

**UNITED STATES DISTRICT COURT  
EASTERN DISTRICT OF TENNESSEE**

**UNIVERSITY OF TENNESSEE RESEARCH  
FOUNDATION AND SAINT MATTHEW  
RESEARCH, LLC,**

*Plaintiffs,*

v.

**CLOUDERA, INC.**

*Defendant.*

**Civil Action No.** \_\_\_\_\_

**JURY TRIAL DEMANDED**

**COMPLAINT FOR PATENT INFRINGEMENT**

The University of Tennessee Research Foundation and Saint Matthew Research, LLC (collectively, “Plaintiffs”), by its undersigned counsel, bring this action and make the following allegations of patent infringement relating to U.S. Patent Nos.: 7,454,411 (the “411 patent”) and 8,099,733 (the “733 patent”) (collectively, the “patents-in-suit”). Defendant Cloudera, Inc. (“Cloudera” or “Defendant”) infringes each of the patents-in-suit in violation of the patent laws of the United States of America, 35 U.S.C. § 1 *et seq.*

**INTRODUCTION**

1. This is an action for infringement of patents awarded to computer scientists for their work at The University of Tennessee. The patents are owned by the University of Tennessee Research Foundation, a non-profit 501(c)(3) organization that promotes, licenses, and commercializes The University of Tennessee’s intellectual property.

2. Originally created in 1935 as the University of Tennessee Research Corporation, the University of Tennessee Research Foundation helps The University of Tennessee fulfill its mission in becoming a national leader in research, discovery, and innovation. The University of Tennessee Research Foundation “was established to protect, manage, and commercialize

university inventions and intellectual property; grow the university research enterprise; develop and support an entrepreneurial culture; and contribute to state and regional economic development.”<sup>1</sup>

3. This case arises from Cloudera’s infringement of a portfolio of patents protecting the groundbreaking work of several of The University of Tennessee’s faculty in the fields of parallel processing and high performance database design. The patents disclose systems and methods for parallel processing computer architectures that evenly distribute computational loads over multiple nodes. The patents also teach using the structural properties of data in large datasets to improve database performance. The inventions taught in the patents employ parallel processing schemes, which drastically improve the storage and retrieval capabilities of databases.

4. Established in 1794, The University of Tennessee is Tennessee’s flagship public research institution based in Knoxville, Tennessee. The University of Tennessee has more than 28,000 students, 1700 faculty members, and over 300 degree programs. In 2016, The University of Tennessee system had \$435 million in research and sponsored program expenditures.<sup>2</sup> To maximize the public benefit that its research generates, in certain instances, technological innovations developed by The University of Tennessee faculty, are assigned to the University of Tennessee Research Foundation, which patents and commercializes these innovations on behalf of The University of Tennessee.

5. In the late 1990’s Professor J. Douglas Birdwell and his colleagues at The University of Tennessee developed technologies for searching large volumes of DNA profile

---

<sup>1</sup> *State of Tennessee Comptroller of the Treasury, THE UNIVERSITY OF TENNESSEE FINANCIAL AND COMPLIANCE AUDIT REPORT* at 19 and 57 (January 20, 2015).

<sup>2</sup> Joe DiPietro, *THE STATE OF THE UNIVERSITY OF TENNESSEE ADDRESS* (February 21, 2017), *available at*: <http://tennessee.edu/state-of-ut/>.

data. Initially, Professor Birdwell and his fellow researchers sought to design database systems that could handle the vast and multi-dimensional nature of datasets associated with storing DNA profiles. DNA profile information presented unique challenges because of the need to search through tens of millions of records using complex match specifications. At the time, standard database structures failed to provide “rapid access to records” and were unable to “take advantage of naturally occurring structure in the data.”<sup>3</sup>

6. To handle large datasets of DNA profile information, The University of Tennessee researchers developed technologies for efficiently distributing computationally intensive query functionality between computer hosts, grouping data records into distinct clusters in a database, and coordinating communication between groups of computer processors. Professor Birdwell and his colleagues developed technologies that used the structural properties of DNA data to improve the retrieval of data records. In addition, The University of Tennessee researchers developed parallel processing computer architectures for balancing computational requests across nodes. “As Search Queues become unbalanced, unprocessed Search Requests are exchanged to bring them back into balance. This exchange occurs randomly with a stochastic selection method utilized to determine the recipient of each exchange.”<sup>4</sup>

7. While Dr. Birdwell and his collaborators initially sought to address the need for database systems capable of handling the size and complex nature of DNA profile information, their work led to groundbreaking innovations applicable to database systems that handle a wide variety of complex and large datasets. The technologies developed for handling large volumes of DNA profile data were identified by The University of Tennessee researchers as providing

---

<sup>3</sup> ‘983 patent, Col. 1:48-51.

<sup>4</sup> *Id.* at Col. 28:2-6.

groundbreaking insights applicable far beyond forensic applications.

8. Highlighting the importance of The University of Tennessee faculty's contributions to the field is the fact that the patents-in-suit have been cited by over 300 U.S. Patents and Patent Applications by a wide variety of the largest companies operating in the field. For example, the patents-in-suit have been cited by companies such as:

- International Business Machines Corporation (IBM)<sup>5</sup>
- Hewlett-Packard Enterprise Company (HPE)<sup>6</sup>
- Google, Inc.<sup>7</sup>
- Oracle Corporation<sup>8</sup>
- SAP SE<sup>9</sup>
- Microsoft Corporation<sup>10</sup>
- EMC Corporation<sup>11</sup>
- Qualcomm, Inc.<sup>12</sup>
- Koninklijke Philips N.V. (Philips)<sup>13</sup>
- Western Digital Corporation<sup>14</sup>
- Teradata Corp.<sup>15</sup>
- Xerox Corporation<sup>16</sup>

#### **THE UNIVERSITY OF TENNESSEE'S LANDMARK PARALLEL PROCESSING TECHNOLOGIES**

9. Through their work at The University of Tennessee, Professor Birdwell and his colleagues developed groundbreaking inventions in the fields of parallel processing and high

---

<sup>5</sup> See, e.g., U.S. Patent Nos. 6,915,289; 7,269,786; 7,684,963; 7,685,584.

<sup>6</sup> See, e.g., U.S. Patent Nos. 8,621,480; 8,185,893; 8,341,626.

<sup>7</sup> See, e.g., U.S. Patent Nos. 8,326,861; 8,316,019; 8,392,396; 8,73,238.

<sup>8</sup> See, e.g., U.S. Patent Nos. 7,174,344; 7,779,008; 8,229,909; 9,436,514.

<sup>9</sup> See, e.g., U.S. Patent Nos. 8,806,016; 8,103,772; 8,499,078; 9,444,732.

<sup>10</sup> See, e.g., U.S. Patent Nos. 7,921,424; 7,502,807; 9,430,508.

