IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

NORTHEASTERN UNIVERSITY and JARG CORPORATION **Plaintiffs,** **v.** **GOOGLE INC.** **Defendant.**	Civil Action No. 2:07-CV-486-CE

GOOGLE INC.'S AMENDED INVALIDITY CONTENTIONS

TABLE OF CONTENTS

Page

I. INTRODUCTION

Pursuant to Patent Local Rule 3-6(b), Google Corporation ("Google") discloses its Amended Preliminary Invalidity Contentions for U.S. Patent No. 5,694,593 as follows:

Google's discovery and investigation in connection with this lawsuit are continuing, and, thus, these disclosures are based on information obtained to date. To the extent that Google obtains additional information, Google reserves the right to supplement these invalidity contentions.

Plaintiff has asserted Claims 1, 2, 3, 8, 9, and 13 of U.S. Patent No. 5,694,593 ("'593"). Google's Amended Invalidity Contentions are based in whole or in part on its present understanding of the asserted claims, Plaintiffs' apparent construction of the claim elements in Plaintiffs' Patent Local Rule 3-1 disclosures (served on Sept. 9, 2008), Plaintiffs' apparent construction of the claim elements in Plaintiffs' Amended Infringement Contentions (amended on March 24, 2010[1]), and Plaintiffs' proposed construction of the claim elements in Plaintiff's Patent Local Rule 4-2 disclosures (served on July 31, 2010). Accordingly, Google's Amended Invalidity Contentions (including the charts attached to Google Inc.'s Invalidity Contentions served on November 7, 2008, which are incorporated herein) take into account alternative and potentially inconsistent positions as to claim construction and scope. In this regard, Plaintiffs' disclosed and apparent constructions, as well as Plaintiffs' application of those constructions, are both confusing and at the very least broader than Google's understanding of the claim terms, and thus references that anticipate under Google's claim constructions would also

[1] The Court granted plaintiffs leave to amend their infringement contentions on March 24, 2010 (Dkt. No. 81).

anticipate under Plaintiffs' claim constructions. Further, by including prior art that would anticipate or render obvious claims based on Plaintiffs' disclosed and apparent claim constructions or any other particular claim construction, Google is not adopting Plaintiffs' claim constructions nor admitting to the accuracy of any particular claim construction. Google reserves all rights to amend these invalidity contentions after the Court issues its claim construction ruling, or if Plaintiff amends its infringement contentions.

The charts attached to Google Inc.'s Invalidity Contentions served on November 7, 2008, which are incorporated herein, list specific examples of where prior art references disclose, either expressly or inherently, each limitation of the asserted claims and/or examples of disclosures in view of which a person of ordinary skill in the art would have considered each limitation, and therefore the claim as a whole, obvious. Google endeavored to identify the most relevant portions of the references. The references, however, may contain additional support for particular claim limitations. Google may rely on uncited portions of the prior art references, other documents including statements in the cited references and file history for the '593 Patent, and expert testimony to provide context or to aid in understanding the cited portions of the references. Furthermore, where Google cites to a particular figure in a reference, the citation should be understood to encompass the caption and description of the figure and any text relating to the figure. Similarly, where Google cites to particular text referring to a figure, the citation should be understood to include the corresponding figure as well.

II. U.S. PATENT NO. ’593

A. Anticipation

Pursuant to P.R. 3-3, Google identifies the following prior art now known to it as anticipating Claims 1, 2, 3, 8, 9, and 13 of the ’593 patent, either expressly, implicitly, or inherently as understood by a person having ordinary skill in the art. Each of these prior art patents, publications, and systems anticipates the asserted claims. In some instances, Google treated certain prior art as anticipatory where certain elements are inherently present in view of based on Plaintiffs’ apparent claim construction in Plaintiffs’ infringement contentions.

The following publications are prior art under at least 35 U.S.C. §§ 102(a) and/or (b) and/or are evidence of prior invention under.

	PRIOR ART PUBLICATIONS
1.	Ijsbrand Jan Aalbersberg, Frans Sijstermans, High Quality and High Performance Full-Text Document Retrieval: The Parallel InfoGuide System, Proceedings of the First International Conference on Parallel and Distributed Information Systems, pp. 142-50, 1991
2.	William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp. 90-98, June 1988
3.	Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990 (“Prototyping_Bubba”)
4.	Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990 (“Burkowski”)
5.	Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992
6.	Itasca Systems, ITASCA: Distributed Object Database Management System,

	Technical Summary for Release 2.1, 1992
7.	Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989
8.	Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991 ("Stanfill_Information_91")
9.	Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991 ("Stanfill_Massively_91")
10.	Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992 ("Tomasic_Caching_92")
11.	Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993 ("Tomasic_Caching_93")
12.	Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds.
13.	Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992
14.	Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993

The following systems are prior art under at least 35 U.S.C. §§ 102(a), (b) and/or (g). Although Google's investigation continues, information available to date indicates that each system was (1) known or used in this country before the alleged invention of the claimed subject matter of the asserted claims, (2) was in public use and/or on sale in this country more than one year before the filing date of the patent, and/or (3) was invented by another who did not abandon, suppress, or conceal, before the alleged invention of the claimed subject matter of the asserted claims.

	PRIOR ART SYSTEMS
15.	ITASCA The ITASCA system was offered for sale, publicly used, and/or known by at least 1992. At least Douglas Barry publicly used ITASCA and/or made ITASCA publicly known. *Associated References* Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992 Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992
16.	BUBBA The BUBBA system was offered for sale, publicly used, and/or known by at least 1988. At least William Alexander, George Copeland, Haran Boral, Larry Clar, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Tom Keller, Herb Schwetman, Chii-Ren Young, and Ellen Boughter publicly used BUBBA and/or made BUBBA publicly known. *Associated References* William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988 Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990 Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989 George Copeland, William Alexander, Ellen Boughter, Tom Keller, Data Placement in Bubba, ACM, pp. 99-108, 1988

	("Data_Placement_in_Bubba")
17.	WAIS The WAIS (Wide Area Information Servers) system was offered for sale, publicly used, and/or known by at least 1991. At least Craig Stanfill, M. McCahil, B. Kahler, and A. Medlar publicly used WAIS and/or made WAIS publicly known. *Associated References* Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991 ("Stanfill_Massively_91") Mark McCahil, The Internet Gopher: A distributed server information system, ConneXions – The Interoperability Rept.Vol. 6, No. 7, pp.10-14, July, 1992 B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991

In addition, the following patents were relied upon by the Examiner during prosecution of the '593 Patent to reject claims under Section 102, and anticipate Claims 1, 2, 3, 8, 9, and 13 of the '593 patent as those claims are now apparently construed by the plaintiffs, either expressly, implicitly, or inherently as understood by a person having ordinary skill in the art:

- U.S. Patent No. 5,006,978, (Apr. 9, 1991) (Neches)
- U.S. Patent No. 4,811,199 (Mar. 7, 1989) (Kuechler et al.)

