Trial@uspto.gov
571-272-7822

Paper 14
Entered: May 17, 2013

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

EMC CORPORATION
Petitioner,

v.

PERSONALWEB TECHNOLOGIES, LLC
Patent Owner.

Case IPR2013-00084 (JYC)
Patent 7,945,544

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

CHANG, *Administrative Patent Judge*

DECISION
Institution of *Inter Partes* Review
37 C.F.R. § 42.108

I. INTRODUCTION

EMC Corporation ("EMC") filed a petition, requesting an *inter partes* review of U.S. Patent 7,945,544 ("the '544 patent"). (Paper 3, "Pet.") In response, PersonalWeb Technologies LLC ("PersonalWeb") filed a patent owner preliminary response. (Paper 15, "Prel. Resp.") We have jurisdiction under 35 U.S.C. § 314.

The standard for instituting an *inter partes* review is set forth in 35 U.S.C. § 314(a) which provides:

> THRESHOLD -- The Director may not authorize an inter partes review to be instituted unless the Director determines that the information presented in the petition filed under section 311 and any response filed under section 313 shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.

Upon consideration of the petition and patent owner preliminary response, we determine that the information presented in the petition demonstrates that there is a reasonable likelihood that EMC would prevail with respect to claim 1, the sole challenged claim. Accordingly, we authorize an *inter partes* review to be instituted for claim 1 of the '544 patent.

A. Related Proceedings

EMC indicates that the '544 patent is the subject of litigation styled *PersonalWeb Technologies LLC v. EMC Corporation and VMware, Inc.*, No 6:11-cv-00660-LED (E.D. Tex.). (Pet. 1.)

EMC also filed five other petitions seeking *inter partes* review of the following patents: Patent 5,978,791 (IPR2013-00082), Patent 6,415,280 (IPR2013-00083), Patent 7,945,539 (IPR2013-00085), Patent 7,949,662 (IPR2013-00086), and Patent 8,001, 096 (IPR2013-00087). (Pet. 1.) According to EMC, those patents and the '544 patent share a common disclosure. (*Id*. citing to Ex. 1008.)

The '544 patent claims the benefit of various applications and patents under 35 U.S.C. § 120, but the earliest priority date claimed by the '544 patent is April 11, 1995, the filing date of U.S. patent application No. 08/425,160 (now abandoned). (Ex 1001, front page; *see also* Ex. 1008.)

B. The '544 Patent

The '544 patent relates to a method for identifying a data item (*e.g.*, a data file or record) in a data processing system, by using an identifier which depends on all of the data in the data item and only on the data in the data item. (Ex. 1001, 1:45-49; 3:53-56.) Thus, the identity of a data item is said to be independent of its name, origin, location, and address. (Ex. 1001, 3:56-59.) According to the '544 patent, it is desirable to have a mechanism for identifying identical data items to reduce duplicate copies of a data item. (Ex. 1001, 3:37-40.)

Figure 10(b) of the '544 patent, reproduced below, is a flow chart for determining an identifier (True Name) of a simple or compound data item.

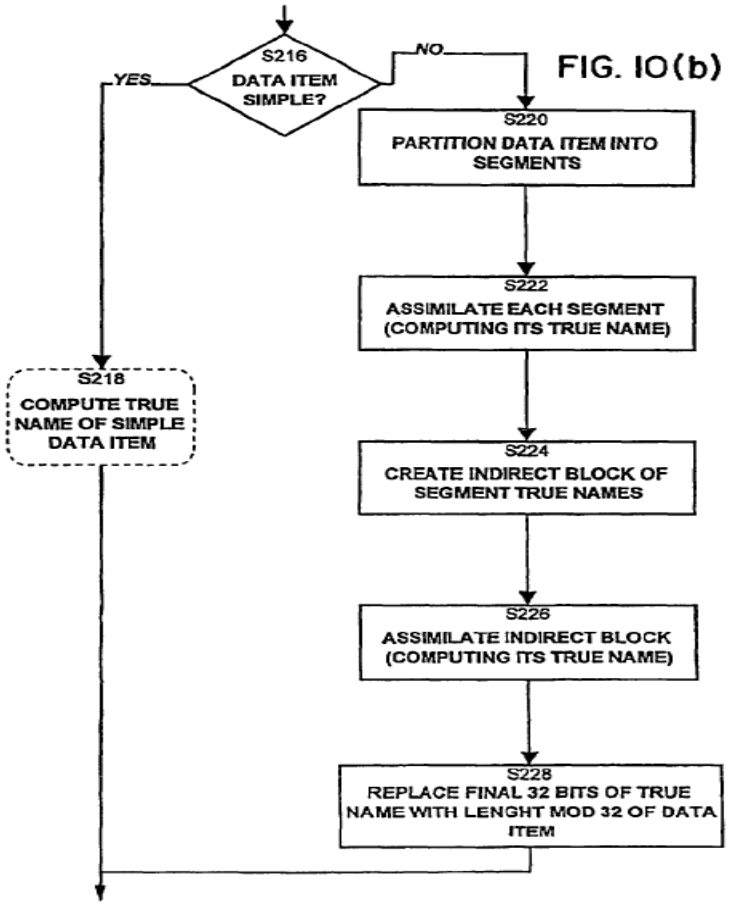


FIG. IO(b)