<sup>11</sup> See, e.g., U.S. Patent Nos. 9,171,042; 9,275,117; 9,454,573.

<sup>12</sup> See, e.g., U.S. Patent No. 9,110,726.

<sup>13</sup> See, e.g., U.S. Patent App. No. 14/416,647.

<sup>14</sup> See, e.g., U.S. Patent Nos. 8,868,487; 8,874,515; 8,954,385.

<sup>15</sup> See, e.g., U.S. Patent Nos. 7,359,913; 8,938,444.

<sup>16</sup> See, e.g., U.S. Patent No. 7,720,848.

performance database design

10. Throughout the 2000's, these developments, which revolutionized the scalability of modern-day database systems, were widely adopted in the market. Dr. Birdwell recognized in a subsequent academic paper that the groundbreaking patents came out of a need to develop technologies for handling large DNA data sets.

***The projected growth of the NDIS [National DNA Index System] database and the demand for searches of the database necessitates migration to a parallel computing platform.*** Effective utilization of a parallel computer architecture requires the computational load to be distributed more or less evenly over the available CEs. The qualifier “more or less” is used because the communications required to distribute the load consume both computational resources and network bandwidth. A point of diminishing returns exists. The distribution of computational load across available resources is referred to as the load balancing problem in the literature.

Douglas J. Birdwell et al, *Dynamic Time Delay Models for Load Balancing Part I: Deterministic Models*, in *ADVANCES IN TIME-DELAY SYSTEMS*, pp. 355-370. Springer Berlin Heidelberg (2004) (emphasis added).

11. In the years following the development of the patents-in-suit, the volume of data has grown exponentially, requiring new technologies such as those disclosed in the patents-in-suit. Many of the companies that cite the patents-in-suit have also acknowledged that the growing volume of data led to unique challenges that required new computer technologies.

With the surging volume of data being sourced from an ever-growing variety of data sources and applications, many streaming with great velocity, organizations are unable to use traditional data integration mechanisms such as ETL (extraction, transformation, and load). Big Data requires new strategies and technologies designed to analyze big data sets at terabyte or even petabyte scale.

*An Enterprise Architect's Guide to Big Data – Reference Architecture Overview*, ORACLE ENTERPRISE ARCHITECTURE WHITE PAPER at 26 (March 2016).

12. Each of the patents-in-suit claim priority to 1999, when typical datasets stored in databases were orders of magnitude smaller than today. Dr. Birdwell and his colleagues invented technologies that today are incorporated into the products of many leading technology companies.



Inventors.<sup>20</sup> Dr. Birdwell received his B.S. and M.S. degrees in Electrical Engineering from The University of Tennessee, and in 1978 he received his Ph.D. in Electrical Engineering from the Massachusetts Institute of Technology (“MIT”). Upon receiving his doctorate, Dr. Birdwell joined The University of Tennessee faculty in 1978.

16. Professor Tsewei Wang is a Professor Emeritus of Chemical and Biomolecular Engineering at The University of Tennessee, Knoxville. Dr. Wang is the author of numerous articles on load balancing, data mining, and clustering data records. Dr. Wang received her Master’s degree in Chemical Engineering from The University of Tennessee and her Ph.D. in Biophysics from MIT. Dr. Wang’s research focuses on the fields of data mining, process monitoring, and bioinformatics.

17. Professor Roger Horn is a Research Professor of Electrical Engineering and Computer Science at The University of Tennessee, Knoxville. Dr. Horn’s research and industrial experience range from digital circuit design, data acquisition, control systems, and the development of application and driver software to turbine control systems, radiation monitoring systems, smart pressure transducers, and medical devices. Dr. Horn is a Registered Professional Engineer in the State of Tennessee, and a member of the National Society of Professional Engineers, the Institute of Electrical and Electronics Engineers, the IEEE Control Systems Society, and the IEEE Robotics and Automation Society. Dr. Horn received his Bachelor’s and Master’s degrees in Electrical Engineering from the University of California, Berkeley in 1978 and 1984, respectively. Dr. Horn received his Ph.D. in Electrical Engineering from The University of Tennessee in 1992.

---

<sup>20</sup> *Third UT Faculty Member Named to National Academy of Inventors*, TENNESSEE TODAY (Dec. 17, 2015), available at, <http://tntoday.utk.edu/2015/12/17/birdwell-ut-named-national-academy-inventors/>.

18. Professor David J. Icove is a UL Professor of Electrical Engineering and Computer Science at The University of Tennessee, Knoxville. Dr. Icove is the author of numerous articles and textbooks on forensic analysis, computational modeling, and pattern recognition. Dr. Icove received his Master's in Electrical Engineering from The University of Tennessee in 1973, and his Ph.D. in Engineering Science and Mechanics from The University of Tennessee in 1979. Dr. Icove is the inaugural Underwriters Laboratories Professor of Practice at The University of Tennessee.<sup>21</sup>

19. Puneet Yadav is a Director of Product Marketing at Lam Research in Fremont, California. Mr. Yadav received his Bachelor's degree in Chemical Engineering from the Indian Institute of Technology in 1998 and a Master's degree in Chemical Engineering from The University of Tennessee, Knoxville in 2001. At the time the inventions disclosed in the patents-in-suit were developed, Mr. Yadav was a graduate student at The University of Tennessee and collaborated with the other inventors of the patents-in-suit.<sup>22</sup>

#### **THE PARTIES**

##### **UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION**

20. The University of Tennessee Research Foundation is the second-oldest university research foundation in the United States. The University of Tennessee Research Foundation's mission is to promote, support, and carry out the research mission of The University of Tennessee, to enhance the competitive position of The University of Tennessee for research and development funding, facilitate expanded research and development activities at The University

---

<sup>21</sup> *Underwriters Laboratories Supporting College of Engineering Course*, TENNESSEE TODAY (April 7, 2014), available at, <http://tntoday.utk.edu/2014/04/07/underwriters-laboratories-supports-ut-engineering-course/>.

<sup>22</sup> Puneet Yadav, *Increasing the Speed and Efficiency of Search in FBI/CODIS DNA Database Through Multivariate Statistical Clustering Approach and Development of a Similarity Ranking Scheme*, MASTER'S THESIS UNIVERSITY OF TENNESSEE (2001).

of Tennessee, and to facilitate the commercialization of University of Tennessee research outcomes and the transfer of research-generated technology from The University of Tennessee to commercial and industrial enterprises in furtherance of the economic development of the State of Tennessee.

21. UTRF is responsible for licensing patents covering University of Tennessee inventions and collecting royalties on behalf of The University of Tennessee from those license agreements. The State of Tennessee Comptroller has described the role of UTRF as:

The foundation's stated purpose is, in conjunction with the university, to grow the University of Tennessee research enterprise; harvest, manage, and market University of Tennessee intellectual property; encourage and support entrepreneurial education and ventures by faculty, staff, students, and commercial partners/affiliates of the University of Tennessee; and to contribute to the well-being of the State of Tennessee through economic development.

