *Id*. (citing Ex. 1005, 9:23-27; Ex. 1081 ¶ 26). In particular, EMC argues that, during Woodhill's backup procedure, Woodhill prevents unwanted

duplicates before they happen by determining which parts of a file have changed, and only backing up that changed data. *Id*. at 12 (Ex. 1005, 9:24-25; Ex. 1081 ¶ 26).

As explained in our Decision to Institute, we identified the corresponding structure for performing the claimed function of "making and maintaining, for a data item in the system, an association between the data item and the identifier of the data item," to be a data processor programmed according to steps S230, S232, and S237-239 illustrated in Figure 11. Dec. 23-25. With respect to steps S230, S232, and S237-239 illustrated in Figure 11, the specification of the '791 patent discloses the following:

> First, determine the True Name of the data item corresponding to the given scratch File ID using the Calculate True Name primitive mechanism (Step S230). Next, look for an entry for the True Name in the True File registry 126 (Step S232) and determine whether a True Name entry, record 140, exists in the True File registry 126. If the entry record includes a corresponding True File ID or compressed File ID (Step S237), delete the file with the scratch File ID (Step S238). Otherwise store the give True File ID in the entry record (step S239).

Ex. 1001, 14:51-60.

Contrary to PersonalWeb's arguments, we construed the claimed function associated with the "data associating means" to encompass detecting and avoiding duplicate True File IDs in True File registry 126. *See, e.g.*, Ex. 1001, 14:51-60, fig. 11, S230, S232, S237-239. The claimed function is not limited to deleting a file in response to comparing a True File ID. *See, e.g.*, Ex. 1001, fig. 11, step S238. We agree with EMC that, during Woodhill's backup procedure, Distribute Storage Manager program 24 detects and avoids duplicate Binary Object Identifiers 74 in File Database 25

by determining which parts of a binary object or file have changed, and only backing up the changed data. Ex. 1005, 9:23-27. Dr. Clark also testifies that the claimed function associated with the "data associating means," and the operations performed during Woodhill's backup procedure, perform the identical function because they each assimilate data items without creating duplicates. Ex. 1081 ¶ 26. We credit Dr. Clark's testimony because it is consistent with our claim construction for the "data associating means," as well as Woodhill's description of the backup procedure.

To the extent PersonalWeb contends that Woodhill teaches away from the claimed function for the "data associating means" because the purpose of its backup procedure is to ensure that backup copies of the binary objects are saved—not deleted or lost—we disagree. PO Resp. 30. PersonalWeb's argument is not persuasive because EMC's proposed ground of unpatentability is based on anticipation by Woodhill. It is well settled that "[t]eaching away is irrelevant to anticipation." *Seachange Int'l, Inc., v. C-Cor, Inc.*, 413 F.3d 1361, 1380 (Fed. Cir. 2005).

In summary, we agree with EMC that Woodhill's backup procedure, which includes detecting and avoiding duplicate Binary Object Identifiers 74 in File Database 25 by determining which parts of a binary object or file have changed, and only backing up the changed data, describes the claimed function associated with the "data associating means," as recited in dependent claim 4. For the foregoing reasons, we conclude that that EMC has demonstrated by a preponderance of the evidence that dependent claim 4 is anticipated by Woodhill.

7. Claim 29

Dependent claim 29 recites "a data item is at least one of a file, a database record, a message, a data segment, a data block, a directory, and an instance [of] an object class." Ex. 1001, 42:54-57. The contentions and supporting evidence presented by EMC that explain how Woodhill describes the claimed subject matter recited in dependent claim 29 have merit and otherwise are unrebutted by PersonalWeb. Pet. 58 (citing Ex. 1005, 7:51-55); *see* Ex. 1009 ¶¶ 83-96. Therefore, we conclude that EMC has demonstrated by a preponderance of the evidence that dependent claim 29 is anticipated by Woodhill.

8. Claim 30

a. "accessing a data item in the system using the identifier of the data item"

Dependent claim 30 recites, in relevant part, "accessing a data item in the system using the identifier of the data item." Ex. 1001, 42:66-67. PersonalWeb relies upon essentially the same argument presented against dependent claim 4 to rebut the explanations provided by EMC as to how Woodhill describes the above-identified method step recited in independent claim 30. *Compare* PO Resp. 23-29 *with* PO Resp. 31-36. For the same reasons discussed above with respect to dependent claim 4, PersonalWeb's arguments are not persuasive. Therefore, we conclude that EMC has demonstrated by a preponderance of the evidence that independent claim 30 is anticipated by Woodhill.

9. Claims 31 and 32

Dependent claim 31 recites:

> making and maintaining, for a plurality of data items present in the system, an association between each of the data items and the identifier of each of the data items, wherein said accessing of a data item accesses a data item via the association.

Ex. 1001, 43:2-6. Dependent claim 32 recites "assimilating a new data item into the system, by determining the identifier of the new data item and associating the new data item with its identifier." Ex. 1001, 43:8-10. The contentions and supporting evidence presented by EMC that explain how Woodhill describes the claimed subject matter recited in dependent claims 31 and 32 have merit and otherwise are unrebutted by PersonalWeb. Pet. 58 (citing Ex. 1005, 7:60-8:65, 18:11-23); *see* Ex. 1009 ¶¶ 83-96. Therefore, we conclude that EMC has demonstrated by a preponderance of the evidence that dependent claims 31 and 32 are anticipated by Woodhill.

10. Claim 41

a. EMC does not switch between different unrelated embodiments when explaining how Woodhill describes the "accessing" method step

Dependent claim 41 recites, in relevant part, "*[t]he method of claim 30*, wherein *said accessing further comprises*: for a given data identifier and for a given current location and a remote location in the system." Ex. 1001, 45:8-10.