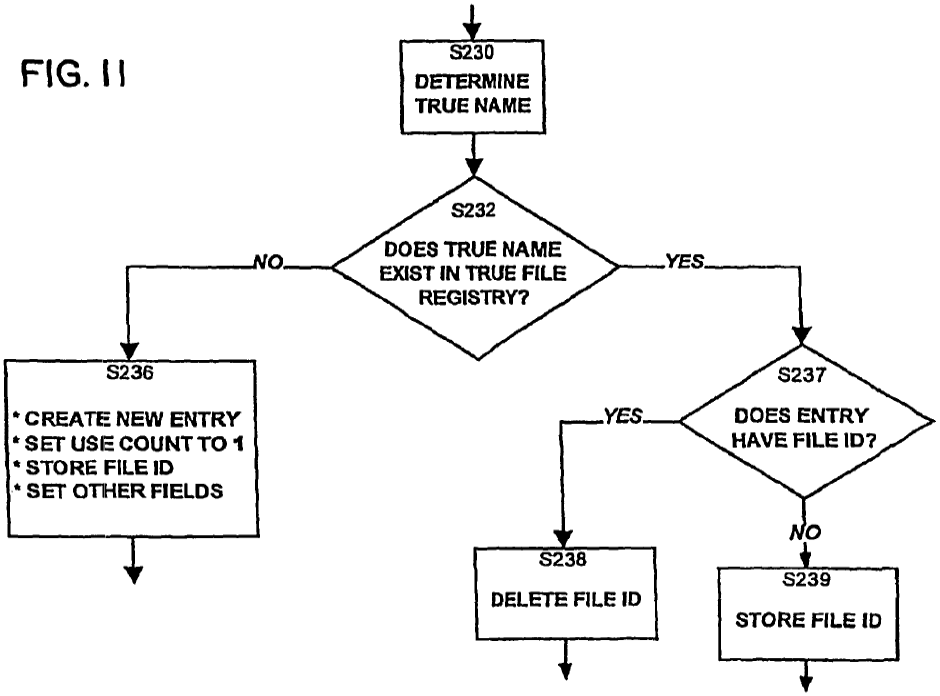
For a simple data item (a data item whose size is less than a particular given size) (S216 and S218), a data identifier (True Name) is computed using a function (*e.g.*, a message digest ("MD") function, such as MD4 or MD5, or a secure hash algorithm ("SHA") function). (Ex. 1001, 12:18-44, 13:31-42, Figs. 10(a) & 10(b).) As a result, a data item that has an arbitrary length is reduced to a relatively small, fixed size identifier (True Name) that represents the data item. (*Id.*)

As shown in Figure 10(b), if the data item is a compound data item (a data item whose size is greater than the particular given size), the system

will partition the data item into segments (S220); assimilate each segment (S222); compute the True Name of the segment; create an indirect block consisting of the computed segment True Names (S224); assimilate the indirect block (S226); and replace the final 32 bits of the resulting True Name by the length MOD 32 of the compound data item (S228). (Ex. 1001, 13:43-61, Fig. 10(b).) The result is the True Name of the compound data item. (*Id*.)

Figure 11 of the '544 patent, reproduced below, depicts a mechanism for assimilating a data item into a file system.

FIG. 11



The purpose of this mechanism is to add a given data item to the True File registry. (Ex. 1001, 14:4-11.) If the data item already exists in the registry, the duplicate will be eliminated. (*Id*.)

To assimilate a data item, the system will determine the True Name of

the data item corresponding to the file (S230); look for an entry for the True Name in the True File Registry (S232); and determine whether a True Name entry exists in the True File Registry (S232). (Ex. 1001, 14:4-27, Fig. 11.) If the entry record includes a corresponding True File ID (Step S237), the system will delete the file (Step S238). Otherwise the system will store the True File ID in the entry record (S239). (*Id*.) If there is no entry in the True File Registry for the True Name (S232), the system will create a new entry in the True File Registry for the True Name (S236).

C. Prior Art Relied Upon

EMC relies upon the following prior art references:

Woodhill U.S. Patent 5,649,196[1] July 15, 1997 (Ex. 1005)

Albert Langer, "*Re: dl/describe (File descriptions)*," post to the "alt.sources" newsgroup on Aug. 7, 1991 ("Langer," Ex. 1003)

Frederick W. Kantor, "*FWKCS (TM) Contents-Signature System Version 1.22*," FWKCS122.REF (Aug. 10, 1993) ("Kantor," Ex. 1004)

S. Browne et al., "*Location-Independent Naming for Virtual Distributed Software Repositories*," University of Tennessee Technical Report CS-95-278 (Feb. 1995) ("Browne," Ex. 1002)

[1] Woodhill claims the benefit of U.S. patent application No. 08/085,596, filed on July 1, 1993.