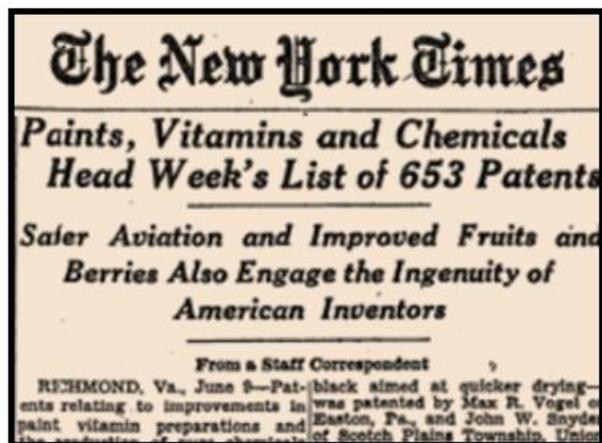
*The University of Tennessee Audit Report for The Year Ending June 30, 2014*, STATE OF TENNESSEE COMPTROLLER OF THE TREASURY at 57 (January 20, 2015).

22. Since its inception as the University of Tennessee Research Corporation in 1935, UTRF has been actively involved in licensing patents whose technologies were first developed at The University of Tennessee. The University of Tennessee Research Corporation's 1935 charter stated that its mission was "to promote, encourage and aid scientific social and/or educational investigation and research."<sup>23</sup> In furtherance of these objectives, the corporation was empowered "[t]o aid in the prosecution of applications for patents, registrations and/or copyrights, foreign and domestic . . . . To prosecute infringements or invasions of any patent, trade-mark, trade name, brand, label, copyright or patent right in which the corporation may be interested."<sup>24</sup>

---

<sup>23</sup> CHARTER OF INCORPORATION OF THE UNIVERSITY OF TENNESSEE RESEARCH CORPORATION (December 31, 1934).

<sup>24</sup> *Id.*



*Safer Aviation and Improved Fruits and Berries Also Engage the Ingenuity of American Inventors*, N.Y. TIMES at 20 (June 10, 1944) (Describing a patent “issued to Brooks D. Drain of Knoxville, Tenn., on a ‘healthy, prolific strawberry plant variety’” and “assigned to the University of Tennessee Research Corporation.”).

23. In its first decades of existence the University of Tennessee Research Foundation licensed its patents to cottonseed, fruit varieties, and water treatment.

Of late, research on a small commercial scale is being carried on by the University of Tennessee Experiment Station and a *University of Tennessee Research Corporation* has been set up to license and control the use of the process. Here then we have a state institution of learning working with a federal government agency, a professional society, a trade association, and private capital all working together for a common purpose now by the upbuilding of regional agricultural economy.”<sup>25</sup>

The largest manufacturer of cottonseed manufacturing equipment in the United States has applied for a license under royalty to be paid to the University of Tennessee Research Corporation.<sup>26</sup>

The University of Tennessee Research Corporation licenses manufacturers to make equipment resulting from the research work for which a patent has been obtained or applied for.”<sup>27</sup>

24. More recently, UTRF has licensed patents pertaining to various fields of

---

<sup>25</sup> THE UNIVERSITY OF TENNESSEE RECORD, Vol. 44, Issue 4 (1941).

<sup>26</sup> UNITED STATES CONGRESS JOINT COMMITTEE HEARING ON THE TENNESSEE VALLEY AUTHORITY at 1451 (1939).

<sup>27</sup> *Statement of David E. Lilienthal, Chairman, Tennessee Valley Authority*, TECHNOLOGICAL MOBILIZATION, HEARINGS BEFORE A SUBCOMMITTEE OF THE COMMITTEE ON MILITARY AFFAIRS OF THE UNITED STATES SENATE S. 2721 Vol. 3 at 910 (December 1942).

technology, including important, lifesaving technologies. UTRF has recently licensed patents for a vaccine for group A streptococcus (Strep-A) and for treatments for prostate cancer.<sup>28</sup>

**SAINT MATTHEW RESEARCH, LLC**

25. Saint Matthew Research, LLC (“SMR”) is a California limited liability company with its principal place of business at 11400 W. Olympic Blvd., Suite 200, Los Angeles, California 90064. Pursuant to a license agreement with UTRF, SMR is the exclusive licensee of the patents-in-suit. SMR facilitates the University of Tennessee Research Foundation’s mission of commercializing UT’s technologies by assisting UTRF’s commercialization efforts.

**CLOUDERA**

26. On information and belief, Cloudera.com, Inc. is a Delaware corporation with a principal place of business at 1001 Page Mill Road, Building 2, Palo Alto, CA 94304. On information and belief, Cloudera is registered to do business in the State of Tennessee and may be served with process by delivering a summons and a true and correct copy of this complaint to its registered agent for receipt of service of process, Corporation Service Company, 2908 Poston Avenue, Nashville, Tennessee 37203.

27. On information and belief, Cloudera regularly conducts business throughout the State of Tennessee, including hosting training sessions on the use products accused of infringement in this case.<sup>29</sup> On information and belief, Cloudera further partners with

---

<sup>28</sup> James B. Dale et al., *Potential Coverage of a Multivalent M Protein-Based Group a Streptococcal Vaccine*, VACCINE 31.12 (2013): 1576–1581 (“The University of Tennessee Research Corporation has licensed the technology to Vaxent, LLC.”); Tom Wilemon, *Powerful Possibilities: GTx Scientists Battle Time, Regulatory Setbacks to Make it Big*, MEMPHIS DAILY NEWS (May 17, 2010) (“The SARM patents, as well as Steiner’s research into tormifene for the prevention of prostate cancer, were licensed to GTx by the University of Tennessee.”).

<sup>29</sup> See, e.g., *Nashville Cloudera Sessions*, CLOUDERA WEBSITE, available at: <https://www.cloudera.com/more/news-and-events/events/sessions/nashville.html> (last visited April 2017).

Tennessee-based businesses to deliver technology solutions incorporating Cloudera's products. For example, Cloudera has partnered with Tennessee-based Digital Reasoning.<sup>30</sup>

#### JURISDICTION AND VENUE

28. This action arises under the patent laws of the United States, Title 35 of the United States Code. Accordingly, this Court has exclusive subject matter jurisdiction over this action under 28 U.S.C. §§ 1331 and 1338(a).

29. Upon information and belief, this Court has personal jurisdiction over Cloudera in this action because Cloudera has committed acts within the Eastern District of Tennessee giving rise to this action and has established minimum contacts with this forum such that the exercise of jurisdiction over Cloudera would not offend traditional notions of fair play and substantial justice. Defendant Cloudera, directly and/or through subsidiaries or intermediaries (including distributors, retailers, and others), has committed and continues to commit acts of infringement in this District by, among other things, offering to sell and selling products and/or services that infringe the patents-in-suit. Moreover, Cloudera is registered to do business in the State of Tennessee, and actively directs its activities to customers located in the State of Tennessee.