In its Petition, EMC contends that Woodhill discloses "for a given data identifier . . . to the current location," as recited in dependent claim 41. Pet. 58 (Ex. 1005, 9:5-23). EMC argues that, during Woodhill's backup procedure, the data processing system only backs up changed binary objects since the previous backup. Ex. 1041, 21. EMC further argues that

Woodhill's data processing system backs up binary objects from local computers 20 on remote backup file server 12. *Id*.

In its Patent Owner Response, PersonalWeb contends that EMC relies upon the self-auditing procedure disclosed in Woodhill to describe the "accessing" method step recited in independent claim 30, yet EMC relies upon the backup procedure disclosed in Woodhill to describe the additional features of the same "accessing" method step recited in dependent claim 41. PO Resp. 36-37 (Ex. 1005, 9:5-23, 18:10-38; Ex. 1009 ¶¶ 94-96). PersonalWeb argues that EMC cannot switch between different unrelated embodiments in Woodhill when explaining how Woodhill describes the "accessing" method step, as recited in independent claim 30, and further recited in dependent claim 41. *Id*. at 37-38.

In its Reply, EMC contends that it relied on only one embodiment in Woodhill to describe the "accessing" method step, as recited in independent claim 30, and further recited in dependent claim 41. Reply 12. In particular, EMC argues that Woodhill's Distributed Storage Manager program 24 is a single structure divided into several distinct functions that are illustrated in Figures 5A through 5L. *Id*. (citing Ex. 1004, 4:62-67). EMC further argues that, when Woodhill describes Distributed Storage Manager program 24, it indicates that "each distinct function operates in cooperation with the other functions to form a unitary computer program." *Id*. (quoting Ex. 1005, 4:67-5:2). We agree with EMC that it only relies upon one embodiment to describe the "accessing" method step, as recited in independent claim 30, and further recited in dependent claim 41.

When determining whether EMC relies on a single embodiment in Woodhill to describe the claimed "accessing" method step, our inquiry is

"not constrained to proceed example-by-example when reviewing an allegedly anticipating prior art reference. Rather, [we] must, while looking at the reference as a whole, conclude whether or not that reference discloses all elements of the claimed invention arranged as in the claim." *Net MoneyIN, Inc. v. VeriSign, Inc.*, 545 F.3d 1359, 1369 n.5 (Fed. Cir. 2008).

The relevant disclosure in Woodhill is reproduced below in its entirety:

> For explanation purposes, the Distributed Storage Manager program 24 is divided into several functions which will be discussed in turn. Those of ordinary skill in the art will recognize, however, that each of the distinct functions operates in cooperation with the other functions to form a unitary program. Those of ordinary skill in the art will also recognize that the following discussion illustrates the operation of the Distributed Storage Manager program 24 on a single local computer 20, although it should be understood that the Distributed Storage Manager program 24 operates in the same fashion on each local computer 20 on the networked computer system 10.

Ex. 1005, 4:64-5:9.

Woodhill then proceeds to provide separate and distinct explanations as to how Distributed Storage Manager program 24 handles the operations of the backup procedure and the self-auditing procedure. Ex. 1005, 9:5-23, 18:10-38. Therefore, contrary to PersonalWeb's argument, Woodhill's backup procedure and self-auditing procedure are not mutually exclusive embodiments, but rather are distinct functions that operate with other functions to form one unitary computer program—namely Woodhill's Distributed Storage Manager program 24. Consequently, we are not persuaded that EMC switches between different unrelated embodiments

when explaining how Woodhill describes the "accessing" method step, as recited in independent claim 30, and further recited in dependent claim 41.

b. Woodhill's back-up procedure discloses the claimed subject matter recited in claim 41

Dependent claim 41 recites, in relevant part:

> determining whether the data item corresponding to the given data identifier is present at the current location, and based on said determining, if said data item is not present at the current location, fetching the data item from a remote location in the system to the current location.

Ex. 1001, 45:11-16.

In its Patent Owner Response, PersonalWeb contends that Woodhill's backup procedure fails to disclose the above-identified features recited in dependent claim 41. PO Resp. 41-43. According to PersonalWeb, the ordinary and customary meaning of the claimed term "fetch" is "to go after and return." *Id*. at 41 (quoting THE AMERICAN HERITAGE DICTIONARY 486 (1975) (Ex. 2004)). Based on that dictionary definition, PersonalWeb argues that, during Woodhill's backup procedure, a new binary object is simply transmitted to remote backup filer server 12—not fetched. *Id*. In its Reply, EMC maintains that it properly relied on the functions performed by Woodhill's Distributed Storage Manager program 24, which it maintains is a single embodiment that incorporates both the operations of the backup procedure and the self-auditing procedure. Reply 13 (citing Ex. 1041, 18, 20-21).

According to Woodhill, both its system and method for managing storage space on network computer system 10 include selectively copying a binary object stored in one storage area to another storage area. Ex. 1005,

Abstract, 2:4-6, 25-27. In our view, this general description of Woodhill's invention applies to its backup procedure. For instance, during Woodhill's backup procedure, Distributed Storage Manager program 24 backs up each binary object by storing a compressed copy of the binary object in the following two locations: (1) on disk drive 19 associated with local computer 20 somewhere other than local computer 20 where the binary object was stored originally; and (2) on remote backup file server 12. Ex. 1005, 9:31-38. Therefore, if a binary object ever was lost or destroyed at an entire site, e.g., disk drive 19 on local computer 20 or remote backup file server 12, Woodhill indicates that a copy of the binary object stored in another storage area may be copied to that site. *See* Ex. 1005, 9:39-45.