D. The Asserted Grounds

EMC challenges the patentability of claim 1of the ’544 patent based on the following grounds:

1. Claim 1 is unpatentable under 35 U.S.C. § 102(e) as anticipated by Woodhill (Pet. 50);
2. Claim 1 is unpatentable under 35 U.S.C. § 102(b) as anticipated by Kantor (Pet. 28);
3. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Kantor and Woodhill (Pet. 36);
4. Claim 1 is unpatentable under 35 U.S.C. § 102(a) as anticipated by Browne (Pet. 37);
5. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Browne and Woodhill (Pet. 43);
6. Claim 1 is unpatentable under 35 U.S.C. § 102(b) as anticipated by Langer (Pet. 44); and
7. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Langer and Woodhill (Pet. 50).

II. ANALYSIS

A. Claim Construction

As a first step in our analysis for determining whether to institute a trial, we determine the meaning of the claims. In an *inter partes* review, claim terms in an unexpired patent are given their broadest reasonable construction in light of the specification of the patent in which they appear.

37 C.F.R. § 42.100(b). Under the broadest reasonable construction standard, claims are to be given their broadest reasonable interpretation consistent with the specification, and the claim language should be read in light of the specification as it would be interpreted by one of ordinary skill in the art. *In re Am. Acad. of Sci. Tech. Ctr.*, 367 F.3d 1359, 1364 (Fed. Cir. 2004). This means that the words of the claim will be given their plain meaning unless the plain meaning is inconsistent with the specification. *In re Zletz*, 893 F.2d 319, 321 (Fed. Cir. 1989). In this regard, an inventor is entitled to be his or her own lexicographer of patent claim terms by providing a definition of the term in the specification with reasonable clarity, deliberateness, and precision. *In re Paulsen*, 30 F.3d 1475, 1480 (Fed. Cir. 1994).

In this proceeding, analysis of the grounds of unpatentability asserted by EMC depends on the meaning of the claim term "data item" as recited in claim 1 (*i.e.*, "a first *data item* comprising a first plurality of parts" and "a *second data* item comprising a second plurality of parts").

Each of the parties proposed a claim construction, and the parties appear to agree that the term "data item" has the meaning of "sequence of bits." (Pet. 6; Prel. Resp. 3.) In particular, PersonalWeb asserts that the specification of the '544 patent provides a special definition for the term. (Prel. Resp. 3, citing Ex. 1001, 2:17-18 "the terms "data" and "data item" as used herein refer to sequences of bits.") EMC also directs our attention to the following portions of the specification of the '544 patent (Pet. 6):

Thus a *data item* may be the contents of a *file*, a portion of a

> file, a *page* in memory, an *object* in an object-oriented program, a digital *message*, a digital scanned *image*, a part of a *video* or audio *signal*, or any *other entity* which can be represented by a sequence of bits.

(Ex. 1001, 2:18-22, emphasis added.)

> In all of the prior data processing systems the names or identifiers provided to identify *data items (the data items being files, directories, records in the database, objects in object-oriented programming, locations in memory or on a physical device, or the like*) are always defined relative to a specific context.

(Ex. 1001, 2:27-32, emphasis added.)

Based on our review of the specification of the '544 patent, we agree with the parties that the claim term "data item" means "sequence of bits," but also clarify that the meaning includes one of the following: (1) the contents of a file; (2) a portion of a file; (3) a page in memory; (4) an object in an object-oriented program; (5) a digital message; (6) a digital scanned image; (7) a part of a video or audio signal; (8) a directory; (9) a record in a database; (10) a location in memory or on a physical device or the like; and (11) any other entity which can be represented by a sequence of bits.

B. Claim 1 – Anticipated by Woodhill

EMC asserts that claim 1 is unpatentable under 35 U.S.C. § 102(e) as anticipated by Woodhill. (Pet. 50-57.) In support of this asserted ground of unpatentability, EMC provides detailed explanations as to how each claim limitation is met by Woodhill and supporting evidence including a

declaration of Professor Douglas W. Clark[2] ("Dr. Clark") and claim charts. (Pet. 50-57, citing to Ex. 1009, ¶¶ 43-49; Ex. 1040.)

PersonalWeb opposes and argues that Woodhill does not describe all of the limitations of claim 1. (Prel. Resp. 14-18.) We are not persuaded by PersonalWeb's arguments. Rather, we determine EMC's explanations and supporting evidence have merit.