B. Obviousness

Pursuant to P.R. 3-3, Google contends that Claims 1, 2, 3, 8, 9, and 13 of the '593 patent are invalid as obvious under 35 U.S.C. § 103.

Each anticipatory prior art reference disclosed in the preceding section, either alone or in combination with other prior art, also renders the asserted claims invalid as obvious. In particular, each anticipatory prior art reference may be combined with (1) information known to persons skilled in the art at the time of the alleged invention, (2) any of the other anticipatory prior art references, and/or (3) any of the additional prior art identified below in this section. To the extent that Plaintiff contends that any of the anticipatory prior art fails to disclose one or more limitations of the asserted claims, Google reserves the right to identify portions of the primary invalidity references and additional prior art references that, when combined with other portions of the primary references, would render the claims obvious despite the allegedly missing limitation.

1. Additional Prior Art

Google identifies the following additional prior art references that either alone or in combination with other prior art (including any of the above anticipatory prior art and additional prior art disclosed in this section) renders the asserted claims invalid as obvious under 35 U.S.C. § 103.

	PRIOR ART PUBLICATIONS
18.	C. Severance, S. Pramanik, P. Wolberg, Distributed Linear Hashing and Parallel Projection in Main Memory Databases, Proceedings of the 16th VLDB Conference, Brisbane, Australia, pp. 674-682, 1990 ("Severance_90")
19.	H. Ammar, Su Deng, A Simple Dynamic Load Balancing Algorithm for Homogeneous Distributed Systems, ACM Conference on Computer Science, pp. 314-319, 1988 ("Ammar_Load_88")
20.	W. W. Chu, M. A. Merzbacher, L. Berkovich, The Design and Implementation of CoBase, SIGMOD, pp. 517-522, May 1993 ("CoBase_93")
21.	B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991
22.	Craig Stanfill, Partitioned Posting Files: A Parallel Inverted File Structure for Information Retrieval, Annual ACM Conference on Research and Development in Information Retrieval, Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 413-428, 1990

23.	Gerald Salton, Automatic Text Processing, Addison-Wesley, 1988 (“Salton_Automatic_89”)
24.	Won Kim, Introduction to Object Oriented Databases, MIT Press, 1990 (“Kim_Introduction”)
25.	Donald Knuth, The Art of Computer Science, Vol. 3, Addison-Wesley, 1973 (“Knuth_73”)
26.	Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993 (“Devine_DDH”)
27.	Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985 (“Faloutsos_Access_85”)
28.	Witold Litwin, Marie-Anne Neimat, Donovan A. Schneider, LH* - Linear Hashing for Distributed Files, International Conference on Management of Data, ACM, pp. 327-336, 1993
29.	George Copeland, William Alexander, Ellen Boughter, Tom Keller, Data Placement in Bubba, ACM, pp. 99-108, 1988 (“Data_Placement_in_Bubba”)
30.	Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992
31.	Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds. (“Frakes_92”)
32.	Erhard Rahm, Robert Marek, Dynamic Multi-Resource Load Balancing in Parallel Database Systems, University of Leipzig, Report No. 2, June 1994
33.	Bharat Bhasker, Csaba J. Egyhazy, Konstantinos P. Triantis, Non-parametric Estimation Techniques in Support of Query Decomposition Strategies for Heterogenous Distributed Database Management Systems ACM 30th Annual Southeast Conference, pp. 37-44, 1992
34.	David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990
35.	Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989)
36.	Anastasia Analyti, Sakti Pramanik, Fast Search in Main Memory Databases, ACM Sigmod, pp. 215-224 (1992)

In addition to the above prior art references, Google identifies the following patents, printed publication, and product literature that are pertinent to invalidity of the asserted claims. Google may rely on these references as invalidating prior art, evidence of the knowledge of those skilled in the art, and/or evidence to support a motivation to combine or modify other prior art. Google reserves all rights to supplement or modify these invalidity contentions and to rely on these references to prove invalidity of the asserted claims in a manner consistent with the Federal Rules of Civil Procedure and the Rules of this Court.

	Additional References
37.	Craig Stanfill, Brewster Kahle, Parallel Free-Text Search on the Connection Machine System, Communications of the ACM, December 1986, Vol. 29, No. 12, pp. 1229-1239, 1986
38.	Gerard Salton, James Allan, Chris Buckley, Automatic Structuring and Retrieval of Large Text Files, Communications of the ACM, February 1994, Vol. 37, No. 2, pp. 97-108, 1994
39.	Radek Vingralek, Yuri Breitbart, Gerhard Weikum, Distributed File Organization with Scalable Cost/Performance, SIGMOD Conference 1994, ACM, pp. 253-264, 1994
40.	Goetz Graefe, Query Evaluation Techniques for Large Databases, ACM Computing Surveys, Vol. 25, No. 2, pp. 73-170, June 1993
41.	Tim Berners-Lee, Robert Cailliau, Jean François Groff, Bernd Pollerman, Electronic Networking, Vol. 2, No. 1, pp. 52-58, Spring 1992
42.	David DeWitt, A Single-User Performance Evaluation of the Teradata Database Machine, Proceedings of the Second International Workshop on High Performance Transaction Systems, pp. 244-269, Sept. 1987
43.	Won Kim, Introduction to Object-Oriented Database Systems, Technology of Object-Oriented Languages and Systems Tools 6, Proceedings of the sixth International Conference TOOLS, p. 240, 1992

44.	Mark McCahil, The Internet Gopher: A distributed server information system, ConneXions – The Interoperability Rept.Vol. 6, No. 7, pp.10-14, July, 1992
45.	Patrick Martin, Ian A. Macleod, Brent Nordin, A Design of a Distributed Full Text Retrieval System, Organization of the 1986-ACM Conference on Research and Development in Information Retrieval, pp. 131-137, 1986
46.	The source code written on, or around, March 1993, identified in Plaintiffs' production, BATES-labeled JAR0255629, JAR0280890, JAR0280892-94, JAR0280896, and other related source code

In addition, Google incorporates by reference each and every prior art reference of record in the prosecution of the '593 Patent and related applications, as well as the prior art discussed in the specification of this patent, including, for example:

- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994 ("Chaturvedi")
- Houtsma et al., "Parallel Hierarchical Evaluation of Transitive Closure Queries," IEEE, 1991 ("Houtsma")

2. Motivation to Combine

The United States Supreme Court recently clarified the standard for what types of inventions are patentable. (*KSR International Co. v. Teleflex Inc.*, 127 S. Ct. 1727 (2007)). In particular, the Supreme Court emphasized that inventions arising from ordinary innovation, ordinary skill, or common sense should not be patentable. (*Id.* at 1732, 1738, 1742-1743, 1746). In that regard, a patent claim may be obvious if the combination of elements was obvious to try or there existed at the time of the invention a known problem for which there was an obvious solution encompassed by the patent's claims. In addition, when a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different one. If a person of ordinary skill can implement a predictable variation, Section 103 likely bars its patentability.