30. Venue is proper in this district under 28 U.S.C. §§ 1391(b)-(d) and 1400(b). Defendant Cloudera is registered to do business in the State of Tennessee and, upon information and belief, has transacted business in the Eastern District of Tennessee and has committed acts of direct and indirect infringement in the Eastern District of Tennessee.

---

<sup>30</sup> *Featured Partner Solutions*, CLOUDERA WEBSITE, available at: <https://www.cloudera.com/partners/solutions.html> (last visited April 2017).

## TECHNOLOGY BACKGROUND

### U.S. PATENT NO. 7,454,411

31. U.S. Patent No. 7,454,411 (“the ‘411 patent”) entitled, *Parallel Data Processing Architecture*, was filed on January 30, 2004, and claims priority to September 28, 1999. The ‘411 patent is subject to a 35 U.S.C. § 154(b) term extension of 385 days. UTRF is the owner by assignment of the ‘411 patent. SMR is the exclusive licensee of the ‘411 patent. A true and correct copy of the ‘411 patent is attached hereto as Exhibit A.

32. The ‘411 patent teaches a novel parallel data processing system for search, storage, and retrieval of data from a database.

33. The ‘411 patent and its underlying patent application have been cited by 83 United States patents and patent applications as relevant prior art. Specifically, patents issued to the following companies have cited the ‘411 patent and its underlying patent application as relevant prior art:

- International Business Machines Corporation
- EMC Corporation
- AT&T, Inc.
- Koutin Technology Inc.
- Hewlett-Packard Enterprise Company
- Google, Inc.
- Infosys Ltd.
- Oracle Corporation
- Qualcomm, Inc.
- Raytheon Company
- Sandbridge Technologies, Inc.

### U.S. PATENT NO. 8,099,733

34. U.S. Patent No. 8,099,733 (“the ‘733 patent”) entitled, *Parallel Data Processing Architecture*, was filed on January 2, 2008, and claims priority to September 28, 1999. The ‘733 patent is subject to a 35 U.S.C. § 154(b) term extension of 1,051 days. UTRF is the owner by

assignment of the '733 patent. SMR is the exclusive licensee of the '733 patent. A true and correct copy of the '733 patent is attached hereto as Exhibit B.

35. The '733 patent teaches a parallel data processing architecture containing two or more host processors. One of the host processors responds to client queries for database records. The host processors are further able to communicate with one another via a communication systems. The parallel data processing architecture balances the workload between the host processors.

36. The '733 patent teaches the balancing of a workload between host processors through the host processors broadcasting load information (e.g., processor capacity and the length of the search queue). The load information is broadcasted by host processors to at least one or more host processors in the system.

37. The parallel data processing architecture disclosed in the '733 patent is capable of balancing the workload between two or more host processors by having the host processors bring their associated search queue of client requests into balance with at least one other host processor. The workload balancing by the host processors use a time constant responsive to receipt of broadcasted capacity and load information.

38. The '733 patent describes balancing workload between host processors as being enabled through the exchange of block search requests between two or more of the host processors. This exchange of search requests between host processors minimizes the time required to respond to client queries by adjusting the block of search requests that are exchanged based on the relative processing speeds of host processors and the communications protocol between host processors.

39. The '733 patent balances workload between computer hosts to equalize the average wait time for computation of requests. For example, blocks of search requests can be exchanged among hosts from hosts with relatively long average waiting times to hosts with shorter waiting times. The sizes of the blocks of exchanged requests can be adjusted to accommodate the relative speeds of the processors and the inter-processor communications protocols and hardware.

40. Another insight for improving the performance of requests on a database system described by the '733 patent is to have each host broadcast measurements of their current load (queue lengths) and measurements of their capacity.

41. The '733 patent and its underlying patent application has been cited by 14 United States patents and patent applications as relevant prior art. Specifically, patents issued to the following companies have cited the '733 patent and its underlying patent application as relevant prior art:

- Oxford University Innovation Ltd.
- San Diego State University Foundation
- Path, Inc.
- Four J's Development Tools, Inc.
- Hewlett-Packard Enterprise Company

**COUNT I**  
**INFRINGEMENT OF U.S. PATENT NO. 7,454,411**

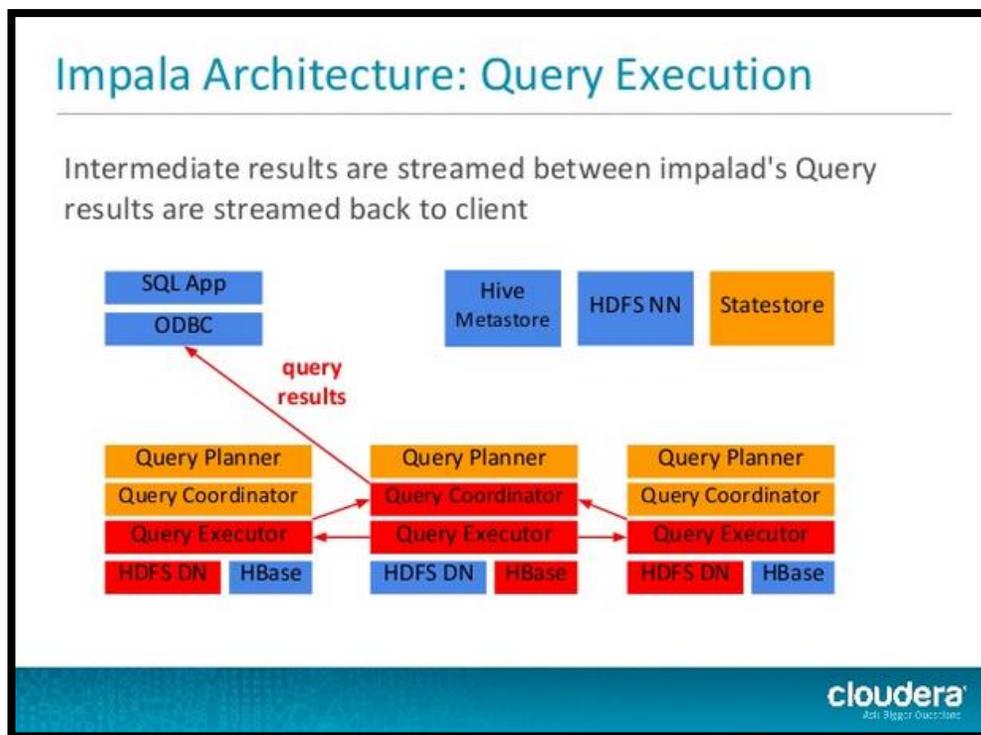
42. Plaintiffs reference and incorporate by reference the preceding paragraphs of this Complaint as if fully set forth herein.