Claimed Subject Matter

According to EMC, claim 1 essentially requires obtaining "values" for two data items, and then comparing these values to ascertain whether the two data items correspond to each other (*e.g.*, whether they are the same). (Pet. 16.) Claim 1 recites the following (emphasis and indentions added):

> 1. A computer-implemented method, the method comprising:
>
> (A) for a first data item comprising a first plurality of parts,
>
> (a1) applying a first function to each part of said first plurality of parts to obtain a corresponding part value for each part of said first plurality of parts,
>
> wherein each part of said first plurality of parts comprises a corresponding sequence of bits, and
>
> wherein the part value for each particular part of said first plurality of parts is based, at least in part, on the corresponding bits in the particular part, and

[2] Dr. Clark has a Ph.D. in computer science and extensive experience in computer systems architecture and design. (Ex. 1009, ¶¶1-6.) We conclude that Dr. Clark is qualified to testify as to the understanding of one skill in the art.

wherein two identical parts will have the same part value as determined using said first function,

wherein said first function comprises a first hash function; and

(a2) ***obtaining a first value for the first data item***, said first value obtained by ***applying a second function to the part values*** of said first plurality of parts of said first data item, said second function comprising a second hash function;

(B) for a second data item comprising a second plurality of parts,

(b1) applying said first function to each part of said second plurality of parts to obtain a corresponding part value for each part of said second plurality of parts,

wherein each part of said second plurality of parts consists of a corresponding sequence of bits, and

wherein the part value for each particular part of said second plurality of parts is based, at least in part, on the corresponding bits in the particular part of the second plurality of parts; and

(b2) ***obtaining a second value*** for the second data item by applying said second function to the part values of said second plurality of parts of said second data item; and

(C) ascertaining whether or not said ***first data item corresponds to said second data item*** based, at least in part, on said first value and said second value.

Prior Art – Woodhill

Woodhill discloses a system for distributed storage management on a computer network system. (Ex. 1005, 1:11-17.) Figure 1 of Woodhill, reproduced below, depicts a computer network system that includes a

distributed storage management system.

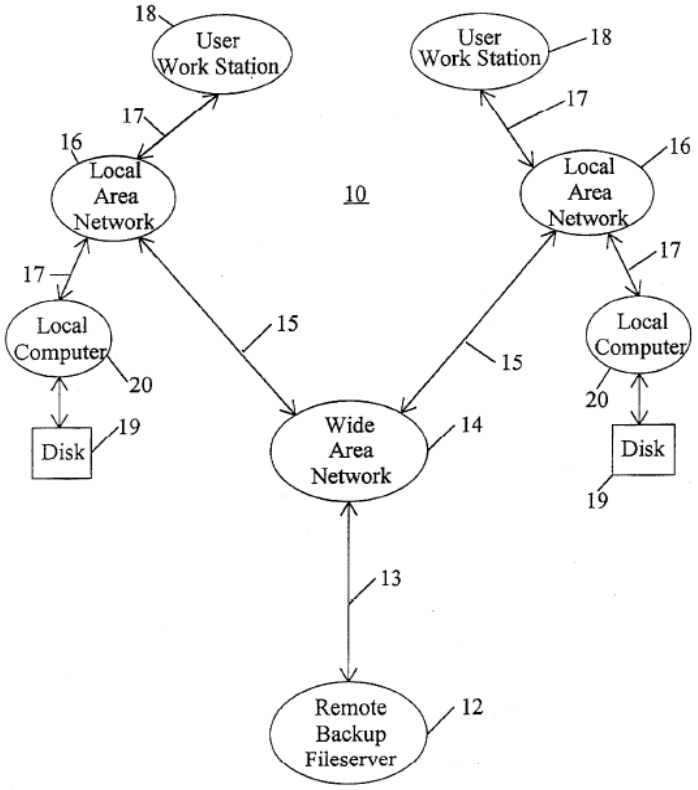


FIG. 1

As illustrated in Figure 1 of Woodhill, each local area network 16 includes multiple user workstations 18 and local computers 20. (Ex. 1005, 3:24-44.) Woodhill's system includes a Distributed Storage Manager (DSM) program for building and maintaining the File Database. (Ex. 1005, 3:44-49.)

The DSM program views a file as a collection of data streams, and divides each data stream into one or more binary objects. (Ex. 1005, 4:13-23; 7:40-43; Fig. 5A, item 132.) More specifically, data streams represent regular data, extended attribute data, access control list data, etc. (Ex. 1005,

7:44-47.) If the size of the data stream is larger than the maximum binary object size, then the DSM program divides the data stream into multiple binary objects; otherwise, a single binary object represents the data stream. (Ex. 1005, 4:23-30; 7:47-59; Fig. 5A, items 134 and 136.) For each binary object being backed up, a Binary Object Identification Record is created in a File Database and includes a Binary Object Identifier to identify a particular binary object uniquely. (Ex. 1005, 7:60-8:1; 8:33-34.)

Binary Object Identifiers are calculated based on the contents of the data instead of from an external and arbitrary source so that the Binary Object Identifier changes when the contents of the binary object changes. (Ex. 1005, 8:57-62; 8:40-42.) Notably, the Binary Object Identifier includes a Binary Object Hash field which is calculated against the contents of the binary object taken one word (16 bits) at a time using a hash algorithm. (Ex. 1005, 8:22-32.) According to Woodhill, duplicate binary objects can be recognized from their identical Binary Object Identifiers, even if the objects are resided on different types of computers in a heterogeneous network. (Ex. 1005, 8:62-65.)