Because the ’593 Patent simply arranges old elements with each performing the same function it had been known to perform and yields no more than what one would expect from such an arrangement, the combination of these old elements is obvious. (*KSR Int’l Co. v. Teleflex Inc.*, 127 S. Ct. 1727, 1742 (2007)). Further, in the prior art, there were well recognized design needs and market pressures to develop information retrieval tools for non-relational distributed database systems. (*See, e.g.*, Tomasic_Caching_93, at 129 (“Information retrieval systems, of the type found in libraries, provide indexed access to the abstract of documents. . . . At the same time, an increasing number of users have access to these databases through the networks. To handle the increased load, a distributed architecture can be used, dispersing the data and index structures across several computers and performing searches in parallel. This paper studies the performance trade-offs in such a shared-nothing distributed system.”); Burkowski, at 71; Stanfill_Massively_91, at 679; Stanfill_Information_91, at 286; Faloutsos_Access_85, at 71-72). Such design needs and market pressures provide ample reason to combine prior art elements in the manner recited in the claims. (*KSR*, 127 S. Ct. at 1742). Moreover, since there were a finite number of predictable solutions, a person of ordinary skill in the art had good reason to pursue the known options. (*Id.*). Indeed, a person skilled in the art would have been familiar with all the claim elements that the patentee used to distinguish the prior art during prosecution. The above identified prior art references merely use those familiar elements for their primary or well known purposes in a manner well within the ordinary level of skill in the art.

Accordingly, common sense and the teachings of the prior art render the claims invalid under either Section 102 or Section 103.

Moreover, a person of ordinary skill would have been motivated to combine the above prior art based on the nature of the problem to be solved, the teachings of the prior art, and the knowledge of persons of ordinary skill in the art. The identified prior art address the same or similar technical issues and suggest the same or similar solutions to those issues. Moreover, some of the prior art refers to or discusses other prior art, illustrating the close technical relationship among the prior art.

To the extent that Plaintiff challenges a combination of prior art with respect to a particular element, Google reserves the right to supplement these contentions to further specify the motivation to combine the prior art. Google may rely on cited or uncited portions of the prior art, other documents, and fact and expert testimony to establish that a person of ordinary skill in the art would have been motivated to modify or combine the prior art so as to render the claims invalid as obvious.

Below are several examples of prior art combinations with respect to particular limitations. These prior art combinations are not exhaustive; rather, they are illustrative examples of the prior art combinations disclosed generally above. These exemplary combinations are alternatives to Google's anticipation and single reference obviousness contentions, and thus, they should not be interpreted as indicating that any of the individual references included in the exemplary combinations are not alone invalidating prior art under 35 U.S.C. §§ 102 and/or 103.

a. "A method for information retrieval using fuzzy queries in a non-relational, distributed database system having a plurality of home nodes and a plurality of query nodes connected by a network" (cl.1); "A non-relational, distributed database system having an information

retrieval tool for handling queries from a user, comprising: a plurality of home nodes; and a plurality of query nodes; said plurality of home nodes and said plurality of query nodes connected by a network" (cl.8); "A non relational, distributed database system having an information retrieval tool for handling queries from a user, comprising: a plurality of home nodes; and a plurality of query nodes, said plurality of home nodes and said plurality of query nodes connected by a network" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 129-130, Fig.2, Table 1
- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 129, 130, Fig. 2, Table 1.
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 243, 244, 247, Fig. 1.
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 1, 2, 4, Fig. 1.
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 8, 9, 10, Fig. 1.
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 71-74, Fig.1
- Craig Stanfill, Brewster Kahle, Parallel Free-Text Search on the Connection Machine System, Communications of the ACM, December 1986, Vol. 29, No. 12, pp. 1229-1239, 1986

- Patrick Martin, Ian A. Macleod, Brent Nordin, A Design of a Distributed Full Text Retrieval System, Organization of the 1986-ACM Conference on Research and Development in Information Retrieval, pp. 131-137, 1986
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 15-21
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 5, 11, 15, 33, 39, Fig.3
- W. W. Chu, M. A. Merzbacher, L. Berkovich, The Design and Implementation of CoBase, SIGMOD, pp. 517-522, May 1993, at 517, 520-521
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 4-6
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 285-287, 289-290, 293-294
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 679-681
- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 90, Fig. 1.1
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 35-37, Fig. 2.1
- Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993, at 101, 104-106, 108
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, V. 17, No. 1, pp. 49-74, 1985, at 50-51, 72
- David DeWitt, A Single-User Performance Evaluation of the Teradata Database Machine, Proceedings of the Second International Workshop on High Performance Transaction Systems, pp. 244-269, Sept. 1987, at 245
- Tim Berners-Lee, Robert Cailliau, Jean François Groff, Bernd Pollerman, Electronic Networking, Vol. 2, No. 1, pp. 52-58, Spring 1992, at 54, 56
- George Copeland, William Alexander, Ellen Boughter, Tom Keller, Data Placement in Bubba, ACM, pp. 99-108, 1988, at 100
- B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991, at 2-3, 5, Fig. 3
- Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992, at 108-109, 131-138, 167-168
- Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds., at 28-43, 363-392
- Bharat Bhasker, Csaba J. Egyhazy, Konstantinos P. Triantis, Non-parametric Estimation Techniques in Support of Query Decomposition Strategies for

Heterogeneous Distributed Database Management Systems ACM 30th Annual Southeast Conference, pp. 37-44, 1992, at 37

- Goetz Graefe, Query Evaluation Techniques for Large Databases, ACM Computing Surveys, Vol. 25, No. 2, pp. 73-170, June 1993, at 74-75, 126-127, 135
- Erhard Rahm, Robert Marek, Dynamic Multi-Resource Load Balancing in Parallel Database Systems, University of Leipzig, Report No. 2, June 1994, at 2
- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994, at Abstract, 196, 199-200
- Gerard Salton, James Allan, Chris Buckley, Automatic Structuring and Retrieval of Large Text Files, Communications of the ACM, February 1994, Vol. 37, No. 2, pp. 97-108, 1994, at 97-99
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, IWDM, pp. 112-126 (1989), at 112-115, 124-125, Figs. 1-3, Fig. 8
- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 1, 2, 6, 19

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the above-identified anticipatory references with one or more of the particular references listed above to potentially improve performance through parallelism by taking advantage of the availability of multiple nodes that exist in distributed database systems. For example, some of the nodes could be used for handling query requests and interactions with users, and some nodes could be used for processing the query and retrieving the answer. (*See*, *e.g.*, Prototyping_Bubba, at 5 ("Bubba contains three types of nodes. . . . The IP's provide communication with external host machines and coordinate execution of user requests. The IR's collectively store the database and perform most of the work in executing transaction programs.")).