Even if we accept PersonalWeb's definition of "fetch" as "to go after and return with" (Ex. 2004), Woodhill's backup procedure still discloses fetching a binary object from a remote location if it is no longer present, e.g., lost or destroyed, at a current location, as required by dependent claim 41. For instance, if a failure occurs at disk drive 19 on local computer 20, Distributed Storage Manager program 24 may determine whether a binary object still is present at that location, i.e., the claimed "current location," by examining the binary objects and their corresponding Binary Object Identifiers 74 stored on disk drive 19. If the binary object and its corresponding Binary Object Identifier 74 have been lost, destroyed, or are no longer present at disk drive 19, Distributed Storage Manager program 24 could fetch a copy of the binary object using its Binary Object Identifier 74 from remote backup file server 12, i.e., the claimed "remote location," and return it to disk drive 19.

Alternatively, if a failure occurs at remote backup file server 12, Distributed Storage Manager program 24 may determine whether a binary object still is present at that location, i.e., the claimed "current location," by examining the binary objects and their corresponding Binary Object Identifiers 74 stored in File Database 25. If the binary object and its corresponding Binary Object Identifier 74 have been lost, destroyed, or are no longer present at remote backup file server 12, Distributed Storage Manager program 24 could fetch a copy of the binary object using its Binary Object Identifier 74 from disk drive 19 on local computer 20, i.e., the claimed "remote location," and return it to remote backup file server 12.

PersonalWeb also reiterates its argument that Woodhill's backup procedure does not "access a data item in the system using the identifier of the data item," as required by dependent claim 41 based on its dependency from independent claim 30. PO Resp. 42-43. For the same reasons discussed above with respect to dependent claim 4 and independent claim 30, PersonalWeb's argument is not persuasive.

c. Woodhill determines whether a particular data item is "not present" at a given location

In its Patent Owner Response, PersonalWeb presents a number of arguments that are predicated on the notion that Woodhill only is capable of analyzing information for a single binary object or file stored at a given location, and is incapable of analyzing the other files stored at that location. PO Resp. 43-49. PersonalWeb also alleges that both parties agree that Woodhill is incapable of determining whether a particular data item is "not present" at a given location. *Id*. at 42-43, 48-49 (citing Ex. 2008, 143, 145, 150-151). Based on those arguments, PersonalWeb asserts that Woodhill

does not disclose "based on said determining, if said data item is not present at the current location, fetching the data item from a remote location in the system to the current location," as recited in dependent claim 41. *Id*. at 43, 49.

In its Reply, EMC contends that PersonalWeb mischaracterizes Dr. Clark's testimony by asserting that he agreed that it is impossible for Woodhill to determine whether a particular data item is not present at a given location. Reply 14. Instead, EMC argues that Dr. Clark only agreed that the hypothetical proposed by PersonalWeb during cross examination made this impossible—not that it was, in fact, impossible for Woodhill to determine whether a particular data item is not present at a given location. *Id*. EMC further contends that Woodhill is capable of determining whether the current version of a binary object or file is not present at remote backup server 12, i.e., the claimed "current location," and transmitting the file to that server. *Id*. (citing Ex. 1081 ¶ 11).

As we explained previously, Woodhill discloses that File Database 25 stores a plurality of Binary Object Identifiers 74 associated with different binary objects or files that have been backed up in the system. *See, e.g.*, Ex. 1005, 3:49-52, 4:30-34, 9:8-22. Dr. Clark testifies that, during Woodhill's backup procedure, Distributed Storage Manager program 24 determines whether a binary object or file is present at remote backup file server 12 by confirming that its Binary Object Identifier 74 already exists among the plurality of Binary Object Identifiers 74 stored in File Database 25. Ex. 1009 ¶ 89; *see* Ex. 1081 ¶ 14. Dr. Clark also testifies that Distributed Storage Manager program 24 only transmits the binary object or file to remote backup file server 12 if it is not already present at that location.

Ex. 1081 ¶ 11. We credit Dr. Clark's testimony because it is consistent with Woodhill's general disclosure of copying binary objects stored in one storage area to another storage area (Ex. 1005, 2:4-6, 25-27), as well as how Woodhill's Distributed Storage Manager program 24 transmits new or changed binary objects or files to remote backup file server 12 (Ex. 1005, 9:36-38).

We are not persuaded by PersonalWeb's allegation that the parties agree that it is "impossible" for Woodhill to determine whether a particular data item is "not present" at a given location. Dr. Clark stated that he did make such an admission and does not agree with PersonalWeb's assertion. Ex. 1081 ¶ 11. Upon reviewing the cited pages in the transcript of PersonalWeb's deposition of Dr. Clark, we note that the questions posed by PersonalWeb's counsel to Dr. Clark are couched in hypotheticals, and not directed to Woodhill's disclosure. *See, e.g.*, Ex. 2008, 143 ("Let's assume we have file A and File B. They're different files. Each of them has a plurality of binary objects. . . . Assume that the exact same binary object is actually present in both file A and file B."); Ex. 2008, 153 ("Assume you are given a sequence of bits, and you have a thousand files stored in a server, and you only have the capability of figuring out if that sequence of bits is in only one of those files, and you do not have the capability of figuring out if that sequence of bits is in the other 999 of those files."); Ex. 1081 ¶ 12. It is not clear to us how each of those constrained hypotheticals relates to the backup procedure disclosed in Woodhill, much less how a conclusion can be drawn from Dr. Clark's response to each hypothetical that he readily admitted it is "impossible" for Woodhill to determine whether a particular data item is "not present" a given location.