For large database files on the network computer system, the DSM program utilizes a technique of subdividing the large database files into "granules" and then tracks changes from the previous backup copy of the "granule" level. (Ex. 1005, 14:53-65.) This technique is used to reduce the amount of data that must be transmitted to the remote backup file server. (Ex. 1005, 15:4-8.) Figure 5G of Woodhill illustrates the "granularization" procedure and is reproduced below.

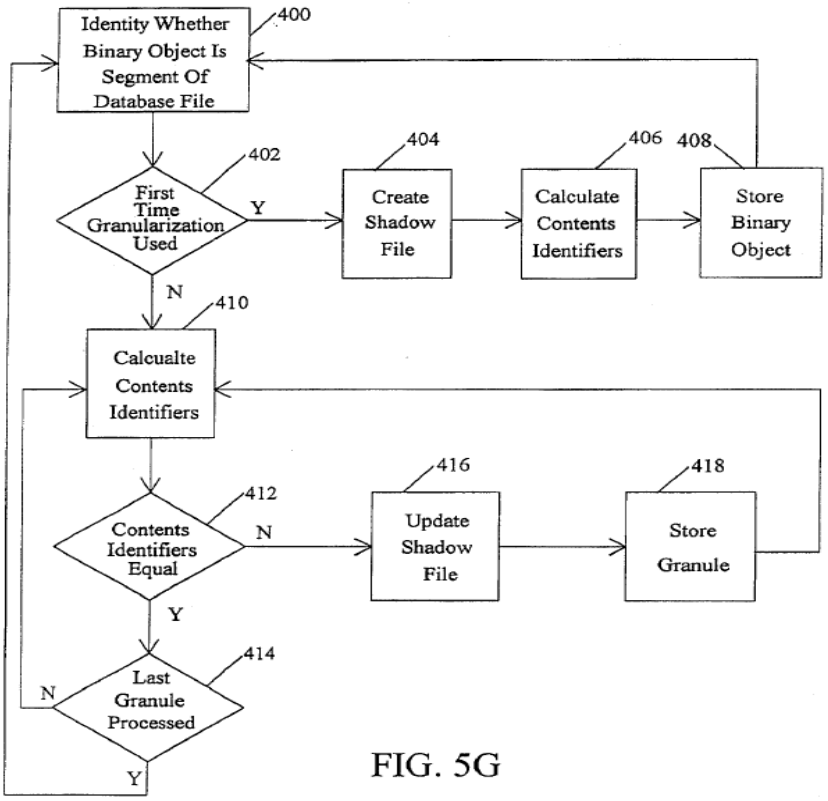


FIG. 5G

As depicted in Figure 5G, if this is the first time that the binary object is being backed up using the "granularization" technique (step 402), the DSM program creates a "shadow file," which contains a "contents identifier" for each "granule" in the binary object (step 404). (Ex. 1005, 15:9-24.) Each "contents identifier" includes a 32-bit hash number which is calculated against the contents of the "granule." (Ex. 1005, 15:24-30; Fig. 5A, step 138)

Each time that the binary object is backed up, the DSM program calculates the "contents identifier" for each "granule" in the binary object, and then compares it to the "contents identifier" of the "granule" the last time the binary object was backed up to determine if the "granule" has

changed. (Ex. 1005, 15:32-38.) At step 406, the DSM program calculates a "change identifier" for each "granule" of the binary object and stores it in the "shadow file" for that binary object. (Ex. 1005, 15:40-45.)

Discussion

PersonalWeb argues that Woodhill does not anticipate claim 1. Specifically, PersonalWeb contends that: (1) "Woodhill fails to disclose applying a second function comprising a hash to either the contents of shadow files or to the contents of Binary Object Identifiers 74 (there is no hash of hashes in Woodhill);" and (2) Woodhill's identifier for the shadow file is not a "first value" for the first data item. (Prel. Resp. 14-18.) PersonalWeb further argues that "there is no disclosure in Woodhill that a Binary Object Identifier is ever created for a 'shadow file.'" (Prel. Resp. 17.) In essence, PersonalWeb argues that Woodhill does not describe step (a2) of claim 1, which recites "obtaining a *first value for the first data item*, said first value obtained by *applying a second function to the part values* of said first plurality of parts of said first data item, said second function comprising a second hash function" (emphasis added).

We are not persuaded by PersonalWeb's arguments because they narrowly focus only on Woodhill's disclosure of the "granularization" technique (Ex. 1005, 14:52-16:15). PersonalWeb fails to recognize that Woodhill's other basic back-up processes also apply to shadow files (*e.g.*, Ex. 1005, 7:60-8:65; fig. 5a, item 138).