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, developers of object-oriented database systems could look to solutions already developed for distributed relational database systems. (Kim_Introduction, at 184 ("Solutions developed for distributed relational database systems apply directly to object-oriented database systems; as such, extending an object-oriented database system from a single-processor or a client server architecture to a distributed architecture is mostly an engineering exercise.")). As another example, in CoBase_93, a distributed database is discussed "that integrates knowledge base technology with database systems to provide cooperative (approximate and conceptual) query answering." (CoBase_93, at 517). That database works in combination with other distributed databases. (*Id.*, at 521).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. It was well known in the art to use some nodes for handling query interaction with a user, and some nodes for processing the query request. (*See* Prototyping_Bubba, at 5). Therefore, a person of ordinary skill in the art would find it obvious to use two types of nodes (*i.e.*, home and query nodes) in a non-relational, distributed database system.

b. "randomly selecting a first one of said plurality of home nodes;" (cl.1)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- George Copeland, William Alexander, Ellen Boughter, Tom Keller, Data Placement in Bubba, ACM, pp. 99-108, 1988, at 102, 106
- H. Ammar, Su Deng, A Simple Dynamic Load Balancing Algorithm for Homogeneous Distributed Systems, ACM Conference on Computer Science, pp. 314-319, 1988, at 314-315, 318
- Erhard Rahm, Robert Marek, Dynamic Multi-Resource Load Balancing in Parallel Database Systems, University of Leipzig, Report No. 2, June 1994, at 2, 19

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because distributed computer systems, especially parallel ones, will benefit from some sort of load balancing technique. (*See*, *e.g.*, Data_Placement_in_Bubba, at 106 (discussing importance of considering load balancing in the early design of a parallel system)).

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, Data_Placement_in_Bubba discloses that "[r]andom-based declustering deliberately destroys most potential block locality due to record locality in order to buy better load balancing, so that only relation locality matters." (Data_Placement_in_Bubba, at 102). It continues, "a particular treatment of locality that [the authors] believe effectively accomplishes [a] compromise for shared-nothing data-intensive architectures" where data placement "boils down to compromising load balancing with overall load reduction, in the face of various kinds of data locality." (*Id.*, at 106). Such "load balancing, in particular declustering and reorganization, must be considered early in the design of such a highly-parallel system as Bubba." (*Id.*) Thus, persons of ordinary skill in the art would have the motivation to combine the load balancing techniques in

Data_Placement_in_Bubba with other shared-nothing architectures, including other references regarding the Bubba system.

As a further example, Ammar_Load_88 discloses "load balancing policies in a homogeneous distributed system," and states "[s]ome of the considerations include: static algorithms (e.g., random and cyclic splitting algorithms, where each source randomly assigns its jobs to the server in the system or assigns its ith arriving job to the ith server), and dynamic algorithms." (Ammar_Load_88 at 314). Thus, persons of ordinary skill in the art would have the motivation to combine the load balancing techniques discussed in Ammar_Load_88 with other references regarding distributed systems.

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. For example, in a distributed system, an incoming job (*e.g.*, query) has to be assigned to a node. The references above disclose methods for assigning that job, including randomly selecting a node.

c. "fragmenting, by said selected home node, a query from a user into a plurality of query fragments" (cl.1); "fragments said query into a plurality of query fragments" (cl.8); "fragmenting a query contained in said query command into a plurality of query fragments" (cl.13); "receiving at said home node, said query from said user, prior to the step of fragmenting said query" (cl.2)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University

Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 2-3, 7-9

- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 130, 132-133.
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 246, 249, 260.
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 3, 5, 12, 13.
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 9, 10, 11, 13, 14.
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9, 16
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 15, 21, 33, 50, Fig. 18
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 285, 290-294
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 680-681
- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 91
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 5-6, 14
- Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993, at 105
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 53, Fig. 1
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 36
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 72-74

- Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992, at 118-122
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989), at 112-115, 124-125, Figs. 1-3, Fig. 8
- Goetz Graefe, Query Evaluation Techniques for Large Databases, ACM Computing Surveys, Vol. 25, No. 2, pp. 73-170, June 1993, at 18, 56, 127
- Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds., at 7-8
- Bharat Bhasker, Csaba J. Egyhazy, Konstantinos P. Triantis, Non-parametric Estimation Techniques in Support of Query Decomposition Strategies for Heterogeneous Distributed Database Management Systems ACM 30th Annual Southeast Conference, pp. 37-44, 1992, at 37
- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994, at Abstract, 195-196, 198
- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 7

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because a distributed system may run more efficiently using parallelism, and fragmenting queries into separate fragments may allow the distributed system to better take advantage of parallelism—*i.e.* by having different parts of the system handle different fragments at the same time. (Tomasic_Caching_92, at 1) (regarding information retrieval systems, "[t]o handle the increased load, a distributed architecture can be used, dispersing the data and index structures across several computers and performing searches in parallel.").