In summary, we agree with EMC that Woodhill's backup procedure, which includes determining whether a binary object or file corresponding to Binary Object Identifier 74 is present at remote backup file server 12 and, if not, transmitting it to that location, describes determining whether a particular data item is "not present" at a given location, as required by dependent claim 41. For the foregoing reasons, we conclude that EMC has demonstrated by a preponderance of the evidence that dependent claim 41 is anticipated by Woodhill.

11. Claim 33

a. Woodhill determines whether a particular data item is "not present" at a destination location

Independent claim 33 recites, in relevant part:

> determining, using the data identifier, whether the data item is present at the destination location; and based on the determining whether the data item is present, providing the destination location with the data item only if the data item is not present at the destination [location].

Ex. 1001, 43:19-23. PersonalWeb relies upon essentially the same arguments presented against dependent claim 41 to rebut the explanations provided by EMC as to how Woodhill describes the above-identified features recited in dependent claim 33. *Compare* PO Resp. 43-51 *with* PO Resp. 51-57. For the same reasons discussed above with respect to dependent claim 41, PersonalWeb's arguments are not persuasive.

b. Woodhill describes providing the destination location with a data item "only if" it is determined that the data item is not present at that destination location

PersonalWeb contends that, because Woodhill cannot determine whether a particular data item is "not present" at a destination location, Woodhill cannot disclose providing the destination location with a data item "only if" it is determined that the data item is not present at that destination location, as required by independent claim 33. PO Resp. 57 (citing Ex. 2013 ¶ 95). As we have explained previously, during Woodhill's backup procedure, Distributed Storage Manager program 24 determines whether a binary object or file corresponding to Binary Object Identifier 74 already exists on remote backup server 12 and, if not, transmits the binary object or file to that location. *See* Ex. 1005, 3:49-52, 4:30-34, 9:1-38; Ex 1009 ¶ 89; Ex. 1081 ¶¶ 11, 12, 14. In that scenario, Woodhill's remote backup file server 12 constitutes the claimed "destination location." For the foregoing reasons, we conclude that EMC has demonstrated by a preponderance of the evidence that independent claim 33 is anticipated by Woodhill.

D. Obviousness over Woodhill—Claims 1-4 and 29

EMC contends that claims 1-4 and 29 are unpatentable under § 103(a) over Woodhill. Pet. 59. In support of that alleged ground of unpatentability, EMC provides explanations as to how Woodhill teaches or suggests each claim limitation. *Id*. (citing Ex. 1041). EMC also submits declarations of Dr. Clark (Ex. 1009 ¶¶ 97-98; Ex. 1081) to support its positions. Upon reviewing EMC's Petition and supporting evidence, as well as PersonalWeb's Patent Owner Response and supporting evidence, we determine that EMC has demonstrated by a preponderance of the evidence

that claims 1-4 and 29 are obvious over Woodhill.

We begin our analysis with the principles of law that generally apply to a ground of unpatentability based on obviousness, and then we turn to the arguments presented by both EMC and PersonalWeb that are directed to whether Woodhill, as a whole, would have taught or suggested the "identity means" recited in independent claim 1 to one with ordinary skill in the art.

1. Principles of Law

A patent claim is unpatentable under § 103(a) if the differences between the claimed subject matter and the prior art are such that the subject matter, as a whole, would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. *KSR Int'l Co. v. Teleflex Inc*., 550 U.S. 398, 406 (2007).
The question of obviousness is resolved on the basis of underlying factual determinations, which include the following: (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of skill in the art; and (4) where in evidence, so-called secondary considerations. *Graham v. John Deere Co.*, 383 U.S. 1, 17-18 (1966). We also recognize that prior art references must be "considered together with the knowledge of one of ordinary skill in the pertinent art." *In re Paulsen*, 30 F.3d 1475, 1480 (Fed. Cir. 1994). We analyze the ground of unpatentability based on obviousness over Woodhill with the above-identified principles in mind.

2. PersonalWeb's Contentions

a. There are no deficiencies in Woodhill to cure

At the outset, PersonalWeb contends that EMC's contentions regarding obviousness do not cure the deficiencies in Woodhill that are

discussed above with respect to claims 1-4 and 29. PO Resp. 58-59. As we explained in our discussion of the ground of unpatentability based on anticipation by Woodhill, there are no such deficiencies in Woodhill to cure.

b. Woodhill, as a whole, would have taught or suggested the claimed "identity means" to one with ordinary skill in the art

Independent claim 1 recites, in relevant part:

> identity means for determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier.

Ex. 1001, 39:16-21.

In its Petition, EMC contends that, to the extent that the claimed "identity means" requires an MD5 hash function, a person of ordinary skill in the art would have found it obvious to calculate Woodhill's Binary Object Identifier 74 for a particular binary object or file using an MD5 hash function. Pet. 59 (citing Ex. 1005, 8:52-58). According to Dr. Clark, this modification to Woodhill would constitute a simple substitution of one known element for another to obtain predictable results. *Id*. (citing Ex. 1009 ¶¶ 97, 98).

In its Patent Owner Response, PersonalWeb contends that it would not have been obvious to calculate Woodhill's Binary Object Identifier 74 using an MD5 has function because there were thousands, if not millions, of possible hash function known at the time of Woodhill's invention, and there would have been no logical reason to select an MD5 hash function for use in Woodhill. PO Resp. 59 (citing Ex. 2013 ¶ 98). PersonalWeb also argues

that an MD5 hash function produces 16-byte hash values, whereas Woodhill desires a 4-byte hash value. *Id*. (citing Ex. 1005, 8:1-3).