As noted by EMC, "shadow files, like all files stored by Woodhill, are divided into binary objects to be backed up." (Pet. 55.) Directing attention to specific portions of Woodhill and the declaration of Dr. Clark (Pet. 55, citing to Ex. 1005, 4:13-34; 5:61-63, Ex. 1009 ¶¶ 46-48), EMC further explains the following:

> By calculating Binary Object Identifiers for shadow file binary objects, *Woodhill "obtain[s] a first value for the first data item" (a Binary Object Identifier for a binary object of the shadow file), and that this "first value" is obtained by applying a "second function" comprising a "second hash function"* (again, a 32-bit hash function) to the part values of the plurality of parts (the Binary Object Identifier is calculated applying the hash function to the contents of a shadow file binary object, that content being granule contents identifiers). (Clark Decl. at ¶¶ 45-48; Ex. 1009; Woodhill at col. 7, l. 60 – col. 8, l. 31; *see also* col. 15, ll. 16-24; Ex. 1005.)

(Pet. 55-56, emphasis added.)

In support of EMC's asserted ground of unpatentability, Dr. Clark testifies the following:

> 46. Prior to backing up a binary object using the granularization technique for the first time, the local computer storing the binary object creates a "shadow file" containing the granule contents identifiers for each granule of that binary object. (*Id*. at col. 15, ll. 16-24; Ex 1005.) *Woodhill also discloses claim portions [1c] and [1e][3] through his process of creating shadow files on local computers to store the latest granule contents identifiers for granularized binary objects, and then backup*

[3] "Claim portions [1c] and [1e]" are referring to steps (a2) and (b2) of claim 1. (Ex. 1009, ¶ 16.)

> *these shadow files*. In particular, a shadow file, including each contents identifier for each granule of a binary object, like any file will be divided into one or more Binary Objects. In some cases, due to the concise nature of a shadow file, a shadow file may be backed up using a single binary object.
>
> 47. As I have illustrated, *each shadow file binary object, like all binary objects, has a corresponding Binary Object Identifier*. Further, each Binary Object Identifier includes a hash of the contents of the Binary Object. Consequently, a Binary Object Identifier for a shadow file binary object satisfies these claim elements because it is *a hash (second function) of the contents identifiers, or granule hashes (i.e., "part values" of the plurality of parts [granules]).*

(Ex. 1009, ¶¶ 46-47, emphasis added.)

On this record, we credit the testimony of Dr. Clark that a shadow file binary object has a corresponding Binary Object Identifier which includes a hash (second function) of the contents of the shadow file binary object.

PersonalWeb also argues that "[w]hile Woodhill uses a hash algorithm, it is not used as a file name or as a substitute for the file name." (Prel. Resp. 13.) However, this argument is not commensurate with the scope of claim 1 ("a first value *for* the first data item"), which does not require the hash value to be a file name or a substitute for the file name. *See In re Self*, 671 F.2d 1344, 1348 (CCPA 1982) (limitations not appearing in the claim cannot be relied upon for patentability).

Accordingly, we are persuaded by EMC's analysis and supporting evidence, and thereby determine that Woodhill describes step (a2) as recited in claim 1. On this record, we conclude EMC has demonstrated that there is

a reasonable likelihood that it would prevail with respect to claim 1 based on the ground that Woodhill anticipates claim 1.

C. Claim 1 – Anticipated by Kantor, and Obvious over Kantor and Woodhill

EMC asserts that claim 1 is unpatentable under 35 U.S.C. § 102(b) as anticipated by Kantor and, alternatively, unpatentable under 35 U.S.C. § 103(a) as obvious over Kantor in view of Woodhill. (Pet. 28-36.) EMC also submits a declaration of Dr. Clark as supporting evidence. (*Id*., citing to Ex. 1009.) EMC contends that Kantor is a published manual that describes a software program called the Frederick W. Kantor Contents Signature System Version 1.22 ("FWKCS"). (Pet. 28, citing to Ex. 1004, Title Page.)

Whether Kantor is a "Printed Publication"

In its preliminary response, PersonalWeb does not dispute the substantive disclosure of Kantor, but rather urges the Board to deny these asserted grounds on the basis that Kantor is not a prior art "printed publication." (Prel. Resp. 4-9.) In particular, PersonalWeb argues that EMC has presented no testimony, declaration, or other evidence that Kantor "was catalogued or indexed in a meaningful way prior to the critical date, or that [it] would have turned up in a customary search prior to the critical date, or that persons interested and ordinarily skilled in the art exercising reasonable diligence would have located [it] prior to the critical date." (Prel. Resp. 5-6.)

We are not persuaded by PersonalWeb's arguments. Rather, on this record, we determine that EMC has made a threshold showing to establish

that Kantor is a "printed publication" within the meaning of 35 U.S.C. § 102(b). As a consequence, Kantor is available as prior art for the purposes of this decision to demonstrate that claim 1 of the '544 patent is unpatentable under 35 U.S.C. §§ 102(b) and 103(a).