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, in Tomasic_Caching_92, the authors studied performance trade-offs in shared-nothing distributed systems, including information retrieval systems, associated with different data placement and query routing techniques. (Tomasic_Caching_92, at 1).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. It was well known in the art to split queries into sub-parts, both in distributed and non-distributed systems. (*See*, *e.g.*, Tomasic_Caching_92, at 3; Stanfill_Massively_91, at 680; Salton_Automatic_89, at 192, 232, 233; Faloutsos_Access_85, at 53). Therefore, a person of ordinary skill in the art would find it obvious to fragment queries into multiple fragments, especially to take advantage of parallelism by using multiple nodes in the distributed system.

d. "hashing, by said selected home node, each said query fragment of said plurality of query fragments, said hashed query fragment having a first portion and a second portion" (cl.1); "hashes each said query fragment of said plurality of query fragments into a hashed query fragment having a first portion and a second portion" (cl.8); "hashing each said query fragment of said plurality of query fragments into a hashed query fragment having a first portion and a second portion" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 7-9
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 36

- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 132
- C. Severance, S. Pramanik, P. Wolberg, Distributed Linear Hashing and Parallel Projection in Main Memory Databases, Proceedings of the 16th VLDB Conference, Brisbane, Australia, pp. 674-682, 1990, at 675-676
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 248
- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 90-91
- Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993, at 105
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 4-6
- Witold Litwin, Marie-Anne Neimat, Donovan A. Schneider, LH* - Linear Hashing for Distributed Files, International Conference on Management of Data, ACM, pp. 327-336, 1993, at 328-333
- David DeWitt, A Single-User Performance Evaluation of the Teradata Database Machine, Proceedings of the Second International Workshop on High Performance Transaction Systems, pp. 244-269, Sept. 1987, at 247-248, 264-265
- Craig Stanfill, Partitioned Posting Files: A Parallel Inverted File Structure for Information Retrieval, Annual ACM Conference on Research and Development in Information Retrieval, Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 413-428, 1990, at 417
- Won Kim, Introduction to Object Oriented Databases, MIT Press, 1990, at 114
- Donald Knuth, The Art of Computer Science, Vol. 3, Addison-Wesley, 1973, at pp. 506-549
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 10-11
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9, 15, 21
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 21, 33, 50, Fig. 18
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 5

- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 72-73
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 53, Fig. 1
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 681
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 293-294
- Ijsbrand Jan Aalbersberg, Frans Sijstermans, High Quality and High Performance Full-Text Document Retrieval: The Parallel InfoGuide System, Proceedings of the First International Conference on Parallel and Distributed Information Systems, pp. 142-50, 1991, at 145
- Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992, at 116-118, 167-168
- Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds., at 20-21, 293-362
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989), at 112-115, 124-125, Figs. 1-3, Fig. 8
- Anastasia Analyti, Sakti Pramanik, Fast Search in Main Memory Databases, ACM Sigmod, pp. 215-224 (1992) at 215-218, Fig. 2
- The source code written on, or around, March 1993, identified in Plaintiffs' production, BATES-labeled JAR0255629, JAR0280890, JAR0280892-94, JAR0280896, and other related source code
- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994, at Abstract
- Houtsma et al., "Parallel Hierarchical Evaluation of Transitive Closure Queries," IEEE, 1991, at 130, 132-133
- Radek Vingralek, Yuri Breitbart, Gerhard Weikum, Distributed File Organization with Scalable Cost/Performance, SIGMOD Conference 1994, ACM, pp. 253-264, 1994, at 254-58
- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 1, 6, 7, 9, 19

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because a distributed database system, like any system, needs a way to structure data. Hashing is an old and well-known solution in computer science, and has been adapted for various uses. (*See*, *e.g.*, Knuth_73, Devine_DDH, Severance_90, Frakes_92). Thus, a person of ordinary skill in the art would have been motivated to combine a reference teaching a hashing method with references about distributed database systems.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, Devine_DDH states that "DDH, a distributed dynamic hashing algorithm, was implemented on a group of workstations to quantify the benefit of using a distributed solution," and that "DDH offers a useful approach for structuring distributed storage systems." (Devine_DDH, at 113-114). Also, Faloutsos_Access_85, discloses that "fast retrieval can be achieved if we invert key words. . . . The index file contains key words, sorted alphabetically." (Faloutsos_Access_85, at 53). Further, Faloutsos_Access_85 indicates "[m]ore sophisticated methods can be used to organize the index file, such as . . . hashing." (Faloutsos_Access_85, at 53).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. For example, developing every aspect of a software program is time consuming and expensive. It is common practice to use or adapt existing software components, or algorithms published in the field, to help develop a new program, rather than try to develop everything from the ground up. For example, Knuth_73, published twenty years before the '593 patent, disclosed a number of standard hashing functions and algorithms. (Knuth_73 at 506-549). Published algorithms are created so developers can implement software and/or systems using those algorithms. (*See*, *e.g.*, Devine_DDH, at 101 (stating that "[a]lgorithms should be devised to work in" the "future file and database systems . . . likely to be constructed as networked clusters of nodes."); JAR0255629, JAR0280890, JAR0280892-94, JAR0280896; Frakes_92, at 293-362). Therefore, a person of ordinary skill in the art would find it obvious to use prior art publications or existing software solutions and adapt them to the task at hand.

e. "transmitting, by said selected home node, each said hashed query fragment of said plurality of query fragments to a respective one of said plurality of query nodes indicated by said first portion of each said hashed query fragment" (cl.1); "transmits each said hashed query fragment to a respective one of said plurality of query nodes indicated by said first portion of said hashed query fragment" (cl.8); "transmitting a query message containing each said hashed query fragment to a respective one of said plurality of query nodes indicated by said first portion of said hashed query fragment" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 2-3, 9, 11
- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 130, 132-133
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 246, 248, 249, 260.
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of

Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 3, 5, 12, 13.

- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 9, 10, 11, 13, 14.
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 50, Fig. 18
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9
- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 91-92, 94-95, Fig. 4.3
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 36-38, Fig. 2.2
- Craig Stanfill, Brewster Kahle, Parallel Free-Text Search on the Connection Machine System, Communications of the ACM, December 1986, Vol. 29, No. 12, pp. 1229-1239, 1986
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 6, 10, 14
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 681
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989), at 113-115, 124-125
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 73-74
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 294
- Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993, at 105
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 72
- C. Severance, S. Pramanik, P. Wolberg, Distributed Linear Hashing and Parallel Projection in Main Memory Databases, Proceedings of the 16th VLDB Conference, Brisbane, Australia, pp. 674-682, 1990, at 675-676

- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 7, 9

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above in order to find more effective algorithms for transmitting query fragments. For example, having fragmented the query, a distributed database can take advantage of parallelism by transmitting the query fragments to various nodes. (*See* Tomasic_Caching_92, at 1) ("a distributed architecture can be used, dispersing the data and index structures across several computers and performing searches in parallel."). Thus, a person of ordinary skill in the art would have been motivated to transmit the query fragments to individual nodes where the data was dispersed.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, in Tomasic_Caching_92, the authors studied performance trade-offs in shared-nothing distributed systems, including information retrieval systems, associated with different data placement and query routing techniques. (Tomasic_Caching_92, at 1).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. It was well known in the art to transmit query fragments to separate nodes for processing a query in parallel. (Tomasic_Caching_92, at 3 ("[t]wo subqueries are issued. One . . . is sent to the host CPU 0, the other . . . is sent to host CPU