In its Reply, EMC contends that PersonalWeb's argument only is relevant if we change our claim construction for "identity means" to include an MD5 hash function. Reply 15. In any event, EMC argues that Dr. Clark confirms that MD5 hash functions were old and well-known at the time of the invention of the '791 patent, and that use of an MD5 hash function in Woodhill's system would be a simple and obvious substitution. *Id*. (citing Ex. 1009 ¶¶ 97, 98; Ex. 1081 ¶¶ 27-29).

To the extent PersonalWeb now argues that the claimed "identity means" requires an MD5 hash function, we disagree. Similar to our explanation in the Decision to Institute, PersonalWeb's argument in that regard is not commensurate in scope with our claim construction of "identity means." *See* Dec. 27-28. The corresponding structure for performing the function of "determining, for any of a plurality of data items present in the system, a substantially unique identifier, the identifier being determined using and depending on all of the data in the data item and only the data in the data item, whereby two identical data items in the system will have the same identifier" is a data processor programmed to perform a hash function, e.g., MD5 or SHA. Neither the specification of the '791 patent, nor the claim itself, indicates that the "identity means" requires an MD5 hash function. Instead, an MD5 hash function is merely one of numerous hash functions capable of being programmed on a data processor that would satisfy this means-plus-function limitation.

Nonetheless, even if we assume that the claimed "identity means" requires an MD5 hash function, we agree with EMC that a person of

ordinary skill in the art would have found it obvious to calculate Woodhill's Binary Object Identifier 74 for a particular binary object or file using an MD5 hash function. As discussed above, PersonalWeb asserts that an MD5 hash function produces 16-byte hash values. Woodhill discloses calculating Binary Object Identifier 74 for a binary object in various ways, including using a binary hash algorithm. Ex.1005, 8:1-31. Of importance here is that Woodhill discloses calculating Binary Object Hash field 70 against the content of the binary object taken one word or 16-bytes at a time. Ex. 1005, 8:23-24. Therefore, similar to PersonalWeb's assertion that an MD5 hash function produces 16-byte hash values, the binary hash algorithm disclosed in Woodhill also produces 16-byte hash values. Dr. Clark's testimony further confirms that Woodhill's Binary Object Identifier 74 and MD5 value are the same byte length, i.e., 16-bytes. Ex. 1081 ¶ 28.

In addition, we agree with PersonalWeb that one with ordinary skill in the art would have substituted an MD5 hash algorithm, which Dr. Clark confirms was old and well-known in the art at the time of the invention of the '791 patent (Ex. 1009 ¶ 97; Ex. 1081 ¶ 27), for Woodhill's binary hash algorithm. In our view, such a substitution is a predictable use of prior art elements according to their established functions—an obvious improvement. *See KSR*, 550 U.S. at 479.

In summary, PersonalWeb's assertion that the claimed "identity means" requires an MD5 hash function is not commensurate in scope with our claim construction for this mean-plus-function limitation. Nonetheless, even if we assume that the claimed "identity means" requires an MD5 hash function, EMC has presented sufficient evidence that Woodhill, as a whole,

would have taught or suggested this means-plus-function limitation to one with ordinary skill in the art.

c. Secondary Considerations of Non-Obviousness—Licenses

In its Patent Owner Response, PersonalWeb contends that, because third parties have licensed the '791 patent, evidence of non-obviousness exists that outweighs the evidence of obviousness based on Woodhill presented by EMC in this proceeding. PO Resp. 59-60. In support of its argument, PersonalWeb directs us to three licensing agreements (Exs. 2010-12), as well as the declaration of Kevin Bermeister (Ex. 2009 ¶¶ 3-9), and then argues that each license granted to a third party was not for the purpose of settling a patent infringement suit. *Id*. at 60.

In its Reply, EMC contends that PersonalWeb has failed to establish a sufficient nexus between claims 1-4 and 29 and the above-identified licensing agreements. Reply 15. EMC argues that each of the three licensing agreements granted rights to more than just claims 1-4 and 29, and involved related parties with interlocking ownership and business interests. *Id*. We agree with EMC that PersonalWeb has failed to establish the requisite nexus between the licensing agreements and the claimed subject matter recited in claims 1-4 and 29.

A party relying on licensing activities as evidence of non-obviousness must demonstrate a nexus between those activities and the subject matter of the claims at issue. *GPAC*, 57 F.3d at 1580. Further, without a showing of nexus, "the mere existence of . . . licenses is insufficient to overcome the conclusion of obviousness" when there is a strong ground of unpatentability based on obviousness. *SIBIA Neurosciences, Inc. v. Cadus Pharm. Corp.*,

225 F.3d 1349, 1358 (Fed. Cir. 2000); *see Iron Grip Barbell Co. v. USA Sports, Inc.*, 392 F.3d 1317, 1324 (Fed. Cir. 2004).

The evidence of non-obviousness presented by PersonalWeb falls short of demonstrating the required nexus in two respects. First, neither PersonalWeb nor the declaration of Mr. Bermeister (Ex. 2009) establishes that the licensing agreements (Exs. 2010-12) are directed to the claimed subject matter recited in claims 1-4 and 29. For instance, PersonalWeb does not present credible or sufficient evidence that the three licensing agreements arose out of recognition and acceptance of the claimed subject matter recited in claims 1-4 and 29. In the absence of an established nexus with the claimed invention, secondary consideration factors are entitled little weight, and generally have no bearing on the legal issue of obviousness. *See In re Vamco Machine & Tool, Inc.*, 752 F.2d 1564, 1577 (Fed. Cir. 1985). Second, even if we assume that the above-identified licenses establish some degree of industry respect for the claimed subject matter recited in claims 1-4 and 29, that success is outweighed by the strong evidence of obviousness over Woodhill discussed above.