43. Cloudera designs, makes, uses, sells, and/or offers for sale in the United States products and/or services for processing data within a database whereby tests based upon selected vectors are performed on each data record and the data records organized into clusters.

44. Cloudera designs, makes, sells, offers to sell, imports, and/or uses Cloudera Enterprise Data Hub and Cloudera Analytic DB (collectively, the “Cloudera ‘411 Product(s)’”).

45. On information and belief, one or more Cloudera subsidiaries and/or affiliates use the Cloudera ‘411 Products in regular business operations.

46. On information and belief, one or more of the Cloudera ‘411 Products are parallel data processing systems for search, storage, and retrieval of data of a database responsive to client queries for specific data from a database.

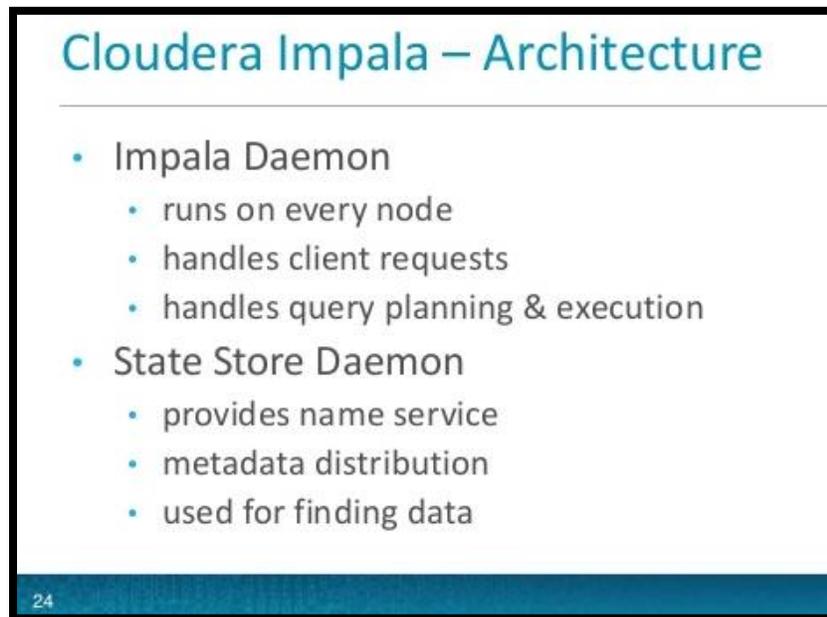


Mark Grover, *Cloudera Impala: A Modern SQL Engine for Apache Hadoop*, CLOUDERA PRESENTATION at 14 (February 27, 2013) (showing that intermediate results from a query being executed are at times streamed back to the requesting client).

47. On information and belief, the Cloudera ‘411 Products include technology comprising a plurality of host processors including a root host processor.

48. On information and belief, the Cloudera ‘411 Products include technology comprising a root host processor that is responsive to client queries for specific data of a

database wherein at least two host processors have a search engine and maintain information of a search queue of the client queries.



Maxime Dumas, *Cloudera Impala*, LAS VEGAS BIG DATA MONTHLY MEETUP at 24 (November 5, 2014).

49. On information and belief, the Cloudera '411 Products include technology comprising at least two host processors having a queue of search requests for specific data of a database, wherein each host processor executes a search engine, communicating capacity and load information between host processors and the at least two host processors exchanging at least one search request.

## Components of the Impala Server

The Impala server is a distributed, massively parallel processing (MPP) database engine. It consists of different daemon processes that run on specific hosts within your CDH cluster.

### The Impala Daemon

The core Impala component is a daemon process that runs on each node of the cluster, physically represented by the `impalad` process. It reads and writes to data files; accepts queries transmitted from the `impala-shell` command, Hue, JDBC, or ODBC; parallelizes the queries and distributes work to other nodes in the Impala cluster; and transmits intermediate query results back to the central coordinator node.

You can submit a query to the Impala daemon running on any node, and that node serves as the **coordinator node** for that query. The other nodes transmit partial results back to the coordinator, which constructs the final result set for a query. When running experiments with functionality through the `impala-shell` command, you might always connect to the same Impala daemon for convenience. For clusters running production workloads, you might load-balance between the nodes by submitting each query to a different Impala daemon in round-robin style, using the JDBC or ODBC interfaces.

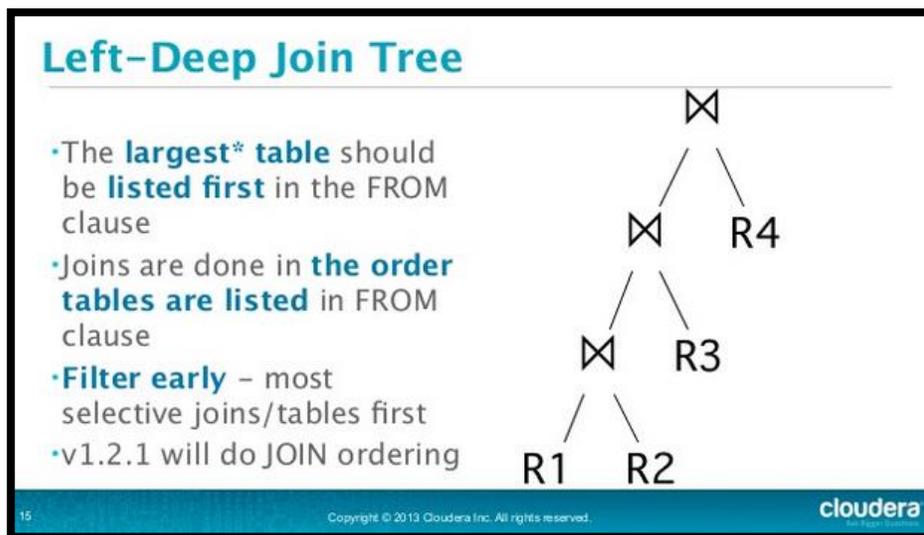
*Impala Concepts and Architecture*, CLUDERA CDH DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017) (describing how the Impala server is a distributed massively parallel processing MPP database engine that contains a daemon process that runs on each node of the cluster.).

50. On information and belief, the Cloudera '411 Products include search engines removing at least one search request from a search queue and generating an additional search request.

51. On information and belief, the Cloudera '411 Products include host and root host processors maintaining a list of available host processors and information about the capacity and load for each available host processor in memory and broadcasts its capacity and load information to the other host processors and bring its search queue into balance with another host processor according to a time constant in response to receipt of the broadcast capacity and load information.

52. On information and belief, the Cloudera '411 Products include a communications system coupling the host and root processors, wherein at least two host processors communicate capacity and load information to other host processors.

53. On information and belief, the Cloudera '411 Products include technology allowing selected host processors to store a database index for the database comprising nodes of a database tree for the database and data accessible via the nodes of the database tree.