1. The subqueries are processed in parallel."); Devine_DDH, at 105 ("A client calls the DDH library routines to do the insert or retrieve operation A client program computes the hash key for the record, locates the likely bucket for that hash key, and then sends the request to the server that owns the bucket.")). Therefore, a person of ordinary skill in the art would find it obvious to transmit the query fragments to different query nodes. And because the fragments are hashed, it is obvious to transmit them according to the hash values.

f. "using, by said query node, said second portion of said respective hashed query fragment to access data according to a local hash table located on said query node" (cl.1); "each said query node uses said second portion of said hashed query fragment to access data according to a local hash table located on said query node" (cl.8); "upon receipt of said query message, using said second portion of said hashed query fragment to access data according to a local hash table located on said query node" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 90-91
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 5
- Robert Devine, Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm, Proceedings of the 4th International Conference, FODO '93, pp. 101-114, 1993, at 104-105
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 41, 50, Fig. 18
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in

Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 36

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 2-3, 7-9
- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 129, 130, 132-133
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 248, 260-261
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 2, 12-13
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 10-11, 13-14
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 53
- C. Severance, S. Pramanik, P. Wolberg, Distributed Linear Hashing and Parallel Projection in Main Memory Databases, Proceedings of the 16th VLDB Conference, Brisbane, Australia, pp. 674-682, 1990, at 675-676
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 285, 290-294
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 681
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 72-73
- Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992, at 116-118, 167-168
- Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds., at 20-21, 293-362
- The source code written on, or around, March 1993, identified in Plaintiffs' production, BATES-labeled JAR0255629, JAR0280890, JAR0280892-94, JAR0280896, and other related source code
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989), at 114-115, 124-125, Figs. 1-3, Fig. 8

- Anastasia Analyti, Sakti Pramanik, Fast Search in Main Memory Databases, ACM Sigmod, pp. 215-224 (1992) at 215-218, Fig. 2
- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994, at 197, 199
- Houtsma et al., "Parallel Hierarchical Evaluation of Transitive Closure Queries," IEEE, 1991, at 130
- Radek Vingralek, Yuri Breitbart, Gerhard Weikum, Distributed File Organization with Scalable Cost/Performance, SIGMOD Conference 1994, ACM, pp. 253-264, 1994, at 254-58
- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 1, 6, 7

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because a distributed database system, like any system, needs a way to structure data. Hashing is an old and well-known solution in computer science, and has been adapted for various uses. (*See*, *e.g.*, Knuth_73, Devine_DDH, Severance_90). Thus, a person of ordinary skill in the art would have been motivated to combine a reference teaching a hashing method with references about distributed database systems.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, Devine_DDH states that "DDH, a distributed dynamic hashing algorithm, was implemented on a group of workstations to quantify the benefit of using a distributed solution," and that "DDH offers a useful approach for structuring distributed storage systems." (Devine_DDH, at 113-114).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. For example, developing every aspect of a software program is time consuming and expensive. It is common practice to use or adapt existing software components, or algorithms published in the field, to help develop a new program, rather than try to develop everything from the ground up – this is precisely why special purpose software modular programs are developed. It was well known in the art to use hashing methods for distributing data, and software practicing those methods were available for purchase. Therefore, a person of ordinary skill in the art would find it obvious to use prior art publications or existing software solutions and adapt them to the task at hand.

g. "returning, by each said query node accessing data according to said respective hashed query fragment, an object identifier corresponding to said accessed data to said selected home node" (cl.1); "returns an object identifier corresponding to said accessed data to said home node" (cl.8); "transmitting a message returning an object identifier corresponding to said accessed data to said home node" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 2-3, 9
- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 130, 132-133
- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No.

3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 246, 249, 260-261

- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 3, 5, 12-13
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 9, 10-11, 13-14
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 73-74
- Craig Stanfill, Brewster Kahle, Parallel Free-Text Search on the Connection Machine System, Communications of the ACM, December 1986, Vol. 29, No. 12, pp. 1229-1239, 1986
- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 91, Fig. 1.2
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 6-7
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at Fig. 2.2
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9, 16, 21
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 33, Fig. 3, Fig. 18
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 293-294
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 680-681
- B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991, at 4
- Sakti Pramanik, Myoung-Ho Kim, Database Processing Models in Parallel Processing Systems, in Database Machines, pp. 112-126, Springer Berlin / Heidelberg (1989), at 112-115, 124-125
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 72

- David J. Dewitt, Shahram Ghandeharizadeh, Donovan Schneider, Allan Bricker, Hui-I Hsiao, Rick Rasmussen, The Gamma Database Machine Project, Computer Sciences Technical Report 921, pp. 1-34, March 1990, at 9

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because a distributed database system must eventually return information responsive to a query command; otherwise the database system is of no use. Thus, a person of ordinary skill in the art would have been motivated to return responsive information, especially in a format that is familiar to users of database systems.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, in Tomasic_Caching_92, the authors studied performance trade-offs in shared-nothing distributed systems, including information retrieval systems, associated with different data placement and query routing techniques. (Tomasic_Caching_92, at 1).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. A person of ordinary skill in the art would find it obvious to return the data requested by a query command, which is the whole point of a database system. It was well known in the art to return to a home node partial answers to a query. (*See*, *e.g.*, Tomasic_Caching_92, at 3). Thus, when query nodes accessed the requested data, it is obvious for those nodes to provide to the home node a means to retrieve the data. Because the database system is using object identifiers to identify data, it is obvious for the query nodes to return object identifiers.

h. "determining, by said home node, a measure of relevance between said accessed data and said query; and returning, to said user, by said home node, accessed data having a predetermined degree of relevance, subsequent to the step of returning said object identifier" (cl.3); "home node determines a measure of relevance between said accessed data and said query and returns to said user accessed data having a predetermined degree of relevance" (cl.9)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 74
- Gerald Salton, Automatic Text Processing, Addison-Wesley, 1988, at Chapter 10, pp. 238-239, 313-314
- Craig Stanfill, Brewster Kahle, Parallel Free-Text Search on the Connection Machine System, Communications of the ACM, December 1986, Vol. 29, No. 12, pp. 1229-1239, 1986
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 286-287, 291-294
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 680
- Christos Faloutsos, Access Methods for Text, ACM Computing Surveys, Vol. 17, No. 1, pp. 49-74, March 1985, at 51
- W. W. Chu, M. A. Merzbacher, L. Berkovich, The Design and Implementation of CoBase, SIGMOD, pp. 517-522, May 1993, at 517, 520-521
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 36-37