Based on the record before us, including the evidence of obviousness based on Woodhill and the evidence of secondary considerations regarding licensing activities, we conclude that EMC has demonstrated by a preponderance of the evidence that claims 1-4 and 29 are obvious over Woodhill.

E. PersonalWeb's Motion to Exclude

PersonalWeb seeks to exclude the following evidence: (1) paragraphs 13, 20, 24, 27, and 28 of the rebuttal declaration of Dr. Clark that rely on, and cite to, Peterson, Tanenbaum, Langer, and RFC 1321 because these

paragraphs are irrelevant, prejudicial, confusing, lacking foundation, and beyond the scope of this proceeding; (2) Langer, because it is not authenticated properly under Federal Rule of Evidence ("FRE") 901; (3) Langer, because it includes impermissible hearsay, in violation of FRE 802; (4) the "capable," "can," and "may" statements in Dr. Clark's rebuttal declaration because these statements are irrelevant, prejudicial, confusing, lacking foundation, and beyond the scope of this proceeding; (5) a new contention that allegedly appears in Dr. Clark's rebuttal declaration because it is prejudicial, outside the scope of this proceeding, lacks foundation, lacks underlying facts and data, is in violation of FREs 702 and 705, and represents a new argument on reply; and (6) paragraphs 17-19 and 23 in Dr. Clark's rebuttal declaration because he relies upon subject matter in Woodhill that does not qualify as prior art to the '791 patent. Paper 62 ("PO Mot."). EMC opposes PersonalWeb's motion to exclude. Paper 71 ("Pet. Opp."). In response, PersonalWeb filed a reply to EMC's opposition to its motion to exclude. Paper 76 ("PO Reply"). For the reasons discussed below, PersonalWeb's motion to exclude is denied.

1. The statements in Dr. Clark's rebuttal declaration regarding Peterson, Tanenbaum, Langer, and RFC 1321 are admissible evidence

PersonalWeb contends that paragraphs 13, 20, 24, 27, and 28 of the rebuttal declaration of Dr. Clark (Ex. 1081) should be excluded because these paragraphs rely upon, and cite to, Peterson (Ex. 1075), Tanenbaum (Ex. 1076), Langer (Ex. 1003), and RFC 1321 (Ex. 1012). PO Mot. 1. PersonalWeb argues that this proceeding was only instituted based on Woodhill—not on Peterson, Tanenbaum, Langer, or RFC1321—and, therefore, EMC's reliance on these documents is outside the scope of this

proceeding and impermissible. *Id*. In response, EMC contends that the statements regarding Peterson, Tanenbaum, Langer, and RFC 1321 are relevant to the instituted grounds of unpatentability based on Woodhill, and simply serve to corroborate the state of the art at the time of the '791 patent. Pet. Opp. 1. We agree with EMC.

The '791 patent has an effective filing date of April 11, 1995. Ex. 1001 at [63]. Peterson has a copyright date of 1983 (Ex. 1075, 2), Tanenbaum has a copyright date of 1987 (Ex. 1076, 2), Langer has publication date of August, 7, 1991 (Ex. 1003, 1), and RFC 1321 is dated April 1992 (Ex. 1012, 1). Each of these references has a publication date prior to April 11, 1995. We recognize that these prior art documents were relied on by EMC's rebuttal declarant, Dr. Clark, and are of the type that experts in the pertinent field reasonably would rely on to formulate their opinions. In other words, EMC may rely on these prior art documents to demonstrate what one with ordinary skill in the art would have known about technical features and developments in the pertinent art at the time of the '791 patent.

For the foregoing reasons, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude paragraphs 13, 20, 24, 27, and 28 of the rebuttal declaration of Dr. Clark.

2. EMC provides sufficient evidence to support a finding that Langer has been authenticated properly

PersonalWeb contends that EMC fails to provide evidence indicating that Langer (Ex. 1003) existed prior to the effective filing date of the '791 patent—April 11, 1995—and, therefore, should be excluded under FRE 901. PO Mot. 2-3. In particular, PersonalWeb argues that Langer allegedly was

downloaded from the Internet in 2003 based on the "7/29/2003" date in the lower, right-hand corner. *Id*. at 2. PersonalWeb also argues that authentication of Langer requires personal knowledge of its existence prior to April 11, 1995. *Id*. at 3. In response, EMC contends that it submitted sworn testimony from Mr. Keith Moore that properly authenticates Langer under FREs 901(b)(1) and (4), 901(b)(3), 901(b)(8), and 901(b)(6) and (7). Pet. Opp. 2-3 (citing Ex. 1052 ¶¶ 5-11). In reply, PersonalWeb contends that Langer is not authenticated properly under the FREs identified by EMC. PO Reply. 1-5.

We agree with EMC that Langer has been authenticated properly under FRE 901(b)(1) and (4) because Mr. Moore testified that Langer is a periodical that was posted on Usenet newsgroups on August 7, 1991 (Ex. 1052 ¶¶ 11-15), and it includes distinct header fields unique to Usenet formatting and content (*id*. at ¶¶ 16,17). Although PersonalWeb presents several theories that attack the authenticity of Langer, PersonalWeb fails to explain adequately why the testimony offered by Mr. Moore does not authenticate Langer. PersonalWeb simply presents mere attorney arguments and does not offer testimony from its own expert contrary to the testimony offered by Mr. Moore. Therefore, based on the record before us, EMC has presented sufficient evidence to support a finding that Langer has been authenticated properly under FRE 901(b)(1) and (4).