Marcel Kornacker, *Practical Performance Analysis and Tuning for Cloudera Impala*, CLOUDERA PRESENTATION at 15 (November 11, 2011) (going on to describe that Impala enables parallel processing).

54. On information and belief, the Cloudera '411 Products are available to businesses and individuals throughout the United States.

55. On information and belief, the Cloudera '411 Products are provided to businesses and individuals located in the Eastern District of Tennessee.

56. On information and belief, Cloudera has directly infringed and continues to directly infringe the '411 patent by, among other things, making, using, offering for sale, and/or selling technology for extracting data from sources of network-based information, including but not limited to the Cloudera '411 Products, which include infringing technology for processing

data within a database whereby tests based upon selected vectors are performed on each data record and the data records organized into clusters. Such products and/or services include, by way of example and without limitation, the Cloudera '411 Products.

57. By making, using, testing, offering for sale, and/or selling products and services, including but not limited to the Cloudera '411 Products, Cloudera has injured Plaintiffs and is liable to Plaintiffs for directly infringing one or more claims of the '411 patent, including at least claim 1, pursuant to 35 U.S.C. § 271(a).

58. On information and belief, Cloudera also indirectly infringes the '411 patent by actively inducing infringement under 35 USC § 271(b).

59. On information and belief, Cloudera has had knowledge of the '411 patent since at least October 16, 2014. On October 16, 2014, the University of Tennessee Research Foundation contacted Cloudera via email in an attempt to secure a license to the patents-in-suit outside the context of litigation.

60. Alternatively, Cloudera has had knowledge of the '411 patent since at least service of this Complaint or shortly thereafter, and on information and belief, Cloudera knew of the '411 patent and knew of its infringement, including by way of this lawsuit.

61. On information and belief, Cloudera intended to induce patent infringement by third-party customers and users of the Cloudera '411 Products and had knowledge that the inducing acts would cause infringement or was willfully blind to the possibility that its inducing acts would cause infringement. Cloudera specifically intended and was aware that the normal and customary use of the accused products would infringe the '411 patent. Cloudera performed the acts that constitute induced infringement, and would induce actual infringement, with knowledge of the '411 patent and with the knowledge that the induced acts would constitute

infringement. For example, Cloudera provides the Cloudera ‘411 Products that have the capability of operating in a manner that infringe one or more of the claims of the ‘411 patent, including at least claim 1, and Cloudera further provides documentation and training materials that cause customers and end users of the Cloudera ‘411 Products to utilize the products in a manner that directly infringe one or more claims of the ‘411 patent.<sup>31</sup> By providing instruction and training to customers and end-users on how to use the Cloudera ‘411 Products in a manner that directly infringes one or more claims of the ‘411 patent, including at least claim 1, Cloudera specifically intended to induce infringement of the ‘411 patent. On information and belief, Cloudera engaged in such inducement to promote the sales of the Cloudera ‘411 Products, e.g., through Cloudera user manuals, product support, marketing materials, and training materials to actively induce the users of the accused products to infringe the ‘411 patent. Accordingly, Cloudera has induced and continues to induce users of the accused products to use the accused products in their ordinary and customary way to infringe the ‘411 patent, knowing that such use constitutes infringement of the ‘411 patent.

62. The ‘411 patent is well-known within the industry as demonstrated by multiple citations to the ‘411 patent in published patents and patent applications assigned to technology companies and academic institutions. Cloudera is utilizing the technology claimed in the ‘411 patent without paying a reasonable royalty. Cloudera is infringing the ‘411 patent in a manner

---

<sup>31</sup> See, e.g., Cloudera Impala Team, *The Impala Cookbook*, CLOUDERA PRESENTATION (Updated January 2017); Mark Grover, *Cloudera Impala: A Modern SQL Engine for Apache Hadoop*, CLOUDERA PRESENTATION (February 27, 2013); CLOUDERA ENTERPRISE DATA SHEET (2016); Marcel Kornacker, *Practical Performance Analysis and Tuning for Cloudera Impala*, CLOUDERA PRESENTATION (November 11, 2011); *Impala Concepts and Architecture*, CLOUDERA CDH DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017); and *Using Cloudera Impala: How Cloudera Impala Works with CDH*, CLOUDERA DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017).

best described as willful, wanton, malicious, in bad faith, deliberate, consciously wrongful, flagrant, or characteristic of a pirate.

63. To the extent applicable, the requirements of 35 U.S.C. § 287(a) have been met with respect to the '411 patent.

64. As a result of Cloudera's infringement of the '411 patent, Plaintiffs have suffered monetary damages, and seek recovery in an amount adequate to compensate for Cloudera's infringement, but in no event less than a reasonable royalty for the use made of the invention by Cloudera together with interest and costs as fixed by the Court.

**COUNT II**  
**INFRINGEMENT OF U.S. PATENT NO. 8,099,733**

65. Plaintiffs reference and incorporate by reference the preceding paragraphs of this Complaint as if fully set forth herein.

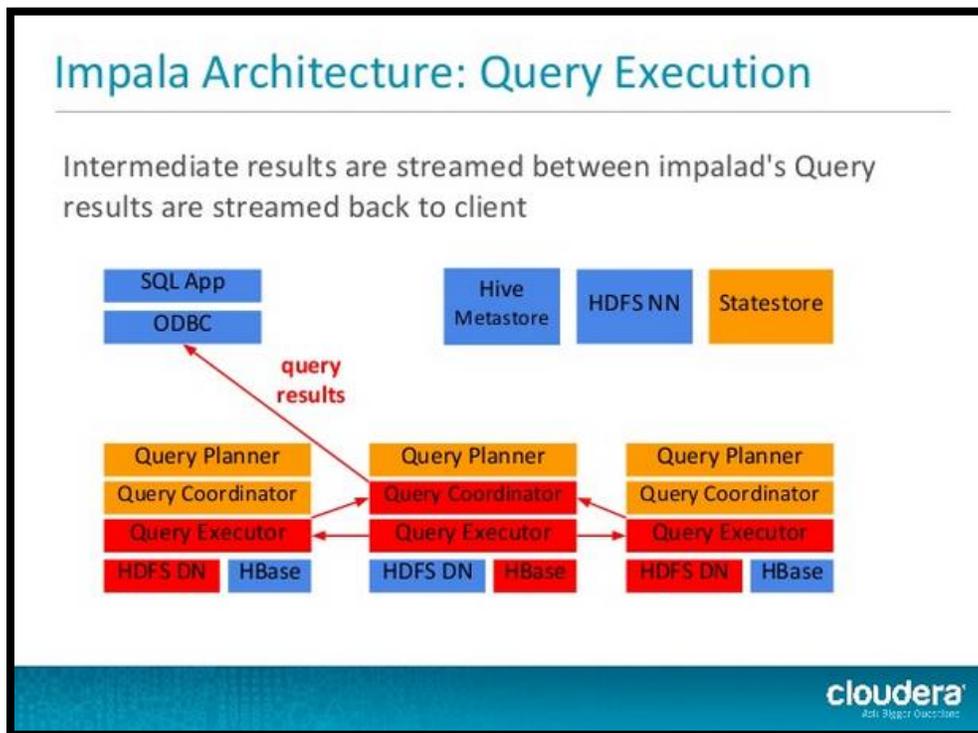
66. Cloudera designs, makes, uses, sells, and/or offers for sale in the United States products and/or services comprising a parallel data processing architecture for search, storage and retrieval of data of a database responsive to queries for specific data of a database where a plurality of host processors capable of communicating with one another and broadcasting load information (processor capacity and search queue length) and bringing client queries into balance with one another based on the receipt of the broadcasted capacity and load information.

67. Cloudera designs, makes, sells, offers to sell, imports, and/or uses Cloudera Enterprise Data Hub and Cloudera Analytic DB (collectively, the "Cloudera '733 Product(s)").

68. On information and belief, one or more Cloudera subsidiaries and/or affiliates use the Cloudera '733 Products in regular business operations.

69. On information and belief, one or more of the Cloudera '733 Products comprise a parallel data processing architecture for search, storage and retrieval of data of a database

responsive to queries for specific data of said database. Specifically, the Cloudera ‘733 Products comprise systems where there are a plurality of host processors and where at least one root host processor is responsive to a client query for data from a database.



Mark Grover, *Cloudera Impala: A Modern SQL Engine for Apache Hadoop*, CLOUDERA PRESENTATION at 14 (February 27, 2013) (showing that intermediate results from a query being executed are at times streamed back to the requesting client).

70. On information and belief, one or more of the Cloudera ‘733 Products comprise a plurality of host processors comprising at least one root host processor responsive to a client query for specific data of a database and at least one other host processor.

71. On information and belief, one or more of the Cloudera ‘733 Products comprise a communication system and method for coupling a plurality of host processors that communicate with one another their capacity and measurements of current loads.

72. On information and belief, one or more of the Cloudera ‘733 Products comprise a method for balancing workload between said plurality of host processors.

73. On information and belief, one or more of the Cloudera '733 Products maintain load information of processor capacity and search queue length for each host processor.

74. On information and belief, one or more of the Cloudera '733 Products maintain information of a search queue of client queries at the host processor for specific data within a database.

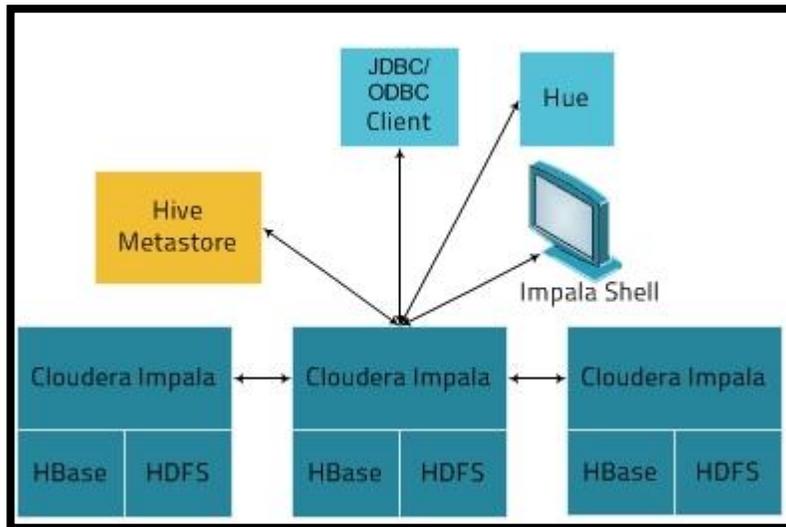
75. The Cloudera '733 Products are systems which contain functionality for a plurality of host processors maintaining load information of processor capacity and search queue length of a host processor. The Cloudera '733 Products enable each host to maintain information regarding both the capacity of the host and the current load on that host.

76. On information and belief, one or more of the Cloudera '733 Products enable each host processor to broadcast load information of its processor capacity and search queue length to at least one other of a plurality of host processors.

77. On information and belief, one or more of the Cloudera '733 Products enable each of the host processors to reconfigure information on available host processors responsive to the receipt of broadcast information of capacity and load information.

78. On information and belief, one or more of the Cloudera '733 Products enable each host processor to bring its search queue of client queries into balance with another of a plurality of host processors according to a time constant responsive to receipt of the broadcast capacity and load information.

79. On information and belief, one or more of the Cloudera '733 Products enable balancing search queues by exchanging unprocessed search requests with a recipient host processor responsive to a stochastic selection process to determine the recipient host processor of an exchanged search request between a root host processor and a recipient host processor.



*Using Cloudera Impala: How Cloudera Impala Works with CDH*, CLOUDERA DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017) (Stating that Impala provides: “Distributed queries in a cluster environment, for convenient scaling and to make use of cost-effective commodity hardware.”).

80. The Cloudera ‘733 Products balance exchanging unprocessed search requests with a recipient host processor responsive to a stochastic selection process to determine the recipient host processor of an exchanged search request between said root host processor and a recipient host processor.

### Fast for Business

Only Cloudera Enterprise enables more insights for more users, all within a single platform. With the most powerful open source tools and the only active data optimization designed for Hadoop, you can move from big data to results faster. Key features include:

- In-Memory Data Processing: The most experience with Apache Spark
- Fast Analytic SQL: The lowest latency and best concurrency for BI with Apache Impala
- Native Search: Complete user accessibility built-into the platform with Apache Solr
- Updateable Analytic Storage: The only Hadoop storage for fast analytics on fast changing data with Apache Kudu (beta)
- Active Data Optimization: Cloudera Navigator Optimizer (limited beta) helps tune data and workloads for peak performance with Hadoop

CLOUDERA ENTERPRISE DATA SHEET at 1-2 (2016) (stating that the “Impala” product is part of Cloudera Enterprise Data Hub Edition and Analytic Database (DB) Edition).

81. On information and belief, one or more of the Cloudera ‘733 Products enable exchanging a block of search requests between a plurality of host processors and adjusting a size of the block of exchanged requests according to relative processing speeds of host processors and inter-processor communications protocol between the host processors.

82. On information and belief, the Cloudera ‘733 Products are available to businesses and individuals throughout the United States.

83. On information and belief, the Cloudera ‘733 Products are provided to businesses and individuals located in the Eastern District of Tennessee.

84. On information and belief, Cloudera has directly infringed and continues to directly infringe the ‘733 patent by, among other things, making, using, offering for sale, and/or selling technology for extracting data from sources of network-based information, including but not limited to the Cloudera ‘733 Products, which include infringing technology comprising a parallel data processing architecture for search, storage, and retrieval of data of a database

responsive to queries for specific data of said database where a plurality of host processors are capable of communicating with one another and broadcasting load information (processor capacity and search queue length) and bringing client queries into balance with one another based on the receipt of the broadcasted capacity and load information. Such products and/or services include, by way of example and without limitation, the Cloudera '733 Products.

85. By making, using, testing, offering for sale, and/or selling products and services, including but not limited to the Cloudera '733 Products, Cloudera has injured Plaintiffs and is liable for directly infringing one or more claims of the '733 patent, including at least claim 18, pursuant to 35 U.S.C. § 271(a).

86. On information and belief, Cloudera also indirectly infringes the '733 patent by actively inducing infringement under 35 USC § 271(b).

87. On information and belief, Cloudera has had knowledge of the '733 patent since at least October 16, 2014. On October 16, 2014, the University of Tennessee Research Foundation contacted Cloudera via email in an attempt to secure a license to the patents-in-suit outside the context of litigation.

88. Alternatively, Cloudera has had knowledge of the '733 patent since at least service of this Complaint or shortly thereafter, and on information and belief, Cloudera knew of the '733 patent and knew of its infringement, including by way of this lawsuit.

89. On information and belief, Cloudera intended to induce patent infringement by third-party customers and users of the Cloudera '733 Products and had knowledge that the inducing acts would cause infringement or was willfully blind to the possibility that its inducing acts would cause infringement. Cloudera specifically intended and was aware that the normal and customary use of the accused products would infringe the '733 patent. Cloudera performed

the acts that constitute induced infringement, and would induce actual infringement, with knowledge of the '733 patent and with the knowledge that the induced acts would constitute infringement. For example, Cloudera provides the Cloudera '733 Products that have the capability of operating in a manner that infringe one or more of the claims of the '733 patent, including at least claim 18, and Cloudera further provides documentation and training materials that cause customers and end users of the Cloudera '733 Products to utilize the products in a manner that directly infringe one or more claims of the '733 patent.<sup>32</sup> By providing instruction and training to customers and end-users on how to use the Cloudera '733 Products in a manner that directly infringes one or more claims of the '733 patent, including at least claim 18, Cloudera specifically intended to induce infringement of the '733 patent. On information and belief, Cloudera engaged in such inducement to promote the sales of the Cloudera '733 Products, e.g., through Cloudera user manuals, product support, marketing materials, and training materials to actively induce the users of the accused products to infringe the '733 patent. Accordingly, Cloudera has induced and continues to induce users of the accused products to use the accused products in their ordinary and customary way to infringe the '733 patent, knowing that such use constitutes infringement of the '733 patent.

90. The '733 patent is well-known within the industry as demonstrated by the multiple citations to the '733 patent in published patents and patent applications assigned to

---

<sup>32</sup> See, e.g., Cloudera Impala Team, *The Impala Cookbook*, CLOUDERA PRESENTATION (Updated January 2017); Mark Grover, *Cloudera Impala: A Modern SQL Engine for Apache Hadoop*, CLOUDERA PRESENTATION (February 27, 2013); CLOUDERA ENTERPRISE DATA SHEET (2016); Marcel Kornacker, *Practical Performance Analysis and Tuning for Cloudera Impala*, CLOUDERA PRESENTATION (November 11, 2011); *Impala Concepts and Architecture*, CLOUDERA CDH DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017); and *Using Cloudera Impala: How Cloudera Impala Works with CDH*, CLOUDERA DOCUMENTATION, available at: <https://www.cloudera.com/documentation/cdh/5-0-x/Impala/> (last visited April 2017).

technology companies and academic institutions. Cloudera is utilizing the technology claimed in the '733 patent without paying a reasonable royalty. Cloudera is infringing the '733 patent in a manner best described as willful, wanton, malicious, in bad faith, deliberate, consciously wrongful, flagrant, or characteristic of a pirate.

91. To the extent applicable, the requirements of 35 U.S.C. § 287(a) have been met with respect to the '733 patent.

92. As a result of Cloudera's infringement of the '733 patent, Plaintiffs have suffered monetary damages, and seeks recovery in an amount adequate to compensate for Cloudera's infringement, but in no event less than a reasonable royalty for the use made of the invention by Cloudera together with interest and costs as fixed by the Court.

**PRAYER FOR RELIEF**

WHEREFORE, Plaintiffs respectfully requests that this Court enter:

- A. A judgment in favor of Plaintiffs that Cloudera has infringed, either literally and/or under the doctrine of equivalents, the '411 and '733 patents;
- B. An award of damages resulting from Cloudera's acts of infringement in accordance with 35 U.S.C. § 284;
- C. A judgment and order finding that Defendant's infringement was willful, wanton, malicious, bad-faith, deliberate, consciously wrongful, flagrant, or characteristic of a pirate within the meaning of 35 U.S.C. § 284 and awarding to Plaintiffs enhanced damages.
- D. A judgment and order finding that this is an exceptional case within the meaning of 35 U.S.C. § 285 and awarding to Plaintiffs their reasonable attorneys' fees against Defendant.
- E. Any and all other relief to which Plaintiffs may show themselves to be entitled.

**JURY TRIAL DEMANDED**

Pursuant to Rule 38 of the Federal Rules of Civil Procedure, University of Tennessee Research Foundation and Saint Matthew Research, LLC requests a trial by jury of any issues so triable by right.

Dated: May 3, 2017

Respectfully submitted,

s/Wayne A. Ritchie II

WAYNE A. RITCHIE II (BPR 013936)

JAMES R. STOVALL (BPR #032512)

RITCHIE, DILLARD, DAVIES

& JOHNSON, P.C.

606 West Main Avenue, Suite 300

P.O. Box 1126

Knoxville, Tennessee 37901-1126

(865) 637-0661

E-mail: [war@rddjlaw.com](mailto:war@rddjlaw.com)

E-mail: [jstovall@rddjlaw.com](mailto:jstovall@rddjlaw.com)

OF COUNSEL:

Dorian S. Berger (CA SB No. 264424)

*pro hac vice application forthcoming*

Daniel P. Hipskind (CA SB No. 266763)

*pro hac vice application forthcoming*

Eric B. Hanson (CA SB No. 254570)

*pro hac vice application forthcoming*

BERGER & HIPSKIND LLP

1880 Century Park East, Ste. 815

Los Angeles, CA 95047

Telephone: 323-886-3430

Facsimile: 323-978-5508

E-mail: [dsb@bergerhipskind.com](mailto:dsb@bergerhipskind.com)

E-mail: [dph@bergerhipskind.com](mailto:dph@bergerhipskind.com)

E-mail: [ebh@bergerhipskind.com](mailto:ebh@bergerhipskind.com)

*Attorneys for Plaintiffs University of  
Tennessee Research Foundation and Saint  
Matthew Research, LLC*