- B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991, at 4
- Charles T. Meadow, Text Information Retrieval Systems, Academic Press, 1992, at 201-204, 274-280
- Information Retrieval, Data Structures and Algorithms, Prentice Hall, 1992, William B. Frakes, Ricardo Baeza-Yates, eds., at 363-392
- Chaturvedi, et al., "Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach," IEEE Transactions on Engineering Management, Vol. 41, No. 2, May 1994, at 199-200
- Houtsma et al., "Parallel Hierarchical Evaluation of Transitive Closure Queries," IEEE, 1991, at 130
- Gerard Salton, James Allan, Chris Buckley, Automatic Structuring and Retrieval of Large Text Files, Communications of the ACM, February 1994, Vol. 37, No. 2, pp. 97-108, 1994, at 97-99

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because reducing the amount of irrelevant answers to a query is desirable, and thus limiting the answers to a query to data that meets a certain threshold of relevance is of importance.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, in CoBase_93, a distributed database is discussed "that integrates knowledge base technology with database systems to provide cooperative (approximate and conceptual) query answering." (CoBase_93, at 517). CoBase works in combination with other distributed databases. (*Id.*, at 521 ("“CoBase uses LOOM as knowledge representation and inference system and supports relational data bases (e.g. Oracle and Sybase).”)). As a further example, the textbook on Automatic Text Processing by Gerald Salton teaches using vectors for information retrieval systems to calculate measures of relevance. (Salton_Automatic_89, at 313-314).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. It was well known in the art to calculate levels of relevance for information retrieval systems. (*See* Salton_Automatic_89, at 313-314). Therefore, a person of ordinary skill in the art would find it obvious to limit responses to a query to data that is at least somewhat relevant (*i.e.*, to data meeting a predetermined degree of relevance) as opposed to returning information that has little or no relevance to the query.

i. "each said home node, upon receiving a command from a user, enqueueing a predetermined task in response to said command" (cl.13)

To the extent Plaintiff contends that any anticipatory prior art reference identified above does not disclose these limitations, the limitations are obvious in light of the background knowledge possessed by a person of ordinary skill in the art, and/or the limitations can be found, for example, in one or more of the following references:

- William Alexander, George Copeland, Process and Dataflow Control in Distributed Data-Intensive Systems, pp 90-98, June 1988, at 92-93
- Haran Boral, William Alexander, Larry Clar, George Copeland, Scott Danforth, Michael Franklin, Brian Hart, Marc Smith, Patrick Valduriez, Prototyping Bubba, A Highly Parallel Database System, IEEE Transactions on Knowledge and Data Engineering, Vol. 2, No. 1, 4-24, March 1990, at 5-6
- Anthony Tomasic and Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, ACM SIGMOD, pp. 129-138, 1993, at 130, 132-133
- Anthony Tomasic, Hector Garcia-Molina, Caching and Database Scaling in Distributed Shared-Nothing Information Retrieval Systems, Stanford University Computer Science Technical Report STAN-CS-92-1456, December 22, 1992, at 3, 7-9
- Itasca Systems, Douglas Barry, Itasca Systems Technical Report Number: TM-92-001, ODBMS Feature Checklist, Rev. 1.0, March 1992, at 9, 16
- Marc Smith, Bill Alexander, Haran Boral, George Copeland, Tom Keller, Herb Schwetman, Chii-Ren Young, An Experiment on Response Time Scalability in

Bubba, IWDM '89: Proceedings of the Sixth International Workshop on Database Machines, G. Goos, J. Hartmanis, eds., June 1989, at 37, Fig. 2.2, Fig. 4.3

- Anthony Tomasic and Hector Garcia-Molina, Query Processing and Inverted Indices in Distributed Text Document Retrieval Systems, VLDB Journal, 2, No. 3, pp. 243-275, 1993, Michael Carey and Patrick Valuriez, eds., at 246, 260
- Itasca Systems, ITASCA: Distributed Object Database Management System, Technical Summary for Release 2.1, 1992, at 15, 21, 33, 50, Fig. 18
- B. Kahle and A. Medlar, An Information System for Corporate Users: Wide Area Information Servers, ConneXions – The Interoperability Report, Vol. 5, No. 11, pp. 2-9, 1991, at 3
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Distributed Text Document Retrieval Systems, Stanford University Department of Computer Science Technical Report Number STAN-CS-92-1434, June 23, 1992, at 5, 12-13
- Anthony Tomasic and Hector Garcia-Molina, Performance of Inverted Indices in Shared-Nothing Distributed Text Document Information Retrieval Systems, Proceedings of the second international conference on Parallel and distributed information systems, IEEE, pp. 8-17, 1993, at 10-11, 13-14
- Forbes J. Burkowski, Retrieval Performance of a Distributed Text Database Utilizing a Parallel Processor Document Server, International Symposium on Databases for Parallel and Distributed Systems, Proceedings of the second international symposium on Databases in parallel and distributed systems, pp. 71-79, IEEE, 1990, at 72, 74
- Ijsbrand Jan Aalbersberg, Frans Sijstermans, High Quality and High Performance Full-Text Document Retrieval: The Parallel InfoGuide System, Proceedings of the First International Conference on Parallel and Distributed Information Systems, pp. 142-50, 1991, at 144-145, 147
- Craig Stanfill, Massively Parallel Information Retrieval for Wide Area Information Servers, pp. 679-682, 1991, at 681
- Craig Stanfill, Robert Thau, Information Retrieval on the Connection Machine: 1 to 8192 Gigabytes, Information Processing & Management, Vol. 27, No. 4, pp. 285-310, 1991, at 293-294

One of ordinary skill in the art would have been motivated to combine these references for at least the following reasons. First, a person of ordinary skill in the art would have been motivated to combine any of the anticipatory references with one or more of the particular references listed above because a node that receives a user command will need to respond in some manner, otherwise the user will receive a bad experience. (*See*, *e.g.*, Burkowski, at 74 ("For a given query originating from Workstation(j), the functionality of the aforementioned steps can be established within various processes or tasks that communicate using a message passing facility.”)). Thus, a person of ordinary skill in the art would have been motivated to enqueue some task to run in response to receiving a user command.

Second, the references themselves suggest their use in a variety of applications and combinations, thus further supporting a finding of obviousness. For example, in Tomasic_Caching_92, the authors studied performance trade-offs in shared-nothing distributed systems, including information retrieval systems, associated with different data placement and query routing techniques. (Tomasic_Caching_92, at 1).