We also are not persuaded by PersonalWeb's argument that the download date of "7/29/2003" in the lower, right-hand corner calls into question whether Langer existed prior to April 11, 1995. The mere fact that a "downloaded" copy of Langer has a date subsequent to the earliest effective filing date is not sufficient to rebut EMC's supporting evidence that

Langer is what it claims to be—namely a periodical posted on Usenet newsgroups on August 7, 1991. *See, e.g.*, Ex. 1052 ¶¶ 11-17.

To the extent PersonalWeb argues that Mr. Moore cannot authenticate Langer because he does not have personal knowledge of its existence prior to April 11, 1995, or that Mr. Albert Langer is the only person that can authenticate Langer properly, we disagree. Neither a declaration from Mr. Langer, nor evidence of someone actually viewing Langer prior to April 11, 1995, is required to support a finding that Langer is what it claims to be. *See In re Wyer*, 655 F.2d 221, 226 (CCPA 1981) (Notwithstanding that there is no evidence concerning actual viewing or dissemination of any copy of the Australian application, the court held that "the contents of the application were sufficiently accessible to the public and to persons skilled in the pertinent art to qualify as a 'printed publication."); *In re Bayer*, 568 F.2d 1357, 1361 (CCPA 1978) (A reference constitutes a "printed publication" under 35 U.S.C. § 102(b) as long as a presumption is raised that the portion of the public concerned with the art would have known of the invention.).

For the foregoing reasons, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude Langer as unauthenticated evidence.

3. Langer is not inadmissible hearsay

PersonalWeb contends that the dates in Langer, or any other information that purports to establish a publication date for Langer, are inadmissible hearsay under FRE 802 and not subject to any hearsay exception. PO Mot. 4. PersonalWeb also argues that, to the extent that EMC contends that any statements in Langer were made prior to the critical date of the '791 patent, the entirety of Langer is inadmissible hearsay. *Id.* In response, EMC contends that Langer is not hearsay because it is being

offered for what it describes—not for the truth of its disclosure. Pet. Opp. 4. EMC also argues that the August 7, 1991, posting date on Langer's header and uniform resource locator ("URL") both were generated automatically by the hosting computer and, therefore, are admissible as non-hearsay to prove Langer's August 1991 publication date. *Id*. (citing Ex. 1052 ¶ 7). In reply, PersonalWeb maintains that the dates and other information in Langer used to establish its availability as of August 1991 amount to inadmissible hearsay. PO Reply 5.

We recognize that EMC's rebuttal declarant, Mr. Moore, reasonably would rely on the date of August 7, 1991, that appears in both Langer's header and URL to formulate his opinion on whether Langer was available publicly as of that date. Accordingly, the date of August 7, 1991, posted in Langer need not be admissible for the testimony of Mr. Moore to be admissible. Nonetheless, we agree with EMC that the date of August 7, 1991, posted on Langer's header and URL, serve a non-hearsay purpose for which it can be admitted—namely to prove that the document was available publicly as of that date.

Moreover, we are not persuaded by PersonalWeb's arguments that Langer, in its entirety, constitutes hearsay. With the exception of the dates in Langer, PersonalWeb does not identify specifically the textual portions of Langer that allegedly are being offered for the truth of the matter asserted, yet does seek to exclude Langer in its entirety. We will not go through the entirety of Langer and determine which portions PersonalWeb believes to be hearsay—this is something that PersonalWeb should have done in its motion to exclude.

Accordingly, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude the dates posted in Langer, or any statements made therein, as impermissible hearsay.

4. The "capable", "can," and "may" statements in Dr. Clark's rebuttal declaration are admissible

PersonalWeb contends that the "capable," "can," and "may" statements in Dr. Clark's rebuttal declaration (Ex. 1081) should be excluded because these statements are irrelevant, prejudicial, confusing, lacking foundation, and beyond the scope of this proceeding. PO. Mot. 5-6 (citing FREs 401, 402, 403). In particular, PersonalWeb argues that the issue in this proceeding is what Woodhill discloses, or what is necessarily present in Woodhill, not what Woodhill is "capable" of or "may" do according to Dr. Clark. *Id*. at 5. In response, EMC contends that the "capable," "can," and "may" statements in Dr. Clark's rebuttal declaration were offered in response to arguments presented by PersonalWeb in its Patent Owner Response. Pet. Opp. 5-6 (citing PO Resp. 43-51; Ex. 2013 ¶¶ 46, 52). EMC argues that Dr. Clark was explaining what a person of ordinary skill in the art, upon reading Woodhill, would have understood Woodhill to disclose. *Id*. at 6.

We recognize that the focus of this proceeding is on the instituted grounds of unpatentability based on anticipation by, or obviousness over, Woodhill. Any statements that Dr. Clark makes regarding those grounds of unpatentability simply would affect how we weigh the testimony offered by Dr. Clark. When weighing the evidence provided by both parties, we are capable of determining whether Woodhill anticipates or renders obvious the

challenged claims without being confused, misled, or prejudiced by Dr. Clark's testimony.

Accordingly, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude the "capable," "can," and "may" statements in Dr. Clark's rebuttal declaration.

5. Dr. Clark's statements are direct rebuttal to an argument raised by PersonalWeb in its patent owner response

PersonalWeb contends that Dr. Clark's rebuttal declaration includes a new contention not presented previously with the Petition that should be excluded because it is prejudicial, outside the scope of this proceeding, lacks foundation, lacks underlying facts and data, is in violation of FREs 702 and 705, and represents a new argument on reply. PO Mot. 5-6 (citing Ex. 1081 ¶ 19). In response, EMC contends that Dr. Clark's testimony properly responds to an argument presented by PersonalWeb in its Patent Owner Response. Pet. Opp. 6 (citing PO Resp. 21).