To determine whether to deny a ground on the basis that a reference is not a "printed publication," we decide each case on the basis of its own facts. More specifically, the determination of whether a given reference qualifies as a prior art "printed publication" involves a case-by-case inquiry into the facts and circumstances surrounding the reference's disclosure to members of the public. *In re Klopfenstein*, 380 F.3d 1345, 1350 (Fed. Cir. 2004).

Here, EMC asserts that Kantor has been publicly available since August 1993, which is prior to the critical date, namely one year before April 11, 1995, the earliest priority date claimed by the '544 patent. (Pet. 3, citing to Ex. 1004.) EMC also proffers the following explanation (*id*.):

> Kantor's FWKCS user manual has been publicly and freely available continuously since August 1993. Kantor distributed the user manual with the FWKCS program as shareware and posted it online to electronic Bulletin Board Systems including "The Invention Factory" and "Channel 1" for an extended period of time, where it could be downloaded by anyone. As such, the document was accessible to others in the relevant community of BBS users and system operators. (*See* Kantor at 3; *see also* 158-59; Ex. 1004.)

Further, the title page of Kantor clearly shows the posted date of August 10, 1993. (Ex. 1004, Title Page "FWKCS (TM) Contents_Signature

System[,] Version 1.22[,] **1993 August 10** [,] (C) Copyright Frederick W. Kantor 1988-1993." Emphasis added.) Kantor also provides the following:

> The FWKCS(TM) Contents_Signature System has become a robust platform for supporting contents_signature functions. FWKCS provides many functions and options for application in a public, commercial, school, institutional, or governmental environment. Extensive technical support is of special value in helping such users to benefit more fully from these many features.
>
> Registered FWKCS hobby BBS users are able to receive a modest amount of assistance, and are invited to participate in the FWKCS conference on The Invention Factory BBS, echoed via Execnet.
>
> Commercial, school, institutional, and governmental users, with their special support needs, are invited to discuss terms for obtaining such assistance.
>
> To get a new version of FWKCS, download FWKCSnnn.ZIP from The Invention Factory BBS, where nnn is the new version number without a decimal point. These special downloads are available at no fee, from a 43_line hunt_up group of USR Dual Standard modems, at 2400-16800 bits/sec (including V32.bis).

(Ex. 1004, 158-159.)

Given that, Kantor appears to convey that the reference was posted on a publicly accessible site well known to those interested in the art – the electronic Bulletin Board Systems – and could be downloaded and retrieved from that site. *In re Wyer*, 655 F.2d 221, 226 (CCPA 1981) (An electronic publication, including an on-line database or Internet publication, is considered to be a "printed publication" "upon a satisfactory showing that such document has been disseminated or otherwise made available to the

extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, can locate it and recognize and comprehend therefrom the essentials of the claimed invention without need of further research or experimentation.").

PersonalWeb cites *Synopsys, Inc. v. Mentor Graphics Corp.*, IPR2012-00042 (Paper No. 16), at *35-36 (PTAB Feb. 22, 2013) for the proposition that any asserted grounds of unpatentability based on an electronic reference should be denied, unless the reference is presented in the petition with a declaration from one of the authors or other evidence that someone accessed or received the reference prior to the critical date. (Prel. Resp. 4-6.) PersonalWeb's reliance on *Synopsys* is misplaced because *Synopsys* did not involve a reference, such as Kantor, that has a posted date. In fact, *Synopsys* involves a brochure that did not include any indication of when it was created or whether it was disseminated publicly, and the only evidence submitted by the petitioner was that it was cited in an Information Disclosure Statement filed in an unpublished patent application. *Synopsys*, IPR2012-00042 (Paper No. 16), at *35.

As to PersonalWeb's argument that there is no evidence that Kantor was catalogued or indexed in a meaningful way prior to the critical date, we are not convinced. "[W]hile often relevant to public accessibility, evidence of indexing is not an absolute prerequisite to establishing online references [] as printed publications within the prior art." *Voter Verified, Inc. v. Premier Election Solutions, Inc.*, 698 F.3d 1374, 1380 (Fed. Cir. 2012).

PersonalWeb further argues that Kantor is inadmissible evidence as the copy of Kantor submitted by EMC has not been authenticated or certified. In that regard, PersonalWeb has not followed the proper procedures for objecting to and/or excluding evidence. *See* 37 C.F.R. § 42.64(b); *LKQ Corp. v. Clearlamp, LLC*, IPR2013-00020 (Paper No. 17), at *3-4 (PTAB Mar. 5, 2013).