Third, the nature of the problem to be solved would have directed persons of ordinary skill in the art to consider the combination of these references to arrive at the pertinent subject matter. It was well known in the art to respond to a user command, and enqueue a task to execute. (*See*, *e.g.*, Prototyping_Bubba, at 6 (“The transaction execution steps are as follows. 1. The execution begins by creating a transaction coordinator (TC) to coordinate the transaction, which resides in the IP that received the execution request from the user. The TC . . . sends a message to IR0 to begin the dataflow execution.”); Tomasic_Caching_93, at 130 (“To answer a query in the system index organization, a subquery is sent to each host relevant to the query.”)). Thus, a person of ordinary skill in the art would find it obvious to enqueue a predetermined task for a home node in response to receiving a user command.

C. Indefiniteness, Enablement, Written Description

Pursuant to P.R. 3-3(d), Google lists below the grounds upon which the asserted claims are invalid based on indefiniteness under 35 U.S.C. § 112(2) or based on failure to meet the enablement or written description requirements under 35 U.S.C. § 112(1). This is particularly so in view of the breadth of the hashing and fragmenting constructions offered by Plaintiffs in their two conflicting infringement theories against IndexServer and TeraGoogle, the testimony of Kenneth Baclawski regarding practicing these limitations, and recent Federal Circuit precedent on written description and enablement. *Ariad Pharmaceuticals, Inc. et al. v. Eli Lily and Company*, ___ F.3d ___, (Fed. Cir. March 22, 2010) (en banc). As Google best understands Plaintiffs' contentions at this time, the asserted claims fail to meet these requirements for at least the following reasons.

Claim Limitation	Claims With the Limitation	Basis for invalidity
"non-relational"	1-3, 8-9, 13	Under Plaintiff's apparent constructions, claims fail to satisfy the requirements of 35 U.S.C. § 112(2) because the limitation is indefinite. Under Plaintiff's apparent constructions, claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation, and fails to provide an adequate written description of the limitation.
"randomly selecting"	1-3	Claims fail to satisfy the requirements of 35 U.S.C. § 112(2) because the limitation is indefinite. Claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation, and fails to provide an adequate written description of the limitation.
"fragmenting"	1-3, 8-9, 13	Claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation, and fails to provide an adequate written description of the limitation.
"hashing"	1-3, 8-9, 13	Claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation, and fails to provide an adequate written description of the limitation.

"fuzzy query"	1-3	Claims fail to satisfy the requirements of 35 U.S.C. § 112(2) because the limitation is indefinite. Claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation, and fails to provide an adequate written description of the limitation.
"access data according to a local hash table"	1-3, 8-9, 13	Claims fail to satisfy the requirements of 35 U.S.C. § 112(1) because the specification fails to provide an enabling disclosure of the limitation.

In addition to the reasons set forth above, claims 8, 9, and 13 are invalid under 35 U.S.C. § 112(2) in view of *IPXL Holdings LLC v. Amazon.com, Inc.*, 430 F.3d 1377, 1383-84 (Fed. Cir. 2005).

Dated: May 17, 2010

Respectfully submitted,

FISH & RICHARDSON P.C.

By: /s/ Frank Albert

Michael E. Jones (SBN 10929400)
mikejones@potterminton.com
POTTER MINTON
A Professional Corporation
110 N. College, Suite 500
Tyler, TX 75702
Telephone: (903) 597-8311
Facsimile: (903) 593-0846

Ruffin B. Cordell (SBN 04820550)
cordell@fr.com
FISH & RICHARDSON P.C.
1425 K Street, N.W., 11th Floor
Washington, DC 20005-3500
Telephone: (202) 783-5070
Facsimile: (202) 783-2331

Jason W. Wolff (CA SBN 215819)
wolf@fr.com
Frank Albert (*Admitted Pro Hac Vice*)
albert@fr.com
FISH & RICHARDSON P.C.
12390 El Camino Real
San Diego, CA 92130
Telephone: (858) 678-5070
Facsimile: (858) 678-5099

Howard G. Pollack (*Admitted Pro Hac Vice*)
pollack@fr.com
Shelley K. Mack (*Admitted Pro Hac Vice*)
mack@fr.com
Jerry T. Yen (CA SBN 247988)
yen@fr.com
Enrique Duarte (*Admitted Pro Hac Vice*)
Duarte@fr.com
FISH & RICHARDSON P.C.
500 Arguello Street, Suite 500
Redwood City, CA 94063
Telephone: (650) 839-5070
Facsimile: (650) 839-5071

Attorneys for Defendant
GOOGLE INC.

CERTIFICATE OF SERVICE

The undersigned hereby certifies that on May 18, 2010, a true and correct copy of the above and foregoing document was served on all counsel of record who are deemed to have consented to electronic service via the Court's CM/ECF system pursuant to Local Rule CV-5(a)(3). Any other counsel of record not deemed to have consented to electronic service were served by email at northeastern@velaw.com.

Stephen Charles Stout Christopher Ryan R. Floyd Walker David B. Weaver Michael Anthony Valek Vinson & Elkins - Austin 2801 Via Fortuna, Suite 100 Austin, TX 78746 sstout@velaw.com; dweaver@velaw.com; fwalker@velaw.com; mvalek@velaw.com;	Attorneys for Plaintiffs NORTHESTERN UNIVERSITY and JARG CORPORATION
William B. Dawson Vinson & Elkins - Dallas 3700 Trammell Crow Center 2001 Ross Avenue Dallas, TX 75201-2975 bdawson@velaw.com	Attorneys for Plaintiffs NORTHESTERN UNIVERSITY and JARG CORPORATION
Otis W. Carroll, Jr. Collin M. Maloney, Esq. Ireland, Carroll & Kelley, PC 6101 South Broadway, Suite 500 [75703] PO Box 7879 Tyler, TX 75711-7879 fedserv@icklaw.com;	Attorneys for Plaintiffs NORTHESTERN UNIVERSITY and JARG CORPORATION
Franklin Jones, Jr. Jones & Jones, Inc., P.C. 201 West Houston Street PO Drawer 1249 Marshall, TX 75671-1249 maizieh@millerfirm.com	Attorneys for Plaintiffs NORTHESTERN UNIVERSITY and JARG CORPORATION
Allen Franklin Gardner Michael E. Jones Potter Minton PC 110 N. College, Suite 500 PO Box 359 Tyler, TX 75710-0359 allengardner@potterminton.com mikejones@potterminton.com	Attorneys for Defendant GOOGLE, INC.

/s/Frank Alber
Frank Albert