Based on our review of the argument presented by PersonalWeb in its Patent Owner Response, as well as the relevant portion of Dr. Clark's rebuttal declaration, we agree with EMC that Dr. Clark's testimony is direct rebuttal to PersonalWeb's argument that Woodhill does not enable comparing Binary Object Identifier 74 with one or more other Binary Object Identifiers 74. *Compare* PO Resp. 21 *with* Ex. 1081 ¶ 19. In other words, Dr. Clark's statement regarding enablement falls within the purview of 37 C.F.R. § 42.23(b), which provides that a petitioner's reply may only respond to arguments raised in the corresponding patent owner response.

Accordingly, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude Dr. Clark's statements regarding enablement.

6. The statements in Dr. Clark's rebuttal declaration that rely on the claim language of Woodhill are admissible

PersonalWeb contends that paragraphs 17-19 and 23 of Dr. Clark's rebuttal declaration (Ex. 1081) that rely upon, and cite to, the claims of Woodhill should be excluded as irrelevant, prejudicial, confusing, lacking foundation, and beyond the scope of this proceeding. PO Mot. 6. In particular, PersonalWeb argues that the "name" of a particular binary object identifier, as recited in the claims of Woodhill, is not prior art to the '791 patent because there is not sufficient written description support in Woodhill's original disclosure for that claimed subject matter. *Id*. at 6-7. In response, EMC contends that Woodhill's specification provides sufficient written description support for the "name" of a particular binary object identifier, as recited in the claims of Woodhill. Pet. Opp. 7 (Ex. 1005, 2:14-17, 7:60-8:65, 18:16-23, fig. 3).

Contrary to PersonalWeb's argument, Woodhill's original disclosure contains sufficient written description support for the "name" of a particular binary object identifier, as recited in the claims of Woodhill. Upon reviewing the description of Binary Object Identification record 58 in Woodhill's original disclosure, the only part of the record that uniquely identifies the binary object associated therewith is Binary Object Identifier 74. Ex. 2007, 26, 33-34. During Woodhill's self-auditing procedure, Distributed Storage Manager program 24 uses Binary Object Identifier 74 to access a randomly selected binary object by retrieving its corresponding Binary Object Identification record 58 in File Database 25. *See* Ex. 2007, 53. Dr. Clark confirms such an operation was routine because it was old and

well-known to access records stored in a database using their identifiers. *See* Ex. 1081, ¶ 20.

Based on the cited portions in Woodhill's original disclosure, as well as Dr. Clark's corroborating testimony, we are persuaded that Woodhill's original disclosure conveys with reasonable clarity to one with ordinary skill in the art that Binary Object Identifier 74 may be considered a "name" for a binary object associated therewith because it uniquely identifies that binary object. *See Ariad Pharms., Inc. v. Eli Lilly & Co.*, 598 F.3d 1336, 1351 (Fed. Cir. 2010) (en banc) (The written description test is whether the original disclosure of the application relied upon reasonably conveys to a person of ordinary skill in the art that the inventor had possession of the claimed subject matter as of the filing date.)

Accordingly, we are not persuaded that PersonalWeb has presented a sufficient basis to exclude paragraphs 17-19 and 23 of Dr. Clark's rebuttal declaration that rely upon, and cite to, the "name" of a particular binary object identifier, as recited in the claims of Woodhill.

F. EMC's Motion to Exclude

EMC seeks to exclude three license agreements (Exs. 2010-12), as well as the two declarations offered by Mr. Kevin Bermeister (Exs. 2009, 2014) relating to those license agreements, because they are irrelevant under FRE 401, highly prejudicial, confusing, and misleading under FRE 403. Paper 65. PersonalWeb opposes EMC's motion to exclude. Paper 72. In response, EMC filed a reply to PersonalWeb's opposition to its motion to exclude. Paper 75.

The current situation does not require us to assess the merits of EMC's motion to exclude. As discussed above, even without excluding the

three license agreements (Exs. 2010-12) and the two declarations offered by Mr. Bermeister (Exs. 2009, 2014), we have concluded that EMC has demonstrated by a preponderance of the evidence that the challenged claims are unpatentable. Accordingly, EMC's motion to exclude evidence is dismissed as moot.

III. CONCLUSION

EMC has demonstrated by a preponderance of the evidence that claims 1-4, 29-33, and 41 of the '791 patent are unpatentable based on the grounds of unpatentability set forth in the table below.

Claims	**Basis**	**Reference**
1-4, 29-33, and 41	§ 102(e)	Woodhill
1-4 and 29	§ 103(a)	Woodhill

IV. ORDER

In consideration of the foregoing, it is

ORDERED that, based on a preponderance of the evidence, claims 1-4, 29-33, and 41 of the '791 patent are unpatentable;

FURTHER ORDERED that PersonalWeb's motion to exclude evidence is DENIED;

FURTHER ORDERED that EMC's motion to exclude evidence is DISMISSED as moot; and

FURTHER ORDERED that, because this is a final written decision, parties to this proceeding seeking judicial review of our decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

For PETITIONERS:

Peter Dichiara
David L. Cavanaugh
WILMER CUTLER PICKERING HALE AND DORR LLP
Peter.Dichiara@wilmerhale.com
David.Cavanaugh@wilmerhale.com

For PATENT OWNERS:

Joseph A. Rhoa
Updeep S. Gill
NIXON & VANDERHYE P.C.
jar@nixonvan.com
usg@nixonvan.com