Under the procedure set forth in 37 C.F.R. § 42.64(b), when a party objects to evidence that was submitted during a preliminary proceeding, such an objection must be served within ten business days of the institution of trial. The objection to the evidence must identify the grounds for the objection with sufficient particularity to allow correction in the form of supplemental evidence. This process allows the party relying on the evidence to which an objection is served timely, the opportunity to correct, by serving supplemental evidence within so many days of the service of the objection. If, upon receiving the supplemental evidence, the opposing party is still of the opinion that the evidence is inadmissible, the opposing party may file a motion to exclude such evidence. The time for filing a motion to exclude is typically several months into a trial. *See, e.g., Office Patent Trial Practice Guide*, 77 *Fed. Reg*. 48756, 48768-69, Scheduling Order – Due Date 4. Therefore, PersonalWeb will have full opportunity to object, serve, and reconsider any supplemental evidence and, finally, file a motion to exclude evidence.

Anticipation and Obviousness

In EMC's view, Kantor discloses all of the limitations as recited in claim 1, including a "data item" (a zipfile) comprising a "first plurality of parts" (the files contained within the zipfile). (Pet. 34-36, citing to Ex. 1004, 2-4, 48-49.)

Kantor describes a method of identifying duplicate files. (*Id*.) In particular, Kantor applies a hash function (*e.g.*, a cyclic residue check or cyclic redundancy check (CRC)) to each file within the zipfile to obtain the contents signature for each file. (Ex. 1004, 6-8, 48-49.) Each contents signature is a string of bits generated from the contents of a file. (*Id*.)

For each zipfile, Kantor creates "zipfile contents signatures" by hashing the contents signatures for the files contained within the zipfile. (Ex. 1004, 2 & 9.) As Kantor points out, this is done by "adding together all the 32_bit CRC's for the files in the zipfile, modulo 2^32, separately adding together their uncompressed file_lengths module 2^32, and then arranging the two resulting hexadecimal number as a single structure." (*Id*. at 9.) Dr. Clark testifies that "additional modulo 2^32" is another well-known simple hashing function that uses addition to calculate a value for a file based on the file's contents. (Ex. 1009, ¶ 20.) Kantor further compares the zipfile contents signatures to check for duplicate files. (Ex. 1004, Preface-2, 5 & 9.)

As to the obviousness ground of unpatentability, EMC contends that in the event that Kantor does not satisfy the claim limitation of a "plurality of parts" of a data item, a person of ordinary skill would have found it

obvious to modify Kantor to meet that limitation in view of Woodhill. (Pet. 36.) As discussed previously, Woodhill discloses dividing a data item into a plurality of parts (*e.g.*, dividing files into "binary objects," and further dividing the binary object into "granules") to reduce the amount of data that needs to be transmitted (*i.e.*, transmitting smaller file segments instead of the entire file). (Ex. 1005, 4:14-30; 14:52-15:8.)

We have reviewed EMC's analysis and supporting evidence, and we determine that EMC's contentions have merit. For the foregoing reasons, we conclude that EMC has demonstrated that it is more likely than not that claim 1 is unpatentable as being anticipated by Kantor and, alternatively, as obvious over Kantor and Woodhill.

D. Other Asserted Grounds

EMC also asserted the following grounds of unpatentability:

1. Claim 1 is unpatentable under 35 U.S.C. § 102(a) as anticipated by Browne;
2. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Browne and Woodhill;
3. Claim 1 is unpatentable under 35 U.S.C. § 102(b) as anticipated by Langer; and
4. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Langer and Woodhill.

However, these asserted grounds are unnecessary in light of the determination that there is a reasonable likelihood that claim 1 is

unpatentable based on the grounds on which we institute an *inter partes* review. We therefore exercise our discretion to deny these grounds as redundant. *See* 37 C.F.R. § 42.208.

III. CONCLUSION

For the forgoing reasons, we determine that the information presented in EMC's petition shows that there is a reasonable likelihood that EMC would prevail with respect to claim 1 of the '544 patent. Accordingly, the petition is granted.

IV. ORDER

It is

ORDERED that pursuant to 35 U.S.C. § 314, an *inter partes* review is hereby instituted for the following grounds:

1. Claim 1 is unpatentable under 35 U.S.C. § 102(e) as anticipated by Woodhill;
2. Claim 1 is unpatentable under 35 U.S.C. § 102(b) as anticipated by Kantor; and
3. Claim 1 is unpatentable under 35 U.S.C. § 103(a) over Kantor and Woodhill;

FURTHER ORDERED that pursuant to 35 U.S.C. § 314(d) and 37 C.F.R. § 42.4, notice is hereby given of the institution of a trial; the trial is commencing on the entry date of this decision; and

FURTHER ORDERED that an initial conference call with the Board is scheduled for 2:00 PM Eastern Time on June 3, 2013; the parties are directed to the Office Trial Practice Guide, 77 *Fed. Reg.* 48756, 48765-66 (Aug. 14, 2012) for guidance in preparing for the initial conference call, and should come prepared to discuss any proposed changes to the Scheduling Order entered herewith and any motions the parties anticipate filing during the trial.

PETITIONER:

Peter M. Dichiara, Esq.
David L. Cavanaugh, Esq.
WILMER CUTLER PICKERING HALE & DORR LLP
peter.dichiara@wilmerhale.com
david.cavanaugh@wilmerhale.com

PATENT OWNER